Monetary Policy under Uncertainty and Learning

Klaus Schmidt-Hebbel
Carl E. Walsh
editors

Central Bank of Chile / Banco Central de Chile

Monetary Policy under Uncertainty and Learning

Klaus Schmidt-Hebbel
Carl E. Walsh
Editors

Central Bank of Chile / Banco Central de Chile

Series on Central Banking, Analysis, and Economic Policies

The Book Series on "Central Banking, Analysis, and Economic Policies" of the Central Bank of Chile publishes new research on central banking and economics in general, with special emphasis on issues and fields that are relevant to economic policies in developing economies. The volumes are published in Spanish or English. Policy usefulness, high-quality research, and relevance to Chile and other economies are the main criteria for publishing books. Most research in this Series has been conducted in or sponsored by the Central Bank of Chile.

Book manuscripts are submitted to the Series editors for a review process with active participation by outside referees. The Series editors submit manuscripts for final approval to the Editorial Board of the Series and to the Board of the Central Bank of Chile. Publication is both in paper and electronic format.

The views and conclusions presented in the books are exclusively those of the authors and do not necessarily reflect the position of the Central Bank of Chile or of the Board members.

Monetary Policy under Uncertainty and Learning

Klaus Schmidt-Hebbel
Carl E. Walsh
Editors

Central Bank of Chile / Banco Central de Chile
Santiago, Chile

Agustinas 1180
Santiago, Chile

Published in Santiago, Chile by the Central Bank of Chile
Manufactured in Chile

ISBN 978-956-7421-32-9
Intellectual Property Registration # 183.782
ISSN # 0717-6686 (Series on Central Banking, Analysis, and Economic Policies)

Production Team

Editors:
Klaus Schmidt-Hebbel
Carl E. Walsh

Supervisors:
Mariana García
Fabián Gredig

Copy Editor:
Jennifer Hoover

Designer:
Mónica Widoycovich

Proof Reader:
Dionisio Vio

Technical Staff:
Carlos Arriagada
María Soledad Gallardo
Claudio Sepúlveda

Printer:
Andros Impresores

Contributors

The articles published in this volume are revised versions of the papers presented at the Eleventh Conference of the Central Bank of Chile on Monetary Policy under Uncertainty and Learning held in Santiago on November 15-16, 2007. The list of contributing authors and conference discussants follows.

Contributing Authors

Marco Del Negro
Federal Reserve Bank of New York
New York, NY, USA

George W. Evans
University of Oregon
Eugene, OR, USA
University of St Andrews
St Andrews, Scotland, United Kingdom

Roger Guesnerie
Collège de France and Paris School of Economics
Paris, France

Seppo Honkapohja
Bank of Finland
Helsinki, Finland
University of Cambridge
Cambridge, United Kingdom

Manuel Marfán
Central Bank of Chile
Santiago, Chile

Bennett T. McCallum
Carnegie Mellon University
Pittsburgh, PA, USA

Juan Pablo Medina
Central Bank of Chile
Santiago, Chile

Martin Melecký
World Bank
Washington, DC, USA

Felipe Morandé
Universidad de Chile
Santiago, Chile

Athanasios Orphanides
Central Bank of Cyprus
Nicosia, Cyprus

Ricardo Reis
Columbia University
New York, NY, USA

Diego Rodríguez Palenzuela
European Central Bank
Frankfurt, Germany

Klaus Schmidt-Hebbel
Pontificia Universidad Católica de Chile
Santiago, Chile

Frank Schorfheide
University of Pennsylvania
Philadelphia, PA, USA

Ulf Söderström
Sveriges Riksbank
Stockholm, Sweden

Claudio Soto
Central Bank of Chile
Santiago, Chile

Lars E.O. Svensson
Sveriges Riksbank
Stockholm, Sweden
Princeton University
Princeton, NJ, USA

Mauricio Tejada
Universidad de Chile
Santiago, Chile

Carl E. Walsh
University of California
at Santa Cruz
Santa Cruz, CA, USA

Volker Wieland
Goethe-Universität Frankfurt
Frankfurt Main, Germany

John C. Williams
Federal Reserve Bank
of San Francisco
San Francisco, CA, USA

Noah Williams
University of Wisconsin
Madison, WI, USA

Conference Discussants

Nicoletta Batini
International Monetary Fund
Washington, DC, USA

Claes Berg
Sveriges Riksbank
Stockholm, Sweden

Alex Bowen
London School of Economics
London, England

Rodrigo Cerda
Pontificia Universidad
Católica de Chile
Santiago, Chile

Rómulo Chumacero
Central Bank of Chile
Santiago, Chile

Sheryl Kennedy
Bank of Canada
Ottawa, Ontario, Canada

Christopher Kent
Reserve Bank of Australia
Sydney, Australia

John McDermott
Reserve Bank of New Zealand
Wellington, New Zealand

Juan Pablo Nicolini
Universidad Torcuato di Tella
Buenos Aires, Argentina

Claudio Soto
Central Bank of Chile
Santiago, Chile

Vicente Tuesta
Deutsche Bank
Lima, Peru

Rodrigo O. Valdés
International Monetary Fund
Washington, DC, USA

Table of Contents

Monetary Policy under Uncertainty and Learning:
An Overview
Klaus Schmidt-Hebbel and Carl E. Walsh 1

Expectations, Learning and Monetary Policy:
An Overview of Recent Research
George W. Evans and Seppo Honkapohja 27

Optimal Monetary Policy under Uncertainty in DSGE Models:
A Markov Jump-Linear-Quadratic Approach
Lars E.O. Svensson and Noah Williams 77

Imperfect Knowledge and the Pitfalls of Optimal Control
Monetary Policy
Athanasios Orphanides and John C. Williams 115

Robust Learning Stability
with Operational Monetary Policy Rules
George W. Evans and Seppo Honkapohja 145

Macroeconomic and Monetary Policies
from the Eductive Viewpoint
Roger Guesnerie 171

Determinacy, Learnability, and Plausibility
in Monetary Policy Analysis: Additional Results
Bennett T. McCallum 203

A Sticky-Information General Equilibrium Model
for Policy Analysis
Ricardo Reis 227

Monetary Policy and Key Unobservables: Evidence from Large
Industrial and Selected Inflation-Targeting Countries
Klaus Schmidt-Hebbel and Carl E. Walsh 285

Inflation Target Transparency and the Macroeconomy
*Martin Melecký, Diego Rodríguez Palenzuela
and Ulf Söderström* 371

Learning, Endogenous Indexation, and Disinflation in the New-Keynesian Model
Volker Wieland 413

Sources of Uncertainty in Conducting Monetary Policy in Chile
Felipe Morandé and Mauricio Tejada 451

Inflation Dynamics in a Small Open Economy Model under Inflation Targeting: Some Evidence from Chile
Marco Del Negro and Frank Schorfheide 511

Overoptimism, Boom-Bust Cycles and Monetary Policy in Small Open Economies
Manuel Marfán, Juan Pablo Medina and Claudio Soto 563

Monetary Policy under Uncertainty and Learning: An Overview

Klaus Schmidt-Hebbel
Pontificia Universidad Católica de Chile

Carl E. Walsh
University of California at Santa Cruz

Uncertainty is not just an important feature of the monetary policy landscape: it is the defining characteristic of that landscape.

—Alan Greenspan

Central bank economists and academic economists conducting research on the design of monetary policy have made significant advances in recent years. This work has led to a clearer understanding of the desirable properties of interest rate rules, the role of announcements and communication, and the consequences of inflation targeting for both inflation and the real economy. Dynamic stochastic general equilibrium (DSGE) models have been extended from the small-scale, often calibrated versions initially employed to address policy issues to much larger models that are estimated using Bayesian techniques. Many central banks now use these models for policy evaluation.[1] Much of this work neglects one of the key issues that policymaker face, however: the pervasive role of uncertainty. The recent global financial crisis and recession serve as the latest example of the policy challenges posed by unexpected and unforeseen events.

At the time of the conference, Klaus Schmidt-Hebbel was affiliated with the Central Bank of Chile.

1. See Galí (2008) for an excellent treatment of the basic New-Keynesian model that has become standard in monetary policy analysis. Examples of estimated DSGE models include Christiano, Eichenbaum, and Evans (2005), Levin and others (2006), Smets and Wouters (2003), Adolfson and others (2008), and Christiano, Motto, and Rostagno (2007).

Monetary Policy under Uncertainty and Learning, edited by Klaus Schmidt-Hebbel and Carl E. Walsh, Santiago, Chile.

The huge swings in oil, food, and other commodity prices in recent years and the dramatic global financial crisis have dominated discussions of monetary policy in the past year. These events provide vivid reminders of how uncertainty, imperfect knowledge of the economy, and the need to learn about new developments in world goods and financial markets affect the macroeconomy and influence the conduct of monetary policy. In this book, leading international scholars address many of the key issues relevant for central banks who must by necessity operate in environments of uncertainty and in which policymakers and the public are continually learning about the economy.

1. Uncertainty and Learning

In this section, we selectively review the literature on uncertainty and learning, focusing specifically on the insights that are important for the conduct of monetary policy. The next section then surveys the new research contained in this volume.

1.1 Types of Uncertainty and their Implications for Monetary Policy

Limitations of economic theory and data, structural changes in the economy, the inherent unobservability of important macroeconomic variables such as potential output and the neutral interest rate, and disagreements over the correct model of the economy and the transmission process of policy are just some of the reasons why central bankers operate in an environment of uncertainty. Research into the effects of uncertainty and the design of optimal policy in the face of uncertainty has broadly focused on three types of uncertainty: additive uncertainty, model uncertainty, and imperfect information.

To illustrate these different forms of uncertainty, suppose that the "true" model of the economy takes the form

$$\mathbf{y}(t+1) = \mathbf{A}\mathbf{y}(t) + \mathbf{B}\mathbf{y}(t \mid t) + \mathbf{C}i(t) + \mathbf{D}\mathbf{u}(t+1), \tag{1}$$

where $\mathbf{y}(t)$ is a vector of macroeconomic variables at time t, $\mathbf{y}(t \mid t)$ is the policymaker's current estimate of $\mathbf{y}(t)$, $i(t)$ is the central bank's instrument, $\mathbf{u}(t)$ is a vector of random, exogenous disturbances, and **A**, **B**, **C**, and **D** are matrices containing the parameters of the model. Most models used for monetary policy analysis can be represented by this linear structure.

Additive uncertainty is represented by the disturbances $\mathbf{u}(t+1)$: when setting its instrument at time t, the central bank does not know what future shocks $\mathbf{u}(t+1)$ will hit the economy. Model uncertainty arises because the central bank does not know the true parameters that characterize the model (that is, the values of $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$); parameter estimates are subject to error, and the policymaker may believe some parameters are zero when they are in fact nonzero. Finally, imperfect information arises because the actual value of $\mathbf{y}(t)$ may be unobserved or only observed with error as a result of measurement error or data lags; as a consequence, the policymaker's best estimate of $\mathbf{y}(t)$, $\mathbf{y}(t \mid t)$, may be wrong. Following Walsh (2003), we discuss each of these sources of uncertainty in turn.

1.1.1 Additive uncertainty

The most extensively studied form of uncertainty is that arising from additive errors to the model's structural equations. In terms of the notation in equation (1), additive uncertainty is represented by $\mathbf{Du}(t+1)$. At the time the central bank must make its policy choice, the value of this term is unknown. Uncertainty about the realized values that $\mathbf{Du}(t+1)$ will take is the only form of uncertainty that typically is included in most models. Modern DSGE models often include random disturbances that enter the equilibrium conditions in nonlinear ways, but these models are then linearized, so that disturbances appear as additive error terms.

The problem of characterizing optimal policy in the face of additive uncertainty is well understood when the policymaker's objectives can be expressed as a quadratic function of various target variables. The standard assumption that central banks desire to minimize the volatility of inflation around its target and real output around potential output lends itself naturally to a representation in terms of a quadratic loss function in which squared deviations of inflation from the target and real output from potential output are penalized. The combination of linear, additive disturbances and quadratic objectives satisfies the well-known principle of certainty equivalence—all that matters for optimal policy are the expected values of the unknowns. Simply replace unknown disturbances with one's best forecast of their values and then treat these forecasts as if they were known with certainty. Thus, again in terms of equation (1), the central bank would replace $\mathbf{Du}(t+1)$ with its expected value,

$\mathbf{D}E\mathbf{u}(t+1)$ and then choose policy as if the true model were known with certainty to be

$$\mathbf{y}(t+1) = \mathbf{A}\mathbf{y}(t) + \mathbf{B}\mathbf{y}(t \mid t) + \mathbf{C}i(t) + \mathbf{D}E\mathbf{u}(t+1). \tag{2}$$

In this case, optimal policy does not require knowledge of the variances of the disturbances or the covariances among the different disturbances. This does not mean that *only* the expected value of the disturbance is relevant. Policymakers will usually need to forecast future values of these exogenous disturbances, and this will require some knowledge of, or at least assumptions about, the persistence of shocks. For example, a forecast that the price of oil will rise is generally not sufficient; the policymaker will need to forecast whether the rise is temporary or whether it is likely to be persistent.

To deal with additive uncertainty, Giannoni and Woodford (2002) propose optimal policies, which they call robustly optimal policies. Robustly optimal policy rules describe how the policy instrument should be set solely in terms of the macroeconomic variables that define the central bank's objective. If the central bank is concerned about maintaining low and stable inflation, stabilizing a measure of output relative to potential (the output gap), and stabilizing interest rate volatility, then the robustly optimal policy rule would show how the policy interest rate should be set as a function of inflation, the output gap, and lagged interest rates. Thus, implementing such a policy does not require information about the time series properties of the exogenous disturbances. Such a property is desirable, as it may be difficult to accurately forecast the degree of persistence in exogenous economic disturbances.

When the central bank is concerned with inflation and output gap stability, the optimal rule can be defined solely in terms of inflation and the output gap. In fact, the optimal policy can be characterized simply, as follows: keep a specific linear combination of inflation (relative to target) and the output equal to zero; if inflation is above target, then the output gap should be negative. The Bank of Norway, for example, describes the desirable properties of an interest rate path as one that ensures that the output gap is negative if the inflation gap (that is, inflation relative to the target) is positive. Adjusting the policy interest rate to maintain this sort of relationship between inflation and the output gap is often called a targeting rule, as it only involves the variables that are directly part of the central bank's objectives.

Unfortunately, robustly optimal policy rules generally require the central bank to make forecasts of inflation and the output gap.

Because monetary policy affects the economy with significant lags, policy must be forward looking, and this forces the central bank to rely on forecasts. To form forecasts of future inflation or real economic activity, however, the policymaker will need to decide whether a shock such as an oil price increase is temporary and will be reversed or is permanent. Thus, robustly optimal rules do not actually eliminate the need to forecast future disturbances.

In contrast to a robustly optimal rule, central bank behavior is often represented by simple instrument rules such as a Taylor rule. These rules typically assume that monetary policy is adjusted systematically in response to current movements in inflation and measures of the output gap. Other variables, such as the exchange rate, are sometimes also included. Given a specification of the central bank's objective, the coefficients in the rule can be chosen optimally. In contrast to fully optimal rules such as Giannoni and Woodford's robustly optimal rules, the best response coefficients in simple Taylor-type rules will depend on the relative variances of the basic disturbances that affect the economy. Designing the optimal "simple" rule thus requires a great deal of information about the additive shocks that hit the economy.

1.1.2 Model uncertainty

Model uncertainty encompasses a wide range of potential sources of error. Model misspecification, parameter uncertainty, and estimation error can all be grouped under this heading. Uncertainty about the values of the coefficient matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ is one reflection of model uncertainty. This uncertainty may arise because the central bank does not know the true values of the parameters in the model and must estimate them, or it could stem from the fact that the central bank's model incorporates incorrect assumptions about how the macroeconomic variables are related. Moreover, the true model may be evolving over time in unknown ways as a result of technological changes and innovations.

To illustrate how model uncertain affects the policy problem, suppose that we can ignore imperfect information, so that $\mathbf{y}(t) = \mathbf{y}(t \mid t)$. Let $\mathbf{A} + \mathbf{B} = \mathbf{H}$, and to keep the example simple, assume only elements of $\mathbf{A}$ and $\mathbf{B}$ are not known with certainty. The model then becomes

$$\mathbf{y}(t+1) = \hat{\mathbf{H}}\mathbf{y}(t) + \mathbf{C}i(t) + \mathbf{v}(t+1), \tag{3}$$

where $\mathbf{v}(t+1) = \mathbf{Du}(t+1) + (\mathbf{H} - \hat{\mathbf{H}})\mathbf{y}(t)$ and $\hat{\mathbf{H}}$ is the central bank's estimate of $\mathbf{H}$. Errors in estimating $\mathbf{H}$ now become part of the equation's error term, but the key difference from the case of additive uncertainty is that the errors represented by $\mathbf{v}(t+1)$ are now correlated with the endogenous variables $\mathbf{y}(t)$. The disturbance terms are no longer exogenous; misspecification is correlated with macroeconomic outcomes. This has important implications for policy choice, as first pointed out by Brainard (1967).

The type of uncertainty represented in equation (3) is called multiplicative uncertainty, since the uncertainty associated with the parameters in $\mathbf{H}$ multiply the endogenous variables. In the example he considered, Brainard (1967) showed that multiplicative uncertainty would make optimal policy less activist. Alan Blinder famously characterized the first step in a preemptive policy for controlling inflation as requiring the central bank to "estimate how much you need to tighten or loosen policy to 'get it right,' then do less" (Blinder, 1998, p. 17). This statement accurately reflected the caution that Brainard found to be appropriate in the face of multiplicative uncertainty.

In research subsequent to the work of Brainard, it was found that caution is not necessarily the best response to model uncertainty (Craine, 1979; Giannoni, 2002; Söderström, 2002). In fact, some forms of multiplicative uncertainty call for a more robust response than otherwise. For example, this may be the case when the uncertainty involves the dynamic response of the economy to shocks. If the central bank is uncertain about the degree to which current inflation may influence future inflation, it may be best to respond strongly to ensure that current inflation remains stable. Thus, an aggressive policy rather than a cautious one may be the best policy. In general, economists have found that there are no clear guidelines about how best to react when faced with this type of uncertainty.

Multiplicative uncertainty is certainly not the only, or even the most important, form of model uncertainty. More commonly, there are competing models for how the economy operates and how monetary policy affects macroeconomic activity and inflation. Within current macroeconomic circles, there are economists who employ models in which monetary policy can have important short-run real effects because of sticky prices and wages and other economists who use models in which monetary policy is impotent in affecting the real economy because all wages and prices are flexible. Faced with these competing models in an environment in which no one knows the true model of the economy, how should policymakers behave?

Clearly, policy is unlikely to contribute to macroeconomic stability if policymakers hold beliefs about the economy that are wrong. Romer and Romer (2002) attribute policy mistakes in the United States in the late 1960s and the 1970s to the use of a wrong model. Specifically, they argue that policymakers in the 1960s believed there was a permanent tradeoff between average unemployment and average inflation, and this led to the onset of the Great Inflation in the United States. Romer and Romer then argue that once inflation had reached high levels, policymakers came to believe that inflation was insensitive to recessions, implying that the cost of reducing inflation would be extremely high. Inflation was therefore allowed to rise, and policymakers delayed reducing it because they based their decisions on models that we now view as incorrect.

The example of model uncertainty provided by equation (3) shows how errors in the central bank's estimate of the parameters in $\mathbf{H}$ would interact with the endogenous variables represented by $\mathbf{y}(t)$. However, if $\mathbf{H} - \hat{\mathbf{H}}$ reflects estimation error or purely random fluctuations in the elements of $\mathbf{H}$, it might not be systematically related to economic developments. Hansen and Sargent (2003, 2004) study optimal policy in environments where the model uncertainty faced by the policymaker is not exogenous, but is designed to be particularly troublesome. They consider the case in which the policymaker fears that model misspecification will yield what, from the policymaker's perspective, is the worst possible outcome. In this environment, the policymaker seeks policies that are robust in the sense that they lead to reasonable outcomes even in the worst-case scenario. In the context of a simple monetary policy problem, Walsh (2004) shows that the worst-case scenario for the central bank involves the occurrence of a positive inflation shock when the economy is already in a recession. Such a scenario pushes the economy further away from the objectives of both low inflation and full employment.

Optimal policy in the face of this malicious misspecification turns out to require the central bank to employ a model of the economy that is deliberately distorted, in the sense that the central bank should assume that inflation shocks will be much more persistent than they are actually expected to be. Thus, in contrast to Gianonni and Woodford (2002), who designed policy rules that do not require the central bank to actually know (or even estimate) the true persistence of inflation shocks, Hansen and Sargent's approach has the central bank behave as if inflation shocks were always very persistent, even if they generally are not.

Worst-case scenarios are, almost by definition, events that occur with low probability, and the Hansen-Sargent approach has been criticized for putting too much weight on the worst-case scenario in policy decisions. However, the idea that a policymaker might want to use a distorted model when designing policy is supported by other lines of research. For example, Levin and Williams (2003) consider what happens when a policy is designed to be optimal for a particular model, but that model turns out to be wrong. They find that policy rules designed to be optimal in models that display high levels of inertia also perform well if the "true" model of the economy is very forward looking. Unfortunately, they find the converse not to be true. Policies designed to do well if forward-looking behavior is important often perform disastrously if the actual economy displays high degrees of inertia. Hence, even if the central bank believes that inflation and real economic activity are heavily influenced by expectations of future inflation and growth, it might still want to act as if the economy were much more backward looking.

In practice, central banks often deal with model uncertainty by employing several models of the economy, using the different models to cross-check forecasts and to ensure that policies are not excessively sensitive to assuming that a particular model is correct. Given competing models of the economy, a sensible approach might be to evaluate alternative polices in several models and to weight the different models based on an assessment of their likelihood. However, Cogley, Colacito, and Sargent (2007) illustrate how model uncertainty can lead to bad policies even when the policymaker is carefully trying to account for the uncertainty by using multiple models to evaluate policies. They consider two simple models. One model, labeled the Samuelson-Solow model, implies that the central bank faces a tradeoff between average unemployment and average inflation. The other incorporates the natural rate hypothesis, implying no tradeoff between average inflation and unemployment. This second model also implies that a credible disinflation would reduce inflation costlessly. The policymaker assigns probabilities to each model, reflecting the likelihood the data assign to each model being the true model. Cogley, Colacito, and Sargent show that by the early 1970s, U.S. data implied that almost all weight should be placed on the natural rate model. This meant that the optimal policy would be to immediately bring down inflation. However, the data still assigned a small but positive probability that the Samuelson-Solow model might be correct, and if that model turned out to be true, the output costs of an immediate disinflation would be enormous. So even though the central bank is almost certain the natural rate model is correct, it fails to reduce inflation out of fear that the Samuelson-Solow model might be correct. Thus, even a model that the data suggest is unlikely to be true can affect policy choices when the policymaker employs several models as a means of seeking robust policies.

1.1.3 Imperfect information

A final type of uncertainty arises from imperfect information. Just about any form of uncertainty could be labeled as being due to imperfect information (about the realizations of the additive disturbances, about the true model, and so on). However, we use the term to refer to a specific aspect of uncertainty—namely, that stemming from the inability to perfectly observe the current state of the economy or macroeconomic variables that are critical for policy design.

Policy decisions are made based on noisy and imperfect data about the economy. A number of authors investigate how data uncertainty affects optimal policy. Intuitively, one would expect that the presence of noise in macroeconomic data would call for responding less strongly to new data. Responding too strongly might simply introduce volatility if the signal-to-noise ratio is small, that is, if much of the variation in the data is simply noise. Rudebusch (2001) explores how data noise would reduce the optimal responses to inflation and the output gap in a standard Taylor rule. Earlier work that ignored data uncertainty found that the optimal response to the output gap was much larger than Taylor found for the Federal Reserve under Alan Greenspan. Rudebusch attributed part of the weaker response found in the data to the presence of noise in measures of the output gap.

Besides the issue of pure measurement error in real time data on observable variables, a further difficulty arises from the fact that many of the variables that play critical roles in theoretical models are not directly observed. The output gap is the best example of this problem. New-Keynesian models define the output gap as the percentage difference between actual output and the output the economy would produce if all wages and prices were flexible, the so-called flexible-price output level. While data on actual output is subject to measurement error and data revisions, it is at least directly measurable. The same cannot be said of the flexible-price output level. Any estimate of the latter will be dependent on a particular theoretical model of how the economy would behave with flexible prices. Older definitions of the output gap that measured output relative to potential output suffered from similar problems. Potential output is not observed but must be estimated, and standard techniques typically relied on simple statistical methods to equate potential output with trend output. This left open the issue of how best to estimate the trend growth rate of real output.

Measures of trend output are inevitably backward looking. They use historical data to estimate trends, so they are likely to have difficulty picking up shifts in underlying growth trends. A case in point was the 1970s, when many countries experienced a decline in trend growth. Orphanides (2003) argues that bad macroeconomic policies in the 1970s in the United States resulted from the failure to recognize this decline in the trend rate of growth. Because it based its estimate of trend growth on historical data, the Federal Reserve was slow to pick up the decline in the growth rate, and it thus overestimated the path of trend output in the 1970s. As a consequence of overestimating trend output, the Federal Reserve believed a negative output gap was opening up. It therefore adopted policies that, in retrospect, were too expansionary. This data-uncertainty hypothesis represents an alternative explanation for the Great Inflation of the 1970s to the interpretation based on the model-uncertainty hypothesis.

Given the difficulties involved in measuring the output gap, McCallum (2001) argues that central banks should not react to it strongly. Alternatively, Orphanides and Williams (2002) find that policy rules that respond to the change in the estimated output gap often perform well and avoid some of the measurement problems that make it difficult to estimate the level of potential output.

Problems with estimating the output gap are only one example of how key variables that modern economic theory suggests should be central to monetary policy are difficult to estimate and may even be unobservable. Another example is the neutral real interest rate, defined as the real interest rate consistent with a zero output gap and a zero deviation of inflation from target. Some modern models imply that the actual real interest rate should move in parallel with this neutral real rate, but the neutral real rate is ultimately unobservable. Several authors attempt to estimate the neutral real rate and the output gap (see Kuttner, 1994; Laubach and Williams, 2003; Garnier and Wilhelmsen, 2005; Benati and Vitale, 2007), but such estimates generally rely on restrictions implied by a particular model of the economy. If policymakers are uncertain about the correct model, they will also be uncertain about how best to measure the neutral real rate and the output gap. Imperfect information is thus a major problem facing policymakers.

1.2 Learning

The uncertainty faced by central banks largely reflects our imperfect understanding of macroeconomics. Economists and policymakers are constantly engaged in a process of learning about the economy. Similarly, members of the public are forming expectations based on their evolving understanding of the economy and the policymaker's behavior. Consequently, learning is pervasive—models are constantly refined and reestimated, new models are developed to reflect the latest progress in economic research, and previously ignored factors suddenly become important and must be incorporated into policy models. At the same time, the public must assess policy decisions and attempt to learn about the central banks' objectives and the way policy is being carried out. In recent years, a large literature has developed that investigates the effects of learning on macroeconomic outcomes and its implications for monetary policy.

Much of the work on learning in macroeconomics is based on the seminal work of Evans and Honkapohja (2001). Evans and Honkapohja (in this volume) provide an excellent overview of this research and its implications for monetary policy. The literature they survey drops the extreme informational assumptions implicit in the rational expectations approach. Instead, individuals (and policymakers) are viewed essentially as econometricians, using the latest data to reestimate and update their models and then using these models to make forecasts of future inflation and other macroeconomic variables. Evans and Honkapohja argue that this view of learning reflects the principle of cognitive consistency, in that it assumes private "agents should be about as smart as (good) economists" (in this volume, page 67). Explicitly incorporating learning allows the authors to study two general issues of relevance for policy. First, will the economy under learning converge to the equilibrium consistent with rational expectations? And second, how are macroeconomic dynamics affected by learning? If rational expectations equilibria are not stable under learning—a property called E-stability or learnability—then the properties of the rational expectations equilibrium becomes irrelevant for describing the economy's behavior once the economy's structure is understood. The standard practice in policy analysis is to study the properties of alternative policies under the assumption that the private sector fully understands how the central bank is behaving. This may be an appropriate assumption in terms of the eventual behavior of the economy, but only if the public eventually learns the true structure of the economy. If the public gradually learns about the different policies the central bank might follow, then the economy may not converge to the rational expectations equilibrium.

As Evans and Honkapohja (in this volume) discuss in their overview chapter, some policy rules for the central bank that appear to be quite reasonable rules under rational expectations can lead to instability under quite reasonable models of learning. However, Bullard and Mitra (2002) show that when the central bank follows a simple Taylor rule for setting the nominal interest rate, the same condition that ensures a unique equilibrium under rational expectations also ensures that the equilibrium is stable under learning. This condition, called the Taylor Principle, requires the central bank to adjust the nominal rate more than one-to-one with inflation.[2] Bullard and Mitra also show that if the central bank responds to expected future inflation rather than current inflation, some policy rules that lead to indeterminacy (multiple equilibria) under rational expectations have equilibria that are stable under learning. In general, Evans and Honkapohja argue that expectations-based policy rules—that is, rules in which the central bank responds to the private sector's inflation expectations and the output gap—have desirable properties. These rules implicitly incorporate the public's learning into the policy rule.

The second broad arena in which the learning literature has contributed to our understanding is macroeconomic dynamics. The manner in which the economy evolves will depend on the way the public learns, and the economy's response to disturbances can differ significantly under learning versus under rational expectations. Incorporating the effects of learning can be particularly important if the central bank is considering changing its policy behavior. The private sector's attempts to learn the new policy can affect the economy's adjustment if the central bank is not explicit or transparent about its policy. For example, Erceg and Levin (2003) study the role of learning in accounting for the steep recessions in the United States associated with the Volcker disinflation of the early 1980s. Under rational expectations, an announced reduction

2. This condition is weakened slightly if the central bank also responds to the output gap.

in the Fed's inflation target should have lowered inflation with little loss in real output. Erceg and Levin show that they can best match the historical experience of a gradual disinflation accompanied by recession when they assume that the Fed's anti-inflation stance lacked credibility and the public engaged in a process of learning about the Fed's target.

The learning literature has also developed new insights that are relevant for the debate over the optimal degree of central bank transparency. In general, greater transparency helps speed learning by providing useful information to the public. In that way, transparency can reduce the volatility that can occur when the public is trying to learn the central bank's objectives. Transparency can also ensure that the economy converges more quickly to the rational expectations equilibrium (Rudebusch and Williams, 2008). Incorporating learning is also relevant for ensuring that policies are robust when private agents and the policymaker may have evolving beliefs about the economy, as in Orphanides and Williams (in this volume).

Perhaps the most important lesson from the learning literature is that in a world of uncertainty and change, both private economic agents and the central bank engage in learning, and this process of learning cannot be ignored when designing policies to ensure determinacy, stability, and robustness.

1.3 Summary

Central banks must make policy decisions in the face of uncertainty based on imperfect and evolving knowledge about the economy. While few general results have emerged from the research on monetary policy in the face of uncertainty and learning, a key lesson is that neither uncertainty nor learning can be ignored. Policymakers must recognize that situations in which the uncertainty associated with forecasts can be ignored—that is, when certainty equivalence holds—are unlikely to hold in practice. Accounting for the role of multiple models and seeking policies that are robust across a range of plausible models is important. Seeking robustness may require using models that are distorted in ways that capture if not the worst-case scenarios, at least the more threatening ones. It is critical to recognize the role of data uncertainty, measurement error, and unobservability of key macroeconomic variables in designing and implementing monetary policy. Finally, policymakers must also account for the way policy actions affect the ability of the private sector to learn and the fact that the process of learning itself will influence the impact policy has on inflation and the real economy.

2. Overview of the Book

The essays in this volume offer both theoretical insight and practical guidance to evaluating monetary policy in the presence of uncertainty and the need to learn. The papers address a number of general questions. Are there practical means for calculating optimal policies in the face of very general specifications of model uncertainty? Does model uncertainty limit the usefulness of optimal control techniques? What types of monetary policy rules ensure stability when private agents employ constant-gain learning strategies? How do alternative notions of learning affect the stability of forward-looking models? How are the costs of disinflations affected by the credibility of the central bank's inflation target and the need for the public to engage in learning? How might disinflations affect the structure of the inflation process as private firms update their beliefs about the behavior of inflation, and do these effects alter the relative costs and benefits of announcing a gradual reduction in inflation targets? Are there general rules for formulating models and policy rules that ensure stability when private agents only have lagged data available? Can alternative models, useful for policy analysis, be developed if the effects of monetary policy arise from sticky information rather than sticky wages and prices? Is it possible to estimate unobservable variables that are key for monetary policy decisions using a simple model applied to different countries—and what does it reveal about international comovement and convergence of the unobservables and their observable counterparts?

The volume also addresses a number of issues specific to Chile's monetary policy. Did Chile's gradual disinflation experience based on annual targets in 1991–2000 contribute to lower costs of disinflation? How empirically important are additive, model, and information uncertainty? How sensitive is monetary policy to the laws of motion of exogenous shocks and to model misspecification? Finally, how sensitive are boom-bust cycles in Chile to alternative monetary policy rules?

The rest of this section briefly summarizes the chapters in the book, exploring how they answer the above set of questions. The second chapter in the volume, by George Evans and Seppo Honkapohja, provides an overview of the lessons for monetary policy derived from the growing literature on learning. Evans and Honkapohja have been the leading figures in developing and applying the notions of adaptive learning to macroeconomic issues. Their work is partly motivated by the idea that economic agents have neither the information nor the information-processing capabilities implicitly assumed by rational expectations approaches. Instead, economists should recognize that individuals are boundedly rational. One means of operationalizing this notion of bounded rationality is to assume that individuals learn adaptively. As the authors note, adaptive learning reflects the way economists typically learn about the empirical structure of the economy—they use new data to update their estimates of the economy's structural relationships or their forecasting equations. Applying this notion of learning to the private sector provides a tractable means of investigating a number of policy-relevant issues without imposing the extreme informational assumptions common to rational expectations models. Using the basic forward-looking New-Keynesian model that has become standard in the literature on monetary policy, the authors discuss a number of policy-related issues such as determinacy and E-stability under alternative policy rules, imperfect information on current variables, imperfect knowledge of structural parameters, and alternative models of adaptive learning. They also study the implications of learning for understanding hyperinflations and liquidity trap environments.

In their chapter, Lars E.O. Svensson and Noah Williams use a benchmark New-Keynesian model to show how policy is affected by the model uncertainty policymakers face. The authors have developed a new methodology for designing optimal monetary policies in the face of model uncertainty. This approach models uncertainty as reflected in shifts in the structural equations that characterize the economy. They represent the economy as jumping randomly between various states. Conditional on each state, the structure of the economy can be described in terms of linear equations and quadratic preferences. The approach is thus called a Markov jump-linear-quadratic model. As the authors argue, this approach can be used to model many types of uncertainty. They also discuss the role of learning, since they assume that the current state of the economy is not observable. The fully optimal policy in their framework will involve some experimentation—that is, deliberate policy actions designed to help the central bank better understand the behavior of the economy. Such policies are difficult to calculate, so Svensson and Williams focus on what they label adaptive optimal policies (AOP). Under these policies, the central bank does not consciously experiment. Svensson and Williams find that the gains from experimentation are typically small, a finding consistent with the reluctance of central banks to experiment with the macroeconomy. To illustrate the applicability of their approach to uncertainty, they employ a small, New-Keynesian model that was originally estimated using U.S. data by Lindé (2005). Using this model, the authors compare the AOP policy with optimal policy without learning, that is, when the central bank does not use the new data it receives to update its knowledge about the economy. Besides illustrating the algorithms they have developed to calculate AOP policies, the paper draws a very important policy conclusion: while learning is important for improving the design of policy in the face of uncertainty, the gains from experimentation are small.

Athanasios Orphanides and John Williams study the implications of alternative policies in the face of uncertainty and learning. They employ a small model estimated using U.S. data, but in evaluating monetary policies, they assume that the central bank must estimate key macroeconomic variables such as the natural rate of unemployment and the equilibrium real interest rate. Private agents are also uncertain about the structure of the model and employ least squares learning to update their beliefs about the economy. The authors show that ignoring uncertainty and learning can be costly in this environment: policies that are optimal when uncertainty is ignored lead to poor macroeconomic outcomes when knowledge is imperfect. Policies that are more robust to imperfect knowledge can be obtained if the central bank acts more conservatively, in the sense of placing greater weight on inflation objectives relative to stabilizing real economic activity. Interestingly, Orphanides and Williams show that simple policy rules that respond to expected future inflation and either lagged unemployment or the change in the unemployment rate perform well in the face of imperfect knowledge.

George Evans and Seppo Honkapohja examine the behavior of monetary policy rules when the private sector is engaged in learning. A huge literature examines the implications of simple policy rules, but this work generally assumes that private agents are fully aware of the rule the central bank is following. If, instead, members of the private sector must learn about the central bank's behavior, some important new issues arise. One issue relates to the stability of policy rules under different assumptions about the way private agents learn. The standard assumption in the literature on adaptive learning is that as agents obtain more observations, they place less weight on each one, a learning process known as decreasing gain. An alternative assumption is that agents use constant-gain least-squares learning, in which the weight on new information does not decrease as more observations are accumulated. Constant-gain learning may be appropriate when structural shifts might occur, making observations from the distant past less informative. Evans and Honkapohja show that some rules that perform well under decreasing-gain learning lead to expectational instability under constant-gain learning. Thus, not only is the fact that the private sector is learning important, but how they learn is also relevant. Finally, the authors show that what they describe as expectations-based optimal policy rules, in which the central bank responds to private sector expectations, have desirable properties.

Roger Guesnerie considers an approach to learning that differs from the adaptive learning models that have become common in monetary policy analysis. Under adaptive learning, individuals behave much like econometricians, using new observations on macroeconomic conditions to update their estimates of key economic relationships. In contrast to this approach, Guesnerie develops the concept of eductive stability. Intuitively, an eductively stable system has the property that if it is common knowledge that the economy is within some neighborhood of the equilibrium, then individuals behave in such a way that the actual equilibrium is within this neighborhood, regardless of their specific beliefs. Eductive stability can then be thought of as a property of an equilibrium such that, if the economic agents' beliefs are in some region, they will remain within that region under a broad set of updating rules. Eductive stability can thus be viewed as a necessary condition for any adaptive learning procedure to be stable. Applying the notion of eductive stability to a simple, cashless forward-looking model, Guesnerie finds that Taylor rules that react too strongly to inflation may not be eductively stable.

Bennett T. McCallum argues that the requirement of stability under least-squares learning is a "compelling necessary condition for a rational expectations equilibrium to be considered plausible." While previous work by McCallum and others demonstrates that monetary policy rules that ensure a unique rational expectations equilibrium (that is, that ensure determinacy) are least-squares learnable, this result is based on the assumption that individuals are able to observe the current equilibrium for the economy. More realistically, individuals may only observe lagged data on the economy, and in this case, the close connection between determinacy and learnability no longer holds. In fact, learnability is ensured only under additional, special assumptions. McCallum also explores the requirement that models be well formulated, where this is interpreted to mean that certain discontinuities in the models' steady state are ruled out. He shows that even when individuals observe current endogenous variables, a well-formulated model does not imply learnability (and vice versa).

Most modern models used for monetary policy analysis assume that nominal prices and wages are sticky, adjusting only slowly over time. In a series of previous papers, Ricardo Reis develops the idea that the economy may be characterized not by sticky prices, but by sticky information. Agents are inattentive to news because they incur costs of acquiring, absorbing, and processing information. In this volume, Reis presents a DSGE model of business cycles and monetary policy, where the only rigidity is pervasive inattention in all markets and where different agents update their information at different dates. The model is estimated on data for the post-1986 United States and the post-1993 euro area and then applied to conduct several counterfactual policy experiments for both regions. Monetary policy shocks have exhibited little persistence, implying a quick response of most macroeconomic variables to monetary shocks. Announcing a policy change in advance increases the response of inflation in comparison with unannounced changes. A gradual policy change has a stronger impact than an expected nongradual change, but only if the gradualist policy is announced and credible. Taylor's (1993) aggressively anti-inflation policy rule would yield higher welfare levels than what is attained by using the actual policy rules estimated for both regions. Finally, compared with flexible inflation targeting under a conventional Taylor rule, welfare would be reduced in both regions if their central banks were to adopt either strict or flexible price-level targeting.

Klaus Schmidt-Hebbel and Carl E. Walsh apply a parsimonious monetary policy model to estimate three key unobservable variables—specifically, the neutral real interest rate, the output gap, and the natural rate of unemployment—for three large non-inflation-targeting economies (namely, the United States, the euro area, and Japan) and seven inflation-targeting countries (namely, Australia, Canada, Chile, New Zealand, Norway, Sweden, and the United Kingdom), using quarterly data for 1970–2006 (at most). Country-by-country estimation closely follows the sequential-step procedure developed by Laubach and Williams (2003) for estimating two unobservables for the United States. The country results reported in this chapter, while mixed, show that trend output growth and the neutral real interest rate vary over time in most countries, and the natural rate of unemployment is found to vary over time in Chile and the United States. As discussed above, policymakers must consider that key unobservables may vary over time if they are to conduct monetary policy efficiently. Regarding common time trends, Schmidt-Hebbel and Walsh show that the volatilities of inflation, output growth, and the real interest rate have declined in their country sample over the last decades, which is consistent with the great moderation observed worldwide since the early 1990s. The three big economies exhibit neither large nor rising comovements of key variables over time. Most smaller inflation-targeting economies, however, exhibit rising comovements of key observables and unobservables with the United States. Finally, on convergence of variable levels observed across countries in the sample period, the authors reject convergence of unobservables in inflation-targeting countries to the levels estimated for the United States and the euro area, but they report convergence of actual growth and interest rates in most inflation-targeting countries to the growth and interest rate levels observed in the United States and the euro area.

In their chapter, Martin Melecký, Diego Rodríquez-Palenzuela, and Ulf Söderström use a model estimated on euro area data to assess the effects of monetary policy transparency and credibility on inflation and output volatility. The key uncertainty faced by private agents in the model arises from shifts in the central bank's policy rule. These shifts might reflect transitory interest rate movements, or they might reflect persistent changes in the central bank's inflation target. The authors employ a forward-looking DSGE model that incorporates sticky prices and sticky wages. They find that the gains from credibly announcing changes in the target inflation rate are relatively small. However, they show that this result depends on the assumption that the private sector fully understands the stochastic process that governs persistence in the target rate. When this aspect of the target rate behavior is not known, the inference problem private agents face is more complicated, and the gains from announcing the target can be much larger, particularly if private agents overestimate the volatility of the target.

Volker Wieland develops a model designed to provide an understanding of the path of gradual disinflation in inflation targeting countries such as Chile. He introduces two new elements into a New-Keynesian model to capture disinflationary experiences. First, private firms engage in adaptive learning; in setting prices, they need to forecast future inflation and, to do so, they employ least squares methods to update estimates of a simple forecasting equation. Second, Wieland develops a model of price indexation in which the degree of indexation is endogenously determined. This approach contrasts with the many models that assume that some prices are partially indexed to past inflation but which treat the degree of indexation as exogenous. Specifically, whenever a firm has an opportunity to optimally reset its price, it also decides whether to index future price changes to past inflation or to the central bank's inflation target. As a consequence, an immediate disinflation via a reduction in the central bank's inflation target causes firms to quickly drop backward-looking indexation and base indexation on the inflation target. The initial impact of this rapid disinflation, however, is a large output decline. The decline in real economic activity can be muted if the central bank carries out a more gradual disinflation. As firms update their assessment of inflation persistence during a gradual disinflation, the real costs of the disinflation decline, but firms are less likely to shift their indexation to the central bank's target in the gradual disinflation scenario. Wieland then goes on to analyze the use of temporary inflation targets that gradually decline toward a low steady-state inflation rate. This situation captures the gradual disinflation strategy based on annual inflation targets adopted by Chile in 1990–2000, similar to several other inflation-targeting countries that adopted annual inflation targets when actual inflation was still high. Meeting short-term targets helps increase the rate at which firms alter their indexation strategies from being based on lagged inflation to being based on the new inflation targets. This helps achieve low inflation.

Felipe Morandé and Mauricio Tejada assess the empirical importance of the three classical sources of uncertainty for monetary policy in Chile. They analyze data uncertainty by comparing real-time estimates for the output gap with each other and with final-data measures; they conclude that the correlations between real-time data and final-data output gap estimates are relatively low. To evaluate the empirical importance of additive uncertainty (associated with the variance of shocks) and multiplicative uncertainty (associate with parameter uncertainty), Morandé and Tejada estimate a small open economy forward-looking New-Keynesian model for Chile, with time-varying parameters and state-dependent variances of disturbances. The results for all model equations show that additive uncertainty dominates multiplicative uncertainty. The estimations support the hypothesis of state-dependent variances linked to two states of either low or high shock volatility. Measures of total uncertainty of both the output gap and inflation have declined over time, and the period of greater stability coincides with full-fledged inflation targeting adopted since 2001.

In previous work, Marco del Negro and Frank Schorfheide (and others) develop the DSGE-VAR model, which relaxes cross-equation restrictions and can be regarded as a structural vector autorgression (VAR) model that retains many features of the underlying DSGE specification. In this volume, Del Negro and Schorfheide present estimation results for a small open economy DSGE-VAR model for Chile in 1999–2007. The authors find it helpful to tilt their VAR estimates toward the restriction generated by their DSGE model because the VAR without tight priors is unlikely to provide good forecasts or sharp policy advice. Observed inflation variability was mostly due to domestic shocks. Regarding monetary policy rules, one finding is that the Central Bank of Chile did not respond significantly to exchange rate and terms-of-trade shocks. A stronger Central Bank response to inflation shocks would have had little effect on inflation volatility, but a weaker response would have led to an inflation volatility spike. Del Negro and Schorfheide derive two more general lessons from their exercise. First, the outcomes of policy experiments are very sensitive to the parameters that reflect the law of motion of exogenous shocks. Second, the presence of misspecification—when the DSGE model is rejected relative to a more loosely parameterized model—does not necessarily imply that the answers to the policy exercises obtained from the DSGE model are not robust.

In the final chapter, Manuel Marfán, Juan Pablo Medina, and Claudio Soto specify and calibrate a DSGE model for Chile to analyze the macroeconomic effects of shocks when private agents suffer from misperceptions about future productivity levels that generate boom-bust cycles, such as those recurrently observed in both emerging market and industrial economies in the 1990s and the 2000s. The model, based on a three-sector small open economy forward-looking DSGE specification with several nominal and real rigidities and a Taylor rule, is used to conduct several simulations. The first simulation shows that a boom-bust cycle can be simulated by an unexpected decline and subsequent reversal in the foreign interest rate, which accounts well for the stylized facts observed in Chile in the 1990s. The second simulation focuses on the effects of overoptimistic expectations about future productivity levels and, alternatively, future productivity trends, which turn out to be wrong ex post.[3] Only overoptimism

3. Overoptimism based, for instance, on the expected outcome of recent economic reforms that is ex-ante hard to evaluate.

about productivity trends (not levels) is able to replicate Chile's cycle, similarly to the foreign-interest-rate-induced cycle. Finally, Marfán, Medina, and Soto contrast the macroeconomic effects of alternative monetary policy reactions in response to an increase in trend productivity. If the central bank follows a stricter inflation-targeting regime, the boom-bust cycle of most macroeconomic variables would be amplified. If the central bank includes the exchange rate as an argument in its policy rule, it may prevent the contraction of the traded sector that occurs under the baseline policy rule, but the volatility of other variables would be amplified. This suggests that the trade-offs faced in the conduct of monetary policy (and exchange rate policy) are not trivial in a boom-bust cycle triggered by misperception about future productivity.

References

Adolfson, M., S. Laséen, J. Lindé, and M. Villani. 2008. "Evaluating an Estimated New Keynesian Small Open Economy Model." *Journal of Economic Dynamics and Control* 32(8): 2690–721.

Benati, L. and G. Vitale, 2007. "Joint Estimation of the Natural Rate of Interest, the Natural Rate of Unemployment, Expected Inflation, and Potential Output." Working paper 797. Frankfurt: European Central Bank.

Blinder, A.S. 1998. *Central Banking in Theory and Practice*. MIT Press.

Brainard, W. 1967. "Uncertainty and the Effectiveness of Policy." *American Economic Review* 57(2): 411–25.

Bullard, J. and K. Mitra. 2002. "Learning about Monetary Policy Rules." *Journal of Monetary Economics* 49(6): 1105–29.

Christiano, L.J., M. Eichenbaum, and C. Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113(1): 1–45.

Christiano, L.J., R. Motto, and M. Rostagno. 2007. "Financial Factors in Business Cycles." Northwestern University, Department of Economics.

Cogley, T., R. Colacito, and T.J. Sargent. 2007. "Benefits from U.S. Monetary Policy Experimentation in the Days of Samuelson and Solow and Lucas." *Journal of Money, Credit, and Banking* 39(S1): 67–99.

Craine, R. 1979. "Optimal Monetary Policy with Uncertainty." *Journal of Economic Dynamics and Control* 1(1): 59–83.

Erceg, C.J. and A.T. Levin. 2003. "Imperfect Credibility and Inflation Persistence." *Journal of Monetary Economics* 50(4): 915–44.

Evans, G.W. and S. Honkapohja. 2001. *Learning and Expectations in Macroeconomics*. Princeton University Press.

Galí, J. 2008. *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*. Princeton University Press.

Garnier, J. and B.R. Wilhelmsen. 2005. "The Natural Real Interest Rate and the Output Gap in the Euro Area: A Joint Estimation." Working paper 546. Frankfurt: European Central Bank.

Giannoni, M. 2002. "Does Model Uncertainty Justify Caution? Robust Optimal Monetary Policy in a Forward-Looking Model." *Macroeconomic Dynamics* 6(1): 111–44.

Giannoni, M. and M. Woodford. 2002. "Optimal Interest-Rate Rules: I. General Theory." Working paper 9419. Cambridge, Mass.: National Bureau of Economic Research.

Greenspan, A. 2003. "Opening Remarks." In *Proceedings: Monetary Policy and Uncertainty: Adapting to a Changing Economy*, 1–7. Federal Reserve Bank of Kansas City.

Hansen, L.P. and T.J. Sargent. 2003. "Robust Control for Forward Looking Models." *Journal of Monetary Economics* 50(3): 581–604.

————. 2004. *Robust Control and Economic Model Uncertainty*. Princeton University Press.

Kuttner, K.N. 1994. "Estimating Potential Output as a Latent Variable." *Journal of Business and Economic Statistics* 12(3): 361–68.

Laubach, T. and J.C. Williams, 2003. "Measuring the Natural Rate of Interest." *Review of Economics and Statistics* 85(4): 1063–70.

Levin, A., A. Onatski, J. Williams, and N. Williams. 2006. "Monetary Policy under Uncertainty in Micro-Founded Macroeconometric Models." In *NBER Macroeconomic Annual 2005*, edited by Mark Gertler and Kenneth Rogoff, 229–87. MIT Press for National Bureau of Economic Research.

Levin, A.T. and J.C. Williams. 2003. "Robust Monetary Policy with Competing Reference Models." *Journal of Monetary Economics* 50(5): 945–75.

Lindé, J. 2005. "Estimating New-Keynesian Phillips Curve: A Full Information Maximum Likelihood Approach." *Journal of Monetary Economics* 52(6): 1135–49.

McCallum, B.T. 2001. "Should Monetary Policy Respond Strongly to Output Gaps?" *American Economic Review* 91(2): 258–62.

Orphanides, A. 2003. "The Quest for Prosperity without Inflation." *Journal of Monetary Economics* 50(3): 633–63.

Orphanides, A. and J.C. Williams. 2002. "Robust Monetary Policy Rules with Unknown Natural Rates." *Brookings Papers on Economic Activity* 2: 63–145.

Romer, C.D. and D.H. Romer. 2002. "The Evolution of Economic Understanding and Postwar Stabilization Policy." In *Proceedings: Rethinking Stabilization Policy*, 11–78. Federal Reserve Bank of Kansas City.

Rudebusch, G.D. 2001. "Is the Fed Too Timid? Monetary Policy in an Uncertain World." *Review of Economics and Statistics* 83(2): 203–17.

Rudebusch, G.D. and J.C. Williams. 2008. "Revealing Secrets of the Temple: The Value of Publishing Central Bank Interest Rate Projections." In *Asset Prices and Monetary Policy*, edited by J.Y. Campbell. University of Chicago Press.

Smets, F. and R. Wouters. 2003. "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area." *Journal of the European Economic Association* 1(5): 1123–75.

Söderström, U. 2002. "Monetary Policy with Uncertain Parameters." *Scandinavian Journal of Economics* 104(1): 125–45.

Taylor, J.B. 2003. "Discretion versus Policy Rules in Practice." *Carnegie-Rochester Conferences Series on Public Policy* 39(1): 195–214.

Walsh, C.E. 2003. "Implications of a Changing Economic Structure for the Strategy of Monetary Policy." In *Proceedings: Monetary Policy and Uncertainty: Adapting to a Changing Economy*, 297–348. Federal Reserve Bank of Kansas City.

————. 2004. "Robustly Optimal Instrument Rules and Robust Control: An Equivalence Result." *Journal of Money, Credit, and Banking* 36(6): 1105–13.

Expectations, Learning, and Monetary Policy: An Overview of Recent Research

George W. Evans
University of Oregon and University of St Andrews

Seppo Honkapohja
Bank of Finland and University of Cambridge

The conduct of monetary policy in terms of interest rate or other rules has been extensively studied in recent research.[1] This literature gives a central role to forecasts of future inflation and output, and the question of whether monetary policy should be forward-looking has been subject to discussion and debate. The Bank of England and the European Central Bank include private sector forecasts and internal macroeconomic projections in their periodic reports (Bank of England, 2007; European Central Bank, 2007). Empirical evidence on Germany, Japan, and the United States since 1979 similarly suggests that central banks are forward-looking in practice (Clarida, Galí, and Gertler, 1998).

The rational expectations hypothesis, the standard benchmark in macroeconomics since the seminal work of Lucas (1976) and Sargent and Wallace (1975), has been employed in most of the research on monetary policy and interest rate rules. The most common formulation of the rational expectations hypothesis is based on the assumption that both private agents and the policymaker know the true model of the economy, except for unforecastable random

We thank Carl E. Walsh for useful comments and gratefully acknowledge financial support from the National Science Foundation grant SES-0617859 and the Economic and Social Research Council grant RES-000-23-1152.

1. Woodford (2003) is a monumental treatise on the subject, while Walsh (2003) provides an accessible graduate-level treatment. For surveys, see Clarida, Galí, and Gertler (1999) and McCallum (1999).

Monetary Policy under Uncertainty and Learning, edited by Klaus Schmidt-Hebbel and Carl E. Walsh, Santiago, Chile.

shocks.[2] The rational expectations assumption is excessively strong: neither private agents nor policymakers have perfect knowledge of the economy. In reality, economists formulate and estimate models that are used to make macroeconomic forecasts and carry out policy analysis. These models are reestimated and possibly reformulated as new data become available. In other words, economists engage in learning processes about the economy as they attempt to improve their knowledge of the economy.

Formal study of these learning processes and their implications for macroeconomic dynamics and policymaking are becoming an increasingly important line of research in macroeconomics.[3] This research is based on a principle of cognitive consistency stating that private agents and policymakers in the economy behave like applied economists and econometricians. It is thus postulated that expectations of macroeconomic variables are formed by using statistical or other formal forecasting models and procedures.

An important policy question is whether the learning processes create new tasks and constraints for macroeconomic policy. An affirmative answer to this question has been demonstrated by the recent work on learning and monetary policy.[4] This view is also reflected in recent speeches by two prominent central bank governors (see Trichet, 2005; Bernanke, 2007). This research shows that interest rate setting by monetary policymakers faces two fundamental problems. First, some of the proposed interest rate rules may not perform well when agents' expectations are out of equilibrium. The consequences of errors in forecasting, and the resulting correction mechanisms, may create instability in the economy. For (usually suboptimal) instrument rules, Bullard and Mitra (2002) consider the stability of the rational expectations equilibrium when monetary policy is conducted using variants of the Taylor rule. These rules work well only under certain parameter restrictions, and Bullard and Mitra suggest that monetary policymaking should take into account the learnability constraints on the parameters of policy behavior. For

2. Some papers do extend the standard notion of rational expectations equilibrium to an equilibrium with limited information. These extensions often assume that economic agents do not observe some variables but know the structure of the economy.

3. Evans and Honkapohja (2001) provide a treatise on the analysis of adaptive learning and its implications in macroeconomics. Evans and Honkapohja (1995, 1999), Marimon (1997), and Sargent (1993, 1999) provide surveys of the field.

4. Evans and Honkapohja (2003a) and Bullard (2006) provide surveys of the recent research.

optimal monetary policy, Evans and Honkapohja (2003c, 2006) show that certain standard forms of optimal interest rate setting by the central bank can lead to expectational instability, as economic agents unsuccessfully try to correct their forecast functions over time. Evans and Honkapohja also propose a new rule for implementing optimal policy that always leads to stability under learning.

Second, monetary policy rules, including some formulations for optimal setting of the instrument and some Taylor rules based on forecasts of inflation and the output gap, can create multiple equilibria, also called indeterminacy of equilibria.[5] Under indeterminacy there are multiple, even continua of rational expectations equilibria and the economy need not settle on the desired equilibrium. The possible rest points have been studied using stability under learning as a selection criterion (see Honkapohja and Mitra, 2004; Carlstrom and Fuerst, 2004; Evans and McGough, 2005a). Indeterminacy is not a critical problem if the fundamental rational expectations equilibrium is the only stable equilibrium under learning. Moreover, indeterminacy need not arise if the forward-looking interest rate rule is carefully designed, as shown by Bullard and Mitra (2002) and Evans and Honkapohja (2003c, 2006). The central message from these studies is that monetary policy has important new tasks when agents' knowledge is imperfect and agents try to improve their knowledge through learning. Policy should be designed to facilitate learning by private agents so that expectations do not create instability in the economy.

Recently, many further aspects of expectations, learning, and monetary policy have been analyzed in the rapidly expanding literature. In this paper, we provide a nontechnical overview of this research program. The first part of the paper reviews the basic theoretical results. We then take up some immediate practical concerns that can arise in connection with rules for interest rate setting, including issues of observability in connection with private forecasts and with current output and inflation data. A second concern is the knowledge of the structure of the economy that is required to implement optimal interest rate policies. The second part of the paper provides an overview of the recent and ongoing developments in the literature. We first summarize research on learnability of rational expectations equilibria when the basic New-Keynesian model is

5. This was first noted by Bernanke and Woodford (1997), Woodford (1999b), and Svensson and Woodford (2005). The problem was systematically explored for Taylor rules by Bullard and Mitra (2002).

extended to incorporate further features of the economy. We then discuss four topics of applied interest in more detail: policy design under perpetual learning, estimated models with learning, recurrent hyperinflations, and macroeconomic policy to combat liquidity traps and deflation.

1. The Model

We conduct our discussion using the New-Keynesian model that has become the workhorse in the analysis of monetary policy, and we directly employ its linearized version. The original nonlinear framework is based on a representative consumer and a continuum of firms producing differentiated goods under monopolistic competition. Nominal stickiness of prices arises from firms' constraints on the frequency of price changes, as originally suggested by Calvo (1983).

The behavior of the private sector is summarized by two equations:

$$x_{\mathrm{t}} = -\varphi\,(i_{\mathrm{t}} - E_t^*\pi_{t+1}) + E_t^* x_{t+1} + g_t, \tag{1}$$

which is the IS curve derived from the Euler equation for consumer optimization, and

$$\pi_t = \lambda x_t + \beta E_t^*\pi_{t+1} + u_t, \tag{2}$$

which is the price setting rule for the monopolistically competitive firms, often called the New-Keynesian Phillips or aggregate supply curve.

Here x_t and π_t denote the output gap and inflation rate for period t, respectively, and i_t is the nominal interest rate, expressed as the deviation from the steady state real interest rate. The determination of i_t is discussed below. Private sector expectations of the output gap and inflation in the next period are denoted $E_t^* x_{t+1}$ and $E_t^*\pi_{t+1}$, respectively. Since our focus is on learning behavior, these expectations need not be rational (E_t without * denotes rational expectations). The parameters φ and λ are positive and β is the discount factor with $0 < \beta < 1$.

For brevity, we do not discuss details of the derivation of equations (1) and (2), which is based on individual Euler equations under (identical) subjective expectations, together with aggregation and definitions of the variables. The Euler equations for the current period give the decisions as functions of the expected state in the next period.

Rules for forecasting the next period's values of the state variables are the other ingredient in the description of individual behavior. We assume that given forecasts, private agents make decisions according to the Euler equations.[6]

The shocks g_t and u_t are assumed to be observable and to follow

$$\begin{pmatrix} g_t \\ u_t \end{pmatrix} = \mathbf{F} \begin{pmatrix} g_{t-1} \\ u_{t-1} \end{pmatrix} + \begin{pmatrix} \tilde{g}_t \\ \tilde{u}_t \end{pmatrix}, \tag{3}$$

where

$$\mathbf{F} = \begin{pmatrix} \mu & 0 \\ 0 & \rho \end{pmatrix},$$

$0 < |\mu| < 1$, $0 < |\rho| < 1$, and $\tilde{g}_t \sim$ i.i.d. $(0, \sigma_g^2)$, $\tilde{u}_t \sim$ i.i.d. $(0, \sigma_u^2)$ are independent white noise. In addition, g_t represents shocks to government purchases or potential output (or both), and u_t represents any cost push shocks to marginal costs other than those entering through x_t. For simplicity, we assume throughout the paper that μ and ρ are known (if not, they could be estimated).

The model is closed by an equation describing the central bank's interest rate setting.[7] One approach examines instrument rules, under which i_t is directly specified in terms of key macroeconomic variables without explicit policy optimization. A prominent example of this type is the standard Taylor (1993) rule, that is,

$$i_t = \pi_t + 0.5(\pi_t - \overline{\pi}) + 0.5x_t,$$

where $\overline{\pi}$ is the target level of inflation and the target level of the output gap is zero. (Recall that i_t is specified net of the real interest rate, which in the standard Taylor rule is usually set at 2 percent).

6. This kind of behavior is boundedly rational, but in our view reasonable, since agents attempt to meet the margin of optimality between the current and the next period. Other models of bounded rationality are possible. Preston (2005, 2006) proposes a formulation in which long horizons matter in individual behavior.

7. We follow the common practice of leaving hidden the government budget constraint and the equation for the evolution of government debt. This is acceptable provided that fiscal policy appropriately accommodates the consequences of monetary policy for the government budget constraint. The interaction of monetary and fiscal policy can be important for the stability of equilibria under learning; see Evans and Honkapohja (2007), McCallum (2003), and Evans, Guse, and Honkapohja (2008). We discuss some aspects of the interaction below.

More generally, Taylor rules are of the form $i_t = \chi_0 + \chi_\pi \pi_t + \chi_x x_t$. For convenience (and without loss of generality), we take the inflation target to be $\bar{\pi} = 0$, so that this class of rules takes the form

$$i_t = \chi_\pi \pi_t + \chi_x x_t, \tag{4}$$

where χ_π, $\chi_x > 0$. Variations of the Taylor rule replace π_t and x_t by lagged values or by forecasts of current or future values.

Alternatively, interest rate policy can be derived explicitly to maximize a policy objective function. This is frequently taken to be of the quadratic loss form, that is,

$$E_t \sum_{s=0}^{\infty} \beta^s \left[(\pi_{t+s} - \bar{\pi})^2 + \alpha x_{t+s}^2 \right], \tag{5}$$

where $\bar{\pi}$ is the inflation target. This type of optimal policy is often called flexible inflation targeting in the current literature (see, for example, Svensson, 1999, 2003). The policymaker is assumed to have the same discount factor, β, as the private sector, while α is the relative weight placed by the policymaker on the output target. The case of $\alpha = 0$ represents strict inflation targeting. The loss function (5) can alternatively be viewed as a quadratic approximation to the welfare function of a representative agent.[8]

The literature on optimal policy under rational expectations distinguishes between optimal discretionary policy, in which the policymaker is unable to commit to policies for future periods, and optimal policy in which such commitment is possible. Under commitment, the policymaker can do better because of the effect on private expectations, but commitment policy exhibits time inconsistency, in the sense that policymakers would have an incentive to deviate from the policy in the future. Assuming that the policy has been initiated at some point in the past (the timeless perspective described by Woodford, 1999a), and setting $\bar{\pi} = 0$, the first-order condition specifies

$$\lambda \pi_t + \alpha(x_t - x_{t-1}) = 0 \tag{6}$$

in every period.

8. See Rotemberg and Woodford, 1999; Woodford, 2003. In this formulation, α is a function of various deep structural parameters in the fully microfounded version of the model.

Condition (6) for optimal policy with commitment is not a complete specification of monetary policy, since one must also provide a reaction function for i_t that implements the policy. A number of interest rate rules are consistent with the model described in equations (1) and (2), the optimality condition (6), and rational expectations. However, some ways of implementing optimal monetary policy can make the economy vulnerable to either indeterminacy or expectational instability or both, while other implementations are robust to these difficulties.

We will consider fundamentals-based and expectations-based rules. The basic fundamentals-based rule depends only on the observable exogenous shocks g_t and u_t and on x_{t-1}:

$$i_t = \psi_x x_{t-1} + \psi_g g_t + \psi_u u_t, \tag{7}$$

where the optimal coefficients are determined by the structural parameters and the policy objective function. The coefficients ψ_i are chosen to neutralize the effects of aggregate demand shocks, g_t, and to strike the optimal balance between output and inflation effects for inflation shocks, u_t. The dependence of i_t on x_{t-1} is optimally chosen to take advantage of the effects on expectations of commitment to a rule.[9]

Expectations-based optimal rules are advocated in Evans and Honkapohja (2003c, 2006) because, as further discussed below, fundamentals-based optimal rules are often unstable under learning. If private expectations are observable, they can be incorporated into the interest rate rule. When this is done appropriately, the rational expectations equilibrium will be stable under learning and optimal policy can thus be successfully implemented. The essence of these rules is that they do not assume rational expectations on the part of private agents, but are designed to feed back on private expectations in such a way that they generate convergence to the optimal rational expectations equilibrium under learning. (If expectations are rational, these rules deliver the optimal equilibrium.)

The optimal expectations-based rule under commitment is

$$i_t = \delta_L x_{t-1} + \delta_\pi E_t^* \pi_{t+1} + \delta_x E_t^* x_{t+1} + \delta_g g_t + \delta_u u_t. \tag{8}$$

9. The coefficients of the interest rate rule (7) are $\psi_x = \bar{b}_x[\varphi^{-1}(\bar{b}_x - 1) + \bar{b}_\pi]$, $\psi_g = \varphi^{-1}$, and $\psi_u = [\bar{b}_\pi + \varphi^{-1}(\bar{b}_x + \rho - 1)]\bar{c}_x + \bar{c}_\pi\rho$. Here $\bar{b}_x = (2\beta)^{-1}[\varsigma - (\varsigma^2 - 4\beta)^{1/2}]$ with $\zeta = 1 + \beta + \lambda^2/\alpha$, and $\bar{b}_\pi = (\alpha/\lambda)(1 - \bar{b}_x)$, $\bar{c}_x = -[\lambda + \beta\bar{b}_\pi + (1 - \beta\rho)(\alpha/\lambda)]^{-1}$, $\bar{c}_\pi = -(\alpha/\lambda)\bar{c}_x$.

The coefficients of equation (8) are

$$\delta_L = \frac{-\alpha}{\varphi(\alpha+\lambda^2)},\ \delta_\pi = 1 + \frac{\lambda\beta}{\varphi(\alpha+\lambda^2)},$$
$$\delta_x = \varphi^{-1},\ \delta_g = \varphi^{-1} \text{ and } \delta_u = \frac{\lambda}{\varphi(\alpha+\lambda^2)}. \tag{9}$$

This rule is obtained by combining the IS curve equation (1), the price-setting equation (2), and the first order optimality condition (6), treating private expectations as given.[10]

Interest rate rules based on observations of x_t and π_t that (outside the rational expectations equilibrium) only approximate the first-order optimality condition (6) are considered by Svensson and Woodford (2005). They suggest a set of hybrid rules, the simplest of which would be

$$i_t = \psi_x x_{t-1} + \psi_g g_t + \psi_u u_t + \theta\left[\pi_t + \frac{\alpha}{\lambda}(x_t - x_{t-1})\right], \tag{10}$$

where $\theta > 0$. This rule combines the fundamentals-based rule of equation (7) with the correction for the first-order condition.[11] Rule (10) delivers the optimal equilibrium under rational expectations. McCallum and Nelson (2004) suggest another hybrid rule, which takes the form

$$i_t = \pi_t + \theta\left[\pi_t + \frac{\alpha}{\lambda}(x_t - x_{t-1})\right], \tag{11}$$

where $\theta > 0$.

2. Determinacy and Stability under Learning

Given an interest rate rule, we can obtain the reduced form of the model and study its properties under rational expectations. Two basic properties of interest are determinacy of the rational expectations

10. Under optimal discretionary policy the first-order condition is $\lambda\pi_t + \alpha x_t = 0$, and the coefficients are identical except that $\delta_L = 0$. The discretionary case is analyzed in Evans and Honkapohja (2003c).

11. The model and the interest rate rule analyzed in Svensson and Woodford (2005) incorporate additional information lags.

solution and stability under learning of the rational expectations equilibrium.

Consider the system given by equations (1), (2), and (3) and one of the i_t policy rules (4), (7), (8), (10), or (11). Defining the vectors

$$\mathbf{y}_t = \begin{pmatrix} x_t \\ \pi_t \end{pmatrix} \text{ and } \mathbf{v}_t = \begin{pmatrix} g_t \\ u_t \end{pmatrix},$$

the reduced form can be written as

$$\mathbf{y}_t = \mathbf{M}E_t^*\mathbf{y}_{t+1} + \mathbf{N}\mathbf{y}_{t-1} + \mathbf{P}\mathbf{v}_t, \tag{12}$$

for appropriate matrices **M**, **N**, and **P**. In the case of policy rule (4), we have **N** = **0** and thus the simpler system

$$\mathbf{y}_t = \mathbf{M}E_t^*\mathbf{y}_{t+1} + \mathbf{P}\mathbf{v}_t. \tag{13}$$

We now briefly describe the concepts of determinacy/indeterminacy and stability under adaptive (least squares) learning using the general frameworks of equations (12) and (13).

The first issue of concern is whether under rational expectations the system possesses a unique stationary equilibrium, in which case the model is said to be determinate. The model is said to be indeterminate if it has multiple stationary solutions. These multiple solutions include sunspot solutions, in which the rational expectations equilibrium depends on extraneous random variables that influence the economy solely through agents' expectations.[12]

The second issue concerns stability under adaptive learning. In the introduction, we stressed the principle of cognitive consistency according to which agents in the model are assumed to behave like econometricians or statisticians when they form their expectations. In the next section, this approach is formalized in terms of the perceived law of motion (PLM) describing the agents' beliefs. These beliefs concern the stochastic process followed by the endogenous (and exogenous) variables that need to be forecasted. The parameters

12. If the model is indeterminate, one can ask whether the sunspot solutions are stable under learning. For a general discussion see Evans and Honkapohja (2001). In general, different forms of sunspot solutions exist, and stability under learning can depend on the particular representation; see Evans and Honkapohja (2003b) and Evans and McGough (2005b).

of the PLM are updated using an appropriate statistical technique, called an adaptive learning rule, and forecasts are made using the estimated PLM at each moment of time. If private agents follow an adaptive learning rule like recursive least squares to update the parameters of their forecasting model, will the rational expectations solution of interest be stable—that is, will it be reached asymptotically by the learning process? If not, the rational expectations equilibrium is unlikely to be attained. This is the focus of the papers by Bullard and Mitra (2002, 2007), Evans and Honkapohja (2003c, 2006), and many others.

2.1 Digression on Methodology

Consider first the simpler reduced-form equation (13) under rational expectations. For determinacy to hold, both eigenvalues of the 2×2 matrix $\mathbf{M}$ must lie inside the unit circle. In the determinate case, the unique stationary solution will be of the minimal state variable (MSV) form:

$$\mathbf{y}_t = \bar{\mathbf{c}}\mathbf{v}_t,$$

where $\bar{\mathbf{c}}$ is a 2×2 matrix that is easily computed. If, instead, one or both roots lie inside the unit circle, then the model is indeterminate. There will still be a solution of the MSV form, but there will also be other stationary solutions.

Next, we consider the system under learning. Suppose that agents believe that the solution is of the form

$$\mathbf{y}_t = \mathbf{a} + \mathbf{c}\mathbf{v}_t, \tag{14}$$

while the the 2×1 vector $\mathbf{a}$ and the the 2×2 matrix $\mathbf{c}$ are not known but instead are estimated by the private agents. Equation (14) is the PLM of the agents. We include an intercept vector because, although we have translated all variables to have zero means for theoretical simplicity, in practice agents will need to estimate intercepts as well as slope parameters.[13]

13. Private agents and the policymaker are here assumed to observe the shocks $\mathbf{v}_t$. If $\mathbf{v}_t$ is not observable then the PLM would be adjusted to reflect relevant available information.

With this PLM and parameter estimates ($\mathbf{a}$, $\mathbf{c}$), agents would form expectations as

$$E_t^* \mathbf{y}_{t+1} = \mathbf{a} + \mathbf{cFv}_t,$$

where $\mathbf{F}$ is either known or also estimated. Inserting these expectations into eqauation (13) and solving for $\mathbf{y}_t$, we get the implied actual law of motion (ALM), that is, the law that $\mathbf{y}_t$ would follow for a fixed PLM ($\mathbf{a}$, $\mathbf{c}$).[14] This is given by

$$\mathbf{y}_t = \mathbf{Ma} + (\mathbf{P} + \mathbf{McF})\mathbf{v}_t.$$

We have thus obtained an associated mapping from PLM to ALM, given by

$$\mathbf{T(a, c) = (Ma, P + McF)v}_t,$$

and the rational expectations solution ($\mathbf{0}$, $\bar{\mathbf{c}}$) is a fixed point of this map.

Under real-time learning, the sequence of events is as follows.[15] Private agents begin period t with estimates ($\mathbf{a}_t$, $\mathbf{c}_t$) of the PLM parameters computed on the basis of data through $t - 1$. Next, exogenous shocks $\mathbf{v}_t$ are realized, and private agents form expectations $E_t^* \mathbf{y}_{t+1} = \mathbf{a}_t + \mathbf{c}_t\mathbf{Fv}_t$ (assuming for convenience that $\mathbf{F}$ is known). Following, for example, policy rule (4), the central bank sets the interest rate i_t, and $\mathbf{y}_t$ is generated according to equations (1) and (2) together with the interest rate rule. This temporary equilibrium is summarized by equation (13). At the beginning of $t + 1$ agents add the new data point to their information set to update their parameter estimates to ($\mathbf{a}_{t+1}$, $\mathbf{c}_{t+1}$) using least squares, for example, and the process continues. The question of interest is whether ($\mathbf{a}_t$, $\mathbf{c}_t$) $\rightarrow$ ($\mathbf{0}$, $\bar{\mathbf{c}}$) over time.

It turns out that the answer to this question is given by the E-stability principle, which advises us to look at the differential equation

$$\frac{d}{d\tau}(\mathbf{a},\mathbf{c}) = \mathbf{T}(\mathbf{a},\mathbf{c}) - (\mathbf{a},\mathbf{c}),$$

14. The ALM describes the temporary equilibrium for given expectations, as specified by the forecasts from the given PLM.

15. Formal analysis of learning and E-stability for multivariate linear models is provided in Evans and Honkapohja (2001, chap. 10).

where τ denotes notional time. If the rational expectations equilibrium $(\mathbf{0}, \bar{\mathbf{c}})$ is locally asymptotically stable under this differential equation, then the equilibrium is stable under least squares and closely related learning rules. Conditions for local stability of this differential equation are known as expectational stability or E-stability conditions. We also refer to these stability conditions as the conditions for stability under adaptive learning, the conditions for stability under learning, or the conditions for learnability of the equilibrium.

For the reduced-form equation (13), it can be shown that the two E-stability conditions are that the eigenvalues of $\mathbf{M}$ have real parts less than one and that all products of eigenvalues of $\mathbf{M}$ times eigenvalues of $\mathbf{F}$ have real parts less than one. It follows that for this reduced form, the conditions for stability under adaptive learning are implied by determinacy, but not vice versa.[16] This is not, however, a general result: sometimes E-stability is a stricter requirement than determinacy, and in other cases neither condition implies the other.

Consider next the reduced-form equation (12). Standard techniques are available to determine whether the model is determinate.[17] In the determinate case, the unique stationary solution takes the MSV form

$$\mathbf{y}_t = \mathbf{a} + \mathbf{b}\mathbf{y}_{t-1} + \mathbf{c}\mathbf{v}_t, \tag{15}$$

for appropriate values $(\mathbf{a}, \mathbf{b}, \mathbf{c}) = (\mathbf{0}, \bar{\mathbf{b}}, \bar{\mathbf{c}})$. In the indeterminate case, there are multiple solutions of this form, as well as non-MSV rational expectations equilibrium.

To examine stability under learning, we treat equation (15) as the agents' PLM. Under real-time learning, agents estimate the coefficients $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ of equation (15). This is a vector autoregression (VAR) with exogenous variables $\mathbf{v}_t$. The estimates $(\mathbf{a}_t, \mathbf{b}_t, \mathbf{c}_t)$ are updated at each point in time by recursive least squares. Once again it can be shown that the E-stability principle gives the conditions for local convergence of real-time learning.

16. See McCallum (2007) for conditions when determinacy implies E-stability.

17. The procedure is to rewrite the model in first-order form and compare the number of nonpredetermined variables with the number of roots of the forward-looking matrix that lie inside the unit circle.

For E-stability, we compute the mapping from the PLM to the ALM as follows. The expectations corresponding to equation (15) are given by

$$E_t^* \mathbf{y}_{t+1} = \mathbf{a} + \mathbf{b}(\mathbf{a} + \mathbf{b}\mathbf{y}_{t-1} + \mathbf{c}\mathbf{v}_t) + \mathbf{c}\mathbf{F}\mathbf{v}_t, \tag{16}$$

where we are treating the information set available to the agents, when forming expectations, as including $\mathbf{v}_t$ and $\mathbf{y}_{t-1}$ but not $\mathbf{y}_t$. (Alternative information assumptions would be straightforward to consider.) This leads to the mapping from PLM to ALM given by

$$\mathbf{T}(\mathbf{a},\mathbf{b},\mathbf{c}) = [\mathbf{M}(\mathbf{I} + \mathbf{b})\mathbf{a},\ \mathbf{M}\mathbf{b}^2 + \mathbf{N},\ \mathbf{M}(\mathbf{b}\mathbf{c} + \mathbf{c}\mathbf{F}) + \mathbf{P}], \tag{17}$$

E-stability is again determined by the differential equation

$$\frac{d}{d\tau}(\mathbf{a},\mathbf{b},\mathbf{c}) = \mathbf{T}(\mathbf{a},\mathbf{b},\mathbf{c}) - (\mathbf{a},\mathbf{b},\mathbf{c}), \tag{18}$$

and the E-stability conditions govern stability under least squares learning.

2.2 Results for Monetary Policy

We now describe the determinacy and stability results for the interest rate rules described in section 1.

2.2.1 Taylor rules

Bullard and Mitra (2002) consider Taylor-type rules and find that the results are sensitive to whether the i_t rule conditions on current, lagged or expected future output and inflation. In addition to assuming that $\chi_\pi, \chi_x \geq 0$, they assume that the serial correlation parameters in $\mathbf{F}$ are nonnegative. The results are particularly straightforward and natural for policy rule (4).[18] Bullard and Mitra (2002) show that the rational expectations equilibrium is determinate and stable under learning if and only if (using our notation)

18. Throughout we assume that we are not exactly on the border of the regions of determinacy or stability.

$$\lambda(\chi_\pi - 1) + (1 - \beta)\chi_x > 0.$$

In particular, determinacy and stability are guaranteed if policy obeys the Taylor principle that $\chi_\pi > 1$, so that nominal interest rates respond at least one for one with inflation.

The situation is more complicated if lagged or forward-looking Taylor rules are used, and full analytical results are not available. For the lagged variable case, they find that for $\chi_\pi > 1$ and a sufficiently small $\chi_x > 0$, the policy leads to a rational expectations equilibrium that is determinate and stable under learning. For $\chi_\pi > 1$ but χ_x too large, the system is explosive.

Bullard and Mitra (2002) also look at forward-looking versions of the Taylor rule, taking the form

$$i_t = \chi_\pi E_t^* \pi_{t+1} + \chi_x E_t^* x_{t+1}, \tag{19}$$

where χ_π, $\chi_x > 0$ and where we can interpret $E_t^* \pi_{t+1}$ and $E_t^* x_{t+1}$ as identical one-step-ahead forecasts, based on least-squares updating, used by both private agents and policymakers. They find that for $\chi_\pi > 1$ and a sufficiently small $\chi_x > 0$, the policy leads to a rational expectations equilibrium that is determinate and stable under learning. Now for $\chi_\pi > 1$ and a large χ_x, the system is indeterminate, yet the MSV solution is stable under learning. E-stable sunspot equilibria are also possible, however, as shown by Honkapohja and Mitra (2004) and discussed further by Carlstrom and Fuerst (2004) and Evans and McGough (2005a).

The Bullard and Mitra (2002) results emphasize the importance of the Taylor principle in obtaining stable and determinate interest rate rules.[19] At the same time, their results show that stability under learning must not be taken for granted, even when the system is determinate so that a unique stationary solution exists. The policymaker must appropriately select the parameters of the policy rule, χ_π, χ_x, when an instrument rule describes policy. Stability under learning provides a constraint for this choice.

19. Bullard and Mitra (2007) extend their analysis to include interest rate inertia, while Kurozumi (2006) considers modifications to the determinacy and E-stability results when the model structure is varied. Mitra (2003) examines performance of the related case of nominal income targeting.

2.2.2 Optimal monetary policy

Evans and Honkapohja (2006) focus on optimal monetary policy under commitment. It turns out that under the fundamentals-based policy rule (7), the economy is invariably unstable under learning. This is the case even though this rule yields regions in which the optimal rational expectations equilibrium is determinate.[20] The basic intuition for this result can be seen from the following reduced-form equation:

$$\begin{pmatrix} x_t \\ \pi_t \end{pmatrix} = \begin{pmatrix} 1 & \varphi \\ \lambda & \beta + \lambda\varphi \end{pmatrix} \begin{pmatrix} E_t^* x_{t+1} \\ E_t^* \pi_{t+1} \end{pmatrix} + \begin{pmatrix} -\varphi\psi_x & 0 \\ -\lambda\varphi\psi_x & 0 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ \pi_{t-1} \end{pmatrix} + \begin{pmatrix} -\varphi\psi_u \\ 1 - \lambda\varphi\psi_u \end{pmatrix} u_t. \quad (20)$$

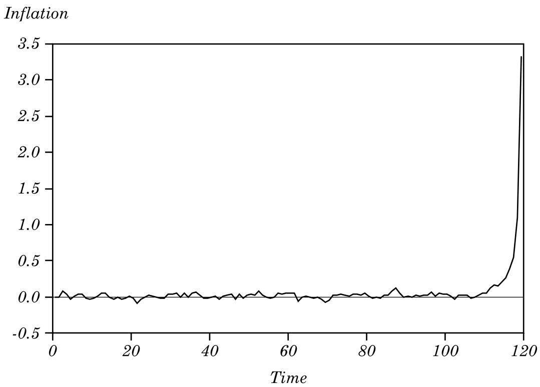
Since typically $\beta + \lambda\varphi > 1$, upward mistakes in $E_t^* \pi_{t+1}$ lead to higher π_t, both directly and indirectly through lower ex ante real interest rates, which under learning sets off a cumulative movement away from the rational expectations equilibrium. The feedback from x_{t-1} under the fundamentals-based i_t rule with commitment (7) does not stabilize the economy. Figure 1 shows how divergence from the optimal rational expectations equilibrium occurs under rule (7).[21] The instability of the fundamentals-based rules, which are designed to obtain optimal policy, serves as a strong warning to policymakers not to automatically assume that rational expectations will be attained. It is necessary to examine explicitly the robustness of contemplated policy rules to private agent learning.

In Evans and Honkapohja (2003c, 2006), we show how the problems of instability and indeterminacy can be overcome if private agents' expectations are observable, so that interest rate rules can be partly conditioned on these expectations. In Evans and Honkapohja (2006), we show that under rule (8), the economy is determinate and the optimal rational expectations equilibrium is stable under private agent learning for all possible structural parameter values. The key to the stability results can be seen from the reduced form,

20. The learning stability results are sensitive to the detailed information assumptions. With the PLM equation (15), if agents can make forecasts conditional also on $\mathbf{y}_t$, then there are regions of both stability and instability under the fundamentals-based rule, depending on the structural parameters.

21. Figures 1 and 2 are based on the calibration by McCallum and Nelson (1999). Using other calibrations would yield similar results.

Figure 1. Instability with a Fundamentals-Based Rule



Source: Authors' calculations.

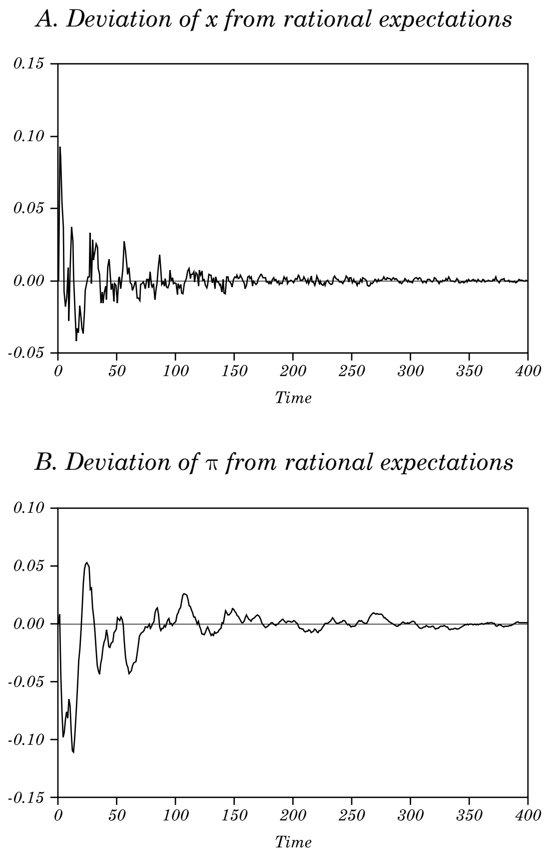
$$\begin{pmatrix} x_t \\ \pi_t \end{pmatrix} = \begin{pmatrix} 0 & -\dfrac{\lambda\beta}{\alpha+\lambda^2} \\ 0 & \dfrac{\alpha\beta}{\alpha+\lambda^2} \end{pmatrix} \begin{pmatrix} E_t^* x_{t+1} \\ E_t^* \pi_{t+1} \end{pmatrix} + \begin{pmatrix} \dfrac{\alpha}{\alpha+\lambda^2} & 0 \\ \dfrac{\alpha\lambda}{\alpha+\lambda^2} & 0 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ \pi_{t-1} \end{pmatrix} + \begin{pmatrix} -\dfrac{\lambda}{\alpha+\lambda^2} \\ \dfrac{\alpha}{\alpha+\lambda^2} \end{pmatrix} u_t. \tag{21}$$

In equation (21), the feedback from inflation expectations to actual inflation is stabilizing since the coefficient $\alpha\beta / (\alpha + \lambda^2)$ is less than one and the influence of x_{t-1} is also weak. Deviations from rational expectations are thus offset by policy in such a way that under learning private agents are guided over time to form expectations consistent with the optimal equilibrium. Our expectations-based rule obeys a form of the Taylor principle, since $\delta_\pi > 1$. Figure 2 illustrates convergence of learning under rule (8).

Our optimal policy rule is conditioned on both private expectations and observable exogenous shocks, as well as lagged output. Moreover, when computing the optimal expectations-based rule, the central bank must use the correct structural model of the IS and price setting relationships, which in turn depend on the specific form of boundedly rational individual behavior. For example, the form of the optimal expectations-based rule would be different if agents followed the long-horizon decision rules considered by Preston (2005, 2006).

Variations of fundamentals-based rules can perform well in some cases, at least for a relevant region of structural parameter values.

Figure 2. Stability with an Expectations-Based Rule



Source: Authors' calculations.

For the hybrid rule suggested by Svensson and Woodford (2005), numerical analysis shows that, in calibrated models, rule (10) yields both determinacy and stability under learning for sufficiently high values of θ. Similarly, the hybrid rule suggested by McCallum and Nelson (2004) appears to deliver E-stability of the rational expectations equilibrium. Another favorable case emerges if policy objective (5) is extended to include a motive for interest rate stabilization. Duffy and Xiao (2007b) show that in this case an optimal Taylor-type rule can deliver determinacy and E-stability for a region of parameter values that includes the usual calibrations used in the literature. We comment further below on stability with constant-gain learning for operational versions of these rules.

Finally, some researchers have proposed monetary policy formulations other than interest rate rules. For example, policy could be formulated as a money supply rule, such as the Friedman proposal for k percent money growth. Evans and Honkapohja (2003d) show that Friedman's rule always delivers determinacy and E-stability in the standard New-Keynesian model, but it does not perform well in terms of the policy objective function. Dennis and Ravenna (2008) examine stability of the economy under optimal discretionary policy, formulated as a targeting rule, for different forms of private agent learning.

2.3 Some Practical Concerns

Many of the i_t rules discussed above may not be operational, as discussed in McCallum (1999). For example, McCallum and Nelson (2004) note that it may be unrealistic to assume that policymakers can condition policy on current x_t and π_t. Similarly, policymakers may not have access to accurate observations on private expectations. We consider these points in reverse order. In the subsequent discussion, we focus on the expectations-based rule (8), the Taylor rule (4), and the hybrid rules (10) and (11).

2.3.1 Observability of private expectations

The expectations-based rule (8) requires observations of current private expectations of future variables. While survey data on private forecasts of future inflation and various measures of future output exist, there are concerns about the accuracy of this data. If observations of expectations are subject to a white noise measurement error, then our stability and determinacy results are unaffected. Furthermore, if measurement errors are small, then the policy will be close to optimal. If measurement errors are large, however, then this will lead to a substantial deterioration in performance. In this case, one might consider substituting a proxy for such observations. Since we are assuming that private agents forecast by running VARs, the most natural proxy is for the central bank to estimate corresponding VARs and use these in equation (8).

Suppose now that agents and the central bank begin with different initial estimates, possibly have different learning rules, and use data sets with different initial dates. When the private agents and the central bank are separately estimating and forecasting VARs, we must distinguish between their expectations. An extended E-stability analysis for economies with heterogenous expectations gives the conditions for convergence of heterogeneous learning, as shown in Honkapohja and Mitra (2006). Honkapohja and Mitra (2005b) analyze this issue for the case of optimal discretionary policy and expectations-based interest rate rules. Evans and Honkapohja (2003a) show that using VAR proxies can also achieve convergence to the optimal rational expectations equilibrium with commitment. Finally, Muto (2008) considers the consequences of learning from the published central bank's forecast.

The form of the extended E-stability conditions for heterogeneous learning depends on the nature of heterogeneity among agents. If the heterogeneities are transient (in the sense described in Honkapohja and Mitra, 2006), then the standard E-stability conditions directly apply. In cases of persistent heterogeneity, the learning stability conditions are somewhat sensitive to the detailed assumptions. Additional restrictions are required for stability in some cases, such as when private agents estimate parameters using stochastic gradient techniques while the central bank uses least squares.

2.3.2 Unavailability of current data

A difficulty with the standard Taylor rule (equation 4), as well as the hybrid rules of Svensson and Woodford (2005) and McCallum and Nelson (2004), is that they presuppose that the policymaker can observe both the current output gap and inflation when setting the interest rate. McCallum (1999) has criticized such policy rules as not being operational. In the case of the Taylor rule, Bullard and Mitra (2002) show that this problem of unobservability can be avoided by the use of "nowcasts" $E_t^* \mathbf{y}_t$ in place of the actual data $\mathbf{y}_t$. Determinacy and E-stability conditions are not affected by this modification.

For the hybrid rules, performance depends on the rule. Numerical analysis suggests that E-stability can still be achieved for the Svensson-Woodford rule under standard values of the parameters. The situation is more complex for the McCallum-Nelson rule. McCallum and Nelson (2004) suggest using forward expectations in place of actual data. Doing so, however, means that determinacy and stability under learning are no longer guaranteed, and sufficiently large values of the policy parameter θ induce both instability under learning and indeterminacy. This is unfortunate since large values of θ are needed to achieve a close approximation to optimal policy. Evans and Honkapohja (2003a) argue that the loss in welfare relative to the optimum is significant if θ is required to satisfy the constraints of E-stability and determinacy.

An additional issue with stability under learning arises when current data are unobservable to the policymaker. If private agents are using constant-gain learning (see section 4.2 for details), the stability conditions are more demanding. As discussed in Evans and Honkapohja (2008), the hybrid rules suggested by Svensson and Woodford (2005) and McCallum and Nelson (2004), as well as the Taylor-type optimal rule of Duffy and Xiao (2007b), are subject to the problem of instability under constant-gain learning for many realistic gain parameter values.

2.3.3 Imperfect knowledge of structural parameters

A third practical concern is that the use of optimal rules requires knowledge of the true values of the structural parameters on the part of the central bank. Evans and Honkapohja (2003a, 2003c) extend the basic analysis to a situation in which the central bank estimates the structural parameters φ and λ in equations (1) and (2) and in each period uses the current estimates in its optimal interest rate rule.[22] The basic results concerning optimal interest rate rules extend naturally to this situation. The fundamental-based rules under commitment and discretion are not learnable, while the corresponding expectations-based rules deliver convergence of simultaneous learning by the private agents and the central bank.

Since optimal monetary policy depends on structural parameters, uncertainty about their values is an issue, even if the central bank can learn their values asymptotically. Evans and McGough (2007) examine optimal Taylor-type rules based on Bayesian model averaging, where determinacy and stability under learning are imposed across all plausible structural parameter values.

Orphanides and Williams (2007) also stressed the importance of structural uncertainty. Their model incorporates both imperfect knowledge about the natural rates of interest and unemployment and constant-gain learning by private agents. They emphasize monetary policy rules that are robust along all of these dimensions.

22. It is natural to assume that the central bank knows the discount factor, β, and the policy weight, α.

3. Further Developments

A great deal of recent work extends the results on monetary policy and learning. One of the more significant issues, from an applied point of view, is the incorporation of constant-gain or perpetual learning, in which private agents update estimates using least squares, but discount past data. Consequently, agents' expectations never fully converge to the rational expectations equilibrium, but they are (asymptotically) in a neighborhood of the equilibrium, provided the equilibrium is stable. Several papers discuss the issue of optimal policy when the learning process itself is incorporated into the optimal policy problem, either during the learning transition or under perpetual learning (Orphanides and Williams, 2005a, 2007; Molnar and Santoro, 2006; Gaspar, Smets, and Vestin, 2005, 2006). A related issue studied by Ferrero (2007) concerns speed of convergence of learning for alternative policy rules. Arifovic, Bullard, and Kostyshyna (2007) consider the implications of social learning for monetary policy rules.

Extensions of the learning stability results to open economy and multi-country settings have been made by Llosa and Tuesta (2008), Bullard and Schaling (2006), Bullard and Singh (2006), Zanna (2006), and Wang (2006), among others. These papers examine both Taylor-type rules and interest rate rules that target real exchange rates. Another extension of the basic model considers determinacy and E-stability of rational expectations equilibrium when long-term interest rates are introduced to the model (see McGough, Rudebusch, and Williams, 2005; Tesfaselassie, Schaling, and Eijffinger, 2008).

In the standard New-Keynesian model, monetary policy works entirely via the demand side. Kurozumi (2006) and Llosa and Tuesta (2007) consider how determinacy and learning conditions are altered when monetary policy has direct effects on inflation. Kurozumi and van Zandweghe (2008b) extend the analysis to the model with search in labor markets, while Wieland (2008) analyzes the role of endogenous indexation for inflation targeting. Kurozumi and van Zandweghe (2008a), Duffy and Xiao (2007a), and Pfajfar and Santoro (2007a) examine in detail how the learning stability conditions for Taylor rules are modified when capital is incorporated into the New-Keynesian model. The results for models with capital depend on precisely how capital is modeled, that is, on whether adjustment costs are included and whether there is firm-specific capital or a rental market for capital. One result that emerges in some of these settings is that determinacy and E-stability require the interest rate rule to have a positive response to the output gap.

Learning plays a key role in a number of detailed policy issues. Some central banks often set monetary policy based on the constant interest rate that is expected to deliver a target inflation rate over a specified horizon. Honkapohja and Mitra (2005a) explore how this affects stability under learning. Transparency and communication of targets and rules are further considered by Berardi and Duffy (2007) and Eusepi and Preston (2007).

While the New-Keynesian model is based on a linearized set-up under Calvo-type pricing, nonlinear settings based on quadratic costs of price adjustments suggested by Rotemberg (1982) have been useful for studying the possibility of liquidity trap equilibria.[23] Benhabib, Schmitt-Grohé, and Uribe (2001) investigate this issue under perfect foresight. Evans and Honkapohja (2005) analyze this set-up under learning for the case of flexible prices, while Evans, Guse, and Honkapohja (2008) focus on a sticky-price version. The latter paper is discussed further below. Sticky-information models that incorporate learning have also been developed (Branch and others, 2007, 2008).

A number of theoretical learning topics have recently been pursued that have a bearing on monetary policy issues. Forward-looking Taylor rules can generate indeterminacy for some choices of parameters. In these cases, can stationary sunspot equilibria be stable under learning? Honkapohja and Mitra (2004), Carlstrom and Fuerst (2004), and Evans and McGough (2005a) examine this issue in the New-Keynesian setting, where conditions for stable sunspots are obtained in linearized models, while Eusepi (2007) looks at the question in a nonlinear setting. Evans, Honkapohja, and Marimon (2007) show that stable sunspot equilibria can arise in a cash-in-advance framework in which part of the government deficit is financed by seigniorage.

Constant-gain learning raises the issue of the appropriate choice of gain parameter (see Evans and Honkapohja, 1993, 2001, chap. 14; Marcet and Nicolini, 2003). Evans and Ramey (2006) consider this problem in a simple monetary set-up in which private agents face an unknown regime-switching process. This paper shows how the Lucas critique, based on rational expectations, can carry over to learning dynamics in which agents have misspecified models.

23. Bullard and Cho (2005) study the possibility of liquidity traps under learning using a linearized New-Keynesian model.

A number of papers model monetary policy with near-rational expectations. Woodford (2005) develops a min-max concept of policy robustness in which policymakers protect against agents' expectations being distorted away from rational expectations within some class of near rational expectations. Bullard, Evans, and Honkapohja (2008) consider the possibility that expert judgement based on extraneous factors believed to be present can become almost self-fulfilling. They show how to alter monetary policy to protect against these near-rational exuberance equilibria. Heterogeneous expectations is another area that is increasingly receiving attention. Theoretical work on monetary policy that allows for learning heterogeneity across private agents, or between policymakers and private agents, includes Evans, Honkapohja, and Marimon (2001), Giannitsarou (2003), and Honkapohja and Mitra (2005b, 2006). Guse (2008) and Berardi (2008) introduce misspecified expectations to the New-Keynesian model, while Tetlow and von zur Muehlen (2008) introduce robustness considerations to the analysis of stability. A related approach emphasizes that private agents may have different types of predictors, with the proportions of agents using the different forecast methods changing over time according to relative forecast performance (see Brock and Hommes, 1997; Branch and Evans, 2006b). For an application to monetary inflation models and monetary policy, see Branch and Evans (2007) and Brazier and others (2008).

Empirical applications of learning to macroeconomics and monetary policy include Bullard and Eusepi (2005) and Orphanides and Williams (2005b), who look at estimated models that focus on the explanation of the large increase in inflation rates in the 1970s. Milani (2005, 2007) incorporates learning as a way to explain persistence in New-Keynesian models, using U.S. data. The first attempts to incorporate learning to applied dynamic stochastic general equilibrium (DSGE) models have recently been undertaken by Slobodyan and Wouters (2007) and Murray (2007). Several papers use least-squares learning models or dynamic predictors to explain expectations data, including Branch (2004), Branch and Evans (2006a), Orphanides and Williams (2005c), Basdevant (2005), Pfajfar (2007), and Pfajfar and Santoro (2007b).

Other important empirical learning papers include Marcet and Nicolini (2003), which studies hyperinflation in South American countries (we discuss this paper in detail below). In addition, Cogley and Sargent (2005), Sargent, Williams, and Zha (2006), Primiceri (2006), Ellison and Yates (2007), and Carboni and Ellison (2007, 2008) emphasize the importance of policymaker model uncertainty and the role of central bank learning in explaining the historical evolution of inflation and unemployment in the post-1950 period.

In the next sections we discuss four recent topics that address important applied questions. Learning plays a crucial role in these analyses, but the main focus in each case goes well beyond the stability of rational expectations equilibrium under learning.

4. Perpetual Learning and Persistence

The preceding sections were concerned with the stability of the rational expectations equilibrium under least squares (LS) learning. That is, we used LS learning to assess whether a rational expectations equilibrium is attainable if we model agents as econometricians. Orphanides and Williams (2005a) show that taking the further step of replacing ("decreasing gain") LS learning with constant-gain learning has important implications for monetary policy, even if the equilibrium is stable under learning.

Orphanides and Williams work with a simple two-equation macroeconomic model. The first equation is a new classical expectations-augmented Phillips curve with inertia:

$$\pi_{t+1} = \phi\pi^e_{t+1} + (1-\phi)\pi_t + \alpha y_{t+1} + e_{t+1}, \tag{22}$$

where π_{t+1} is the rate of inflation between period t and period $t+1$, π^e_{t+1} is the rate of inflation over this period expected at time t, y_{t+1} is the level of the output gap in $t+1$, e_{t+1} is a white noise inflation shock, and $(1-\phi)\pi_t$ represents intrinsic inflation persistence. We assume $0 < \phi < 1$.

The second equation is an aggregate demand relation that embodies a lagged policy effect:

$$y_{t+1} = x_t + u_{t+1}.$$

Here, x_t is set by monetary policy at t, and u_{t+1} is white noise. Through monetary policy it is assumed that one period ahead, policymakers are able to control aggregate output up to the unpredictable random disturbance u_{t+1}. This equation basically replaces the IS and LM curves. It is convenient for the task at hand, but suppresses issues of monetary control.

4.1 Optimal Policy under Rational Expectations

At time t the only state variable is π_t. Policymakers have a target inflation rate, π^*, and care about the deviation of π_t from π^*. Their instrument is x_t, and they are assumed to follow a rule of the form,

$$x_t = -\theta(\pi_t - \pi^*). \tag{23}$$

Policymakers also care about the output gap, y_{t+1}. Since stable inflation requires $Ey_t = 0$, policymakers are assumed to choose θ to minimize

$$L = (1 - \omega)Ey_t^2 + \omega E(\pi_t - \pi^*)^2.$$

This is a standard quadratic loss function. We can think of ω as reflecting policymakers preferences, which may (or may not) be derived from the preferences of the representative agent.

Under rational expectations, $\pi_{t+1}^e = E_t\pi_{t+1}$, and it follows that

$$\pi_{t+1}^e = \pi_t + \frac{\alpha}{1-\phi}x_t.$$

Substituting into equation (22) yields

$$\pi_{t+1} = \pi_t + \frac{\alpha}{1-\phi}x_t + \alpha u_{t+1} + e_{t+1}.$$

Substituting in policy rule (23) yields

$$\tilde{\pi}_{t+1} = \left(\frac{1-\phi-\alpha\theta}{1-\phi}\right)\tilde{\pi}_t + \alpha u_{t+1} + e_{t+1},$$

where $\tilde{\pi}_t = \pi_t - \pi^*$.

Computing $E\tilde{\pi}_t^2$ and Ey_t^2, it is straightforward to minimize L over θ to get θ^P, the optimal choice of θ under rational expectations. Orphanides and Williams (2005a) show that

$$\theta^P = \theta^P\left(\omega, \frac{1-\phi}{\alpha}\right),$$

and that θ^P is increasing in both ω and in the degree of inertia, $1-\phi$. Varying ω leads to an efficiency frontier, described by a familiar trade-off between σ_π and σ_y, which is sometimes called the Taylor curve.

For this choice of feedback parameter, in the rational expectations equilibrium inflation follows the process

$$\pi_t = c_0^P + c_1^P \pi_{t-1} + noise_t$$

and

$$E_t \pi_{t+1} = c_0^P + c_1^P \pi_t,$$

where $c_0^P = \alpha\theta^P/(1-\phi)$ and $c_1^P = 1-[\alpha\theta^P/(1-\phi)]$. Here $noise_t$ is white noise. The superscript P refers to perfect knowledge, which Orphanides and Williams use as a synonym for rational expectations.

The problem is thus quite straightforward under rational expectations. How "aggressive"policy should be with respect to deviations of inflation from target depends naturally on the structural parameters ϕ and α and on the policymaker preferences as described by ω.

4.2 Least-Squares Learning

We now make the crucial step of backing away from rational expectations. Instead of assuming that agents are endowed a priori with rational expectations, we model the agents as forecasting in the same way that an econometrician might: by assuming a simple time series model for the variable of interest, estimating its parameters, and using the estimated model to forecast. Specifically, suppose private agents believe that inflation follows a first-order autoregressive, or AR(1), process, as it does in a rational expectations equilibrium, but that they do not know c_0^P, c_1^P. Instead they estimate the parameters of

$$\pi_t = c_0 + c_1\pi_t + v_t$$

by a least-squares-type regression, and at time t they forecast

$$\pi_{t+1}^e = c_{0,t} + c_{1,t}\pi_t.$$

The estimates $c_{0,t}$, $c_{1,t}$ are updated as new data become available. We consider two cases for this updating. First, suppose that agents literally do least squares using all the data. We assume that policymakers do not explicitly take account of private agent learning and follow the feedback rule with $\theta = \theta^P$. Then, with infinite memory (that is, no discounting of observations), one can show that

$$c_{0,t}, c_{1,t} \rightarrow c_0^P, c_1^P$$

with probability 1. Asymptotically, we get the optimal rational expectations equilibrium.

Orphanides and Williams (2005a) make a small but significant change to the standard least squares updating formula. With regular LS, each data point counts equally. When expressed in terms of a recursive algorithm (that is, recursive least squares, or RLS), the coefficient estimates $c_{0,t}$, $c_{1,t}$ are updated in response to the most recent data point with a weight proportion to the sample size $1/t$. We often say that RLS has a decreasing gain since the gain, or weight, on each data point is $\kappa_t = 1/t$, which declines towards 0 as $t \rightarrow \infty$. Orphanides and Williams instead consider constant-gain RLS, in which past data is discounted. In terms of the RLS algorithm, this is accomplished technically by setting the gain—the weight on the most recent observation used to update estimates—to a small constant, that is, by setting $\kappa_t = \kappa$ (for example, 0.05). This is equivalent to using weighted least squares with weights declining geometrically in time as we move backward from the current date.

Why would it be natural for agents to use a constant rather than decreasing gain? The main rationale for this procedure is that it allows estimates to remain alert to structural shifts. As economists, and as econometricians, we tend to believe that structural changes occasionally occur, and we might therefore assume that private agents also recognize and allow for this. Although in principle one might attempt to model the process of structural change, this tends to unduly strain the amount of knowledge we have about the economic structure. A reasonable alternative is to adjust parameter estimators to reflect the fact that recent observations convey more accurate information on the economy's law of motion than do data further in the past, and constant-gain estimators are one very natural way of accomplishing this down-weighting of past data. Another approach that is sometimes used in practice is to implement a rolling data-window of finite length.[24]

24. Honkapohja and Mitra (2003) discuss the implications of bounded memory as a model of learning.

4.3 Implications of Constant-Gain Least Squares

With constant-gain procedures, estimates no longer fully converge to the rational expectations equilibrium. The estimators $c_{0,t}$, $c_{1,t}$ converge instead to a stochastic process. Orphanides and Williams (2005a) therefore use the term perpetual learning to refer to the constant-gain case.

If the gain parameter κ is very small, then estimators will be close to the equilibrium values most of the time with a high probability, and output and inflation will be near their equilibrium paths. Nonetheless, small plausible values like $\kappa = 0.05$ can lead to very different outcomes in the calibrations Orphanides and Williams consider. They analyze the results using simulations, with $\phi = 0.75$ and $\alpha = 0.25$. They consider $\theta \in \{0.1, 0.6, 1.0\}$, which corresponds to weights $\omega = 0.01$, 0.50, and 1.00, respectively, under rational expectations.

Their main findings are threefold. First, the standard deviations of $c_{0,t}$ and $c_{1,t}$ are large even though forecast performance remains good. Second, the persistence of inflation increases substantially, compared with the rational expectations equilibrium, as measured by the AR(1) coefficient for π_t. Finally, the policy frontier shifts out very substantially and sometimes in a nonmonotonic way.

4.4 Policy Implications

Under perpetual learning by private agents, if policymakers keep to the same class of rules,

$$x_t = -\theta^S(\pi_t - \pi^*),$$

then they should choose a different θ than under rational expectations. Here the notation θ^S indicates that we restrict policymakers to choosing from the same "simple" class of policy rules. There are four main implications for policy in the context of constant-gain (perpetual) learning by private agents. First, the "naive" policy choice, that is, the policy that assumes rational expectations (perfect knowledge) on the part of agents, can be strictly inefficient when the agents are, in fact, following perpetual learning with $\kappa > 0$: there are cases in which increasing θ^S above θ^P would decrease the standard deviations of both inflation and output. Second, policy should generally be more hawkish—that is, under perpetual learning the monetary authorities should pick a larger θ^S than if agents had rational expectations.

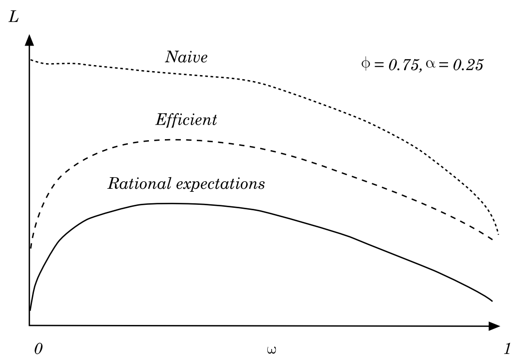
Third, following a sequence of unanticipated inflation shocks, inflation doves (that is, policymakers with a low θ, reflecting a low ω) can do very poorly, as these shocks can lead expectations to temporarily but persistently deviate substantially from rational expectations. Finally, if the inflation target, π^*, is known to private agents, so that they need estimate only the slope parameter c_1 using the PLM,

$$\pi_{t+1} - \pi^* = c_1(\pi_t - \pi^*) + v_{t+1},$$

then the policy frontier is more favorable than when the intercept c_0 is not known. One way to interpret this is that central bank transparency is useful.

Figure 3 indicates how the performance of policy depends on expectations formation and what the policymaker assumes about it. The middle curve is the efficient policy under learning, while "naive" refers to the case in which policy presumes rational expectations while agents are in fact learning with gain $\kappa = 0.05$.

Figure 3. The Policymaker's Loss



Source: Authors' drawing, adapted from Orphanides and Wiliams (2005a), figure 7.

Perpetual learning thus turns out to have major implications for policy, even when the deviation from the rational expectations equilibrium might not be thought to be too large. The main policy implication is that with perpetual learning, there should be a policy bias toward hawkishness. The intuition for this result that a more hawkish policy (high θ) helps to keep inflation expectations, π^e_{t+1}, in line, or closer to rational expectations values. This qualitative result also emerges in the more general setting in Orphanides and Williams (2007).

5. Estimated Models with Learning

The Orphanides and Williams (2005a) results suggest another implication of learning that goes beyond policy, namely, that learning itself can be a source of persistence in macroeconomic dynamics. The starting point for this line of thought, as pursued by Milani (2005, 2007), is that inflation persistence in the data is much higher than arises from the basic New-Keynesian model. For a good empirical fit to the data, a backward-looking component is needed in the New-Keynesian Phillips curve under the rational expectations assumption. The source of the backward-looking component used in these hybrid models, however, is controversial. Milani (2005) considers the question of whether learning dynamics can provide some or all of the persistence needed to fit the data.

To investigate this, consider the most frequently used modification to the basic New-Keynesian model, namely, adding indexation to a Calvo price setting; that is, firms that do not optimize in any given period set prices that are indexed to past inflation. This yields

$$\pi_t = \frac{\gamma}{1+\beta\gamma}\pi_{t-1} + \frac{\beta}{1+\beta\gamma}E_t^*\pi_{t+1} + \frac{\delta}{1+\beta\gamma}x_t + u_t,$$

where x_t is the output gap and γ measures the degree of indexation. Earlier work under rational expectations empirically finds values of γ that are close to one.

For expectations, we assume a PLM of the form

$$\pi_t = \phi_0 + \phi_1\pi_{t-1} + \varepsilon_t,$$

and agents at t are assumed to use data $\{1,\pi_i\}_0^{t-1}$ to estimate ϕ_0, ϕ_1 using constant-gain least squares. For time t estimates $\phi_{0,t}$, $\phi_{1,t}$, the agents' forecasts are given by

$$\begin{aligned} E_t^*\pi_{t+1} &= \phi_{0,t} + \phi_{1,t}E_t^*\pi_t \\ &= \phi_{0,t} + \phi_{1,t}(\phi_{0,t} + \phi_{1,t}\pi_{t-1}), \end{aligned}$$

where we assume that the aggregate inflation rate, π_t, is not included in the agents information set at the time of their forecasts.

The implied ALM is

$$\pi_t = \frac{\beta\phi_{0,t}(1+\phi_{1,t})}{1+\beta\gamma} + \frac{\gamma+\beta\phi_{1,t}^2}{1+\beta\gamma}\pi_{t-1} + \frac{\delta}{1+\beta\gamma}x_t + u_t.$$

Alternatively, Milani (2005) also considers using real marginal cost as the driving variable in place of the output gap, x_t. To estimate the model for the United States, Milani computes inflation from the GDP deflator and the output gap as detrended GDP, while real marginal costs are proxied by the deviation of the labor income share from 1960:1 to 2003:4. Agents' initial parameter estimates are obtained by using presample data from 1951–59.

A two-step procedure is used. First, the PLM is estimated from constant-gain learning using an assumed constant gain of $\kappa = 0.015$. This is in line with earlier empirical estimates. Milani then estimates the ALM using nonlinear least squares. This procedure allows us to estimate the structural source of persistence, γ, taking into account the learning effects. The PLM parameter estimates show the following pattern: $\phi_{1,t}$ was initially low in the 1950s and 1960s, before rising (up to 0.958) and then declining somewhat to values above 0.8; $\phi_{0,t}$ was also initially low before rising sharply and then gradually declining after 1980.

The ALM structural estimates, in particular, generate a degree of indexation of $\gamma = 0.139$ (with the output gap). The results are fairly robust to other choices of gain κ that appear appropriate based on Schwartz's Bayesian information criterion. The estimate of γ is not significantly different from zero, and it constrasts sharply with the high levels of γ found under the rational expectations assumption. It thus appears that the data are consistent with the learning interpretation of the sources of persistence for inflation.

Milani (2007) estimates the full New-Keynesian model under learning. He finds that the degree of habit persistence is also low in the IS curve. This contrasts with the usual extension of the New-Keynesian model under rational expectations that is often employed to improve the empirical fit of the model. Milani's work can be seen as a starting point for the very recent attempts by Slobodyan and Wouters (2007) and Murray (2007) to incorporate learning into DSGE models.

6. Recurrent Hyperinflations

Marcet and Nicolini (2003) start from the standard hyperinflation model with learning and extend it to an open economy setting. Their aim is to provide a unified theory to explain the recurrent hyperinflations experienced by many countries in the 1980s.

6.1 The Basic Hyperinflation Model

The starting point is the theoretical model sometimes known as the seigniorage model of inflation (see Evans and Honkapohja, 2001, chap. 11). The Cagan model is based on the linear money demand equation, which can be obtained from an overlapping generations (OG) endowment economy with log utility. Specifically,

$$\frac{M_t^d}{P_t} = \phi - \phi\gamma \frac{P_{t+1}^e}{P_t},$$

if $1 - \gamma(P_{t+1}^e / P_t) > 0$ and 0 otherwise. This equation is combined with exogenous government purchases, $d_t > 0$, that are entirely financed by seigniorage:

$$M_t = M_{t-1} + d_t P_t.$$

Rewriting this as $M_t/P_t = (M_{t-1}/P_{t-1})(P_{t-1}/P_t) + d$, setting $M_t^d = M_t$, and assuming $d_t = d$, we get

$$\frac{P_t}{P_{t-1}} = \frac{1 - \gamma(P_t^e / P_{t-1})}{1 - \gamma(P_{t+1}^e / P_t) - d/\phi}.$$

Under perfect foresight (that is, $P_{t+1}^e / P_t = P_{t+1} / P_t$) there are two steady states, $\beta_L < \beta_H$, provided $d \geq 0$ is not too large. If d is above a critical value, then there are no perfect foresight steady states. There is also a continuum of perfect foresight paths converging to β_H. Some early theorists suggested that these paths might provide an explanation for actual hyperinflation episodes.

Consider now the situation under adaptive learning. Suppose the PLM is that the inflation process is a steady state, that is, $P_{t+1} / P_t = \beta + \eta_t$, where η_t is perceived white noise. Then PLM expectations are

$$\left(\frac{P_{t+1}}{P_t}\right)^e = \beta,$$

and the corresponding ALM is

$$\frac{P_t}{P_{t-1}} = \frac{1-\gamma\beta}{1-\gamma\beta-d/\phi} \equiv T(\beta;d).$$

Under steady-state learning, agents estimate β based on past average inflation, that is, $(P_{t+1}/P_t)^e = \beta_t$, where

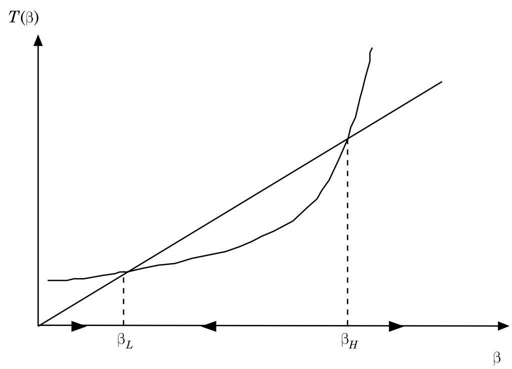
$$\beta_t = \beta_{t-1} + t^{-1}\left(\frac{P_{t-1}}{P_{t-2}} - \beta_{t-1}\right).$$

This is simply a recursive algorithm for the average inflation rate, which is equivalent to a least-squares regression on a constant.[25] The stability of this learning rule is governed by the E-stability differential equation

$$\frac{d\beta}{d\tau} = T(\beta;d) - \beta,$$

where d is a fixed parameter. Since $0 < T'(\beta_L) < 1$ and $T'(\beta_H) > 1$, β_L is E-stable, and therefore locally stable under learning, while β_H is not. This is illustrated in figure 4.

Figure 4. Steady-State Learning in the Hyperinflation Model



Source: Authors' drawing.

25. One can consider more general classes of PLM. Adam, Evans, and Honkapohja (2006) study the circumstances in which autoregressive PLMs can converge to hyperinflation paths.

An increase in d shifts $T(\beta)$ up, so the comparative statics of β_L are natural while those of β_H are counterintuitive. This, together with the fact that the steady state β_H is not stable under learning, suggests problems with the rational expectations version of this model as a theoretical explanation for hyperinflations.

6.2 Empirical Background

Marcet and Nicolini (2003) list four stylized facts about hyperinflation episodes in the 1980s in a number of South American countries (as well as some episodes in other places and at other times): (1) hyperinflation episodes are recurrent; (2) exchange rate rules stop hyperinflations, although new hyperinflations eventually occur; (3) during a hyperinflation, seigniorage and inflation are not highly correlated; and (4) average inflation and seigniorage are strongly positively correlated across countries, with hyperinflations only occurring in countries where seigniorage is high, on average. Stabilization plans to deal with hyperinflation have been based either on heterodox policy (exchange rate rules) or orthodox policy (permanently reducing the deficit). Policies that combine both elements appear to have been successful in stopping hyperinflations permanently.

6.3 The Marcet-Nicolini Model

Marcet and Nicolini (2003) use an open economy version of the overlapping-generations hyperinflation model. This is a flexible price model with purchasing power parity (PPP), so that

$$P_t^f e_t = P_t,$$

where P_t^f is the foreign price of goods, which is assumed to be exogenous. There is a cash-in-advance constraint for local currency on net purchases of consumption. This generates the demand by young agents for the local currency. Hence, we continue to have the money demand equation as in the basic model. Government expenditure, d_t, is assumed to be i.i.d.

There are two exchange rate regimes. In the floating regime the government does not buy or sell foreign exchange, and its budget constraint is as in the basic model. There is no foreign trade, and the economy behaves just like the closed economy model, with PPP determining the price of foreign currency by $e_t = P_t / P_t^f$.

In the exchange rate rule regime, the government buys or sells foreign exchange, R_t, as needed to meet a target exchange rate, e_t. Sales of foreign exchange generate revenue in addition to seigniorage that the government can use to finance government purchases, that is, $(M_t - M_{t-1})/P_t = d_t + [(R_t - R_{t-1})e_t\ ]/P_t$. In equilibrium, any increase in reserves must be matched by a trade surplus, that is, $(R_t - R_{t-1})e_t = TB_t{\cdot}P_t$, where TB_t is total endowment minus total private consumption minus d_t.

The key question is the form of the exchange rate rule. When an exchange rate rule is adopted, it is assumed that the object is to stabilize inflation at a targeted rate, $\overline{\beta}$. This is accomplished by setting e_t to satisfy

$$\frac{P_t^f}{P_{t-1}^f}\frac{e_t}{e_{t-1}} = \overline{\beta},$$

which by PPP guarantees

$$\frac{P_t}{P_{t-1}} = \overline{\beta}.$$

Under the exchange rate rule, this last equation determines P_t. Given expectations, money demand determines M_t. Reserves, R_t, must then adjust to satisfy the flow government budget constraint.

The remaining question is how the government chooses its exchange rate regime. We assume there is a maximum inflation rate tolerated, β_U. The exchange rate regime is imposed only in periods when inflation would otherwise exceed this bound (or if no positive P_t would otherwise clear the market).

6.4 Learning

Marcet and Nicolini (2003) argue that under rational expectations, the model cannot properly explain the stylized facts of hyperinflation outlined above. An adaptive learning formulation will be more successful. They use a variation of the simple (decreasing gain) steady-state learning rule, given above, in which the gain is made state contingent:

$$\beta_t = \beta_{t-1} + \frac{1}{\alpha_t}\left(\frac{P_{t-1}}{P_{t-2}} - \beta_{t-1}\right),$$

with a given β_0. Here $1/\alpha_t = \kappa_t$ is what we have called the gain, $\alpha_t = \alpha_{t-1} + 1$ corresponds to decreasing gain learning, and $\alpha_t = \bar{\alpha} > 1$ is a constant-gain algorithm (α_t can also be thought of as the effective sample size). Marcet and Nicolini consider a version in which agents switch between decreasing and constant gain according to recent performance. Specifically,

$$\alpha_t = \alpha_{t-1} + 1 \quad \text{if} \quad \left| \left(\frac{P_{t-1}}{P_{t-2}} - \beta_{t-1} \right) \Big/ \beta_{t-1} \right|$$

falls below some bound v, and otherwise $\alpha_t = \bar{\alpha}$.

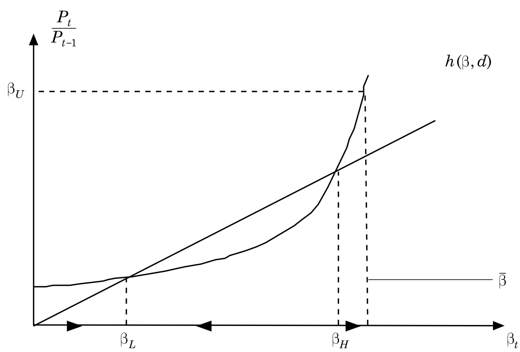
The qualitative features of the model are approximated by the system

$$\frac{P_t}{P_{t-1}} = h(\beta_{t-1}, d_t), \tag{24}$$

where $h(\beta, d) = \begin{cases} T(\beta, d) & \text{if} \quad 0 < T(\beta, d) < \beta_U \\ \bar{\beta} & \text{otherwise} \end{cases}$

Figure 5 describes the dynamics of system (24).

Figure 5. Inflation as a Function of Expected Inflation



Source: Authors' drawing, adapted from Marcet and Nicolini (2003), figure 3.

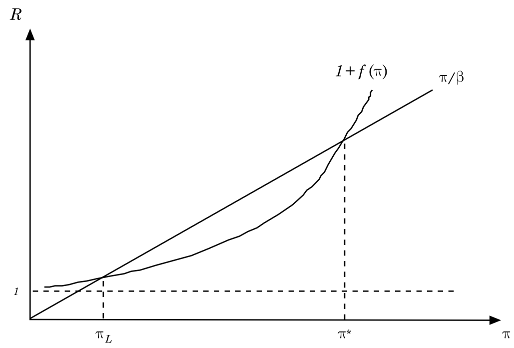
There is a stable region consisting of values of β below the unstable high inflation steady state, β_H, and an unstable region that lies above it. Here we set $\overline{\beta} = \beta_L$, the low inflation steady state. β_U is set at a value above β_H. This gives rise to very natural recurring hyperinflation dynamics: starting from β_L, a sequence of random shocks may push β_t into the unstable region, at which point the gain is revised upward to $1/\overline{\alpha}$ and inflation follows an explosive path until it is stabilized by the exchange rate rule. Then the process begins again. The model with learning has the following three features. First, there may be eventual convergence to rational expectations. This can occur if the random shocks/learning dynamics do not push β_t into the unstable region for a long time. Decreasing gain may then lead to asymptotic convergence to β_L. Second, a higher $E(d_t)$ raises both average inflation and the frequency of hyperinflations. A combination of orthodox and heterodox policies make sense as a way to end hyperinflations. Third, all four stylized facts listed above can be matched using this model, and simulations of a calibrated model look very plausible. Overall this appears to be a very successful application of boundedly rational learning to a major empirical issue.

7. Liquidity Traps and Deflationary Spirals

Deflation and liquidity traps have been a concern in recent times. Evans, Guse, and Honkapohja (2008) consider issues of liquidity traps and deflationary spirals under learning in a New-Keynesian model. As we have seen, contemporaneous Taylor-type interest rate rules should respond to the inflation rate more than one for one to ensure determinacy and stability under learning. As emphasized by Benhabib, Schmitt-Grohé, and Uribe (2001), however, if one considers the interest rate rule globally, rather than in a neighborhood of the target inflation rate, the requirement that net nominal interest rates must be nonnegative implies that the rule must be nonlinear and that, for any continuous rule, there exists a second steady state at a lower (possibly negative) inflation rate. This is illustrated in figure 6, which shows the interest rate policy $R = 1 + f(\pi)$ as a function of π.[26] The straight line in the figure is the Fisher equation, $R = \pi/\beta$, which is obtained from the usual Euler equation for consumption in a steady state.

26. Taylor rules usually also include a dependence on aggregate output, which we omit for simplicity.

Figure 6. Multiple Steady States with a Global Taylor Rule



Source: Authors' drawing.

Here we are now using R to stand for the interest rate factor (so that the net interest rate is $R - 1$), and $\pi_t = P_t/P_{t-1}$ is the inflation factor, so that $\pi - 1$ is the net inflation rate. In the figure, π^* denotes the intended steady state, at which the Taylor principle of a more than one-for-one response is satisfied, and π_L is the unintended steady state. In addition, π_L may correspond to either a very low positive inflation rate or to a negative net inflation rate, that is, deflation. The zero lower bound corresponds to $R = 1$. Benhabib, Schmitt-Grohé, and Uribe (2001) show that under rational expectations, there is a continuum of liquidity trap paths that converge on π_L. The pure rational expectations analysis thus suggests a serious risk of the economy following these liquidity trap paths.

What happens under learning? In Evans and Honkapohja (2005), we analyzed a flexible-price perfect competition model. We show that deflationary paths are possible, but that the real risk, under learning, involves paths in which inflation slips below π_L and then continue to fall further. For this flexible-price model, we show that this can be avoided by a change in monetary policy at low inflation rates. The required policy is to switch to an aggressive money supply rule at some inflation rate between π_L and π^*. Such a policy would successfully avoid liquidity traps and deflationary paths.

Evans, Guse, and Honkapohja (2008) reconsider the issues in a model that allows for sticky prices and deviations of output from flexible-price levels. They consider a representative-agent infinite-horizon dynamic stochastic general equilibrium model with monopolistic competition and price-adjustment costs. Monetary policy follows a global Taylor-rule as above. Fiscal policy is standard: exogenous government purchases, g_t, and Ricardian tax policy that depends on real debt level. The model is essentially a New-Keynesian model, except that, in line with Benhabib, Schmitt-Grohé, and Uribe (2001), it has Rotemberg (1982) costs of price adjustment as the friction rather than Calvo pricing. The model equations are nonlinear, and the nonlinearity in its analysis under learning is retained.

The key equations are

$$\frac{\alpha\gamma}{\nu}(\pi_t - 1)\pi_t = \beta\frac{\alpha\gamma}{\nu}(\pi^e_{t+1} - 1)\pi^e_{t+1} + (c_t + g_t)^{(1+\varepsilon)/\alpha} \\ -\alpha\left(1 - \frac{1}{\nu}\right)(c_t + g_t)c_t^{-\sigma_1}$$

and

$$c_t = c^e_{t+1}\left(\frac{\pi^e_{t+1}}{\beta R_t}\right)^{\sigma_1}.$$

The first equation is the New-Keynesian Phillips curve, relating π_t positively to π^e_{t+1} and to measures of aggregate activity. The second equation is the New-Keynesian IS curve, obtained from the usual household Euler equation. When linearized around a steady state, both of these equations are identical in form to the standard New-Keynesian equations. There are also money and debt evolution equations.

There are two stochastic steady states at π_L and π_H. If the random shocks are i.i.d., then steady-state learning is appropriate for both c^e and π^e, that is,

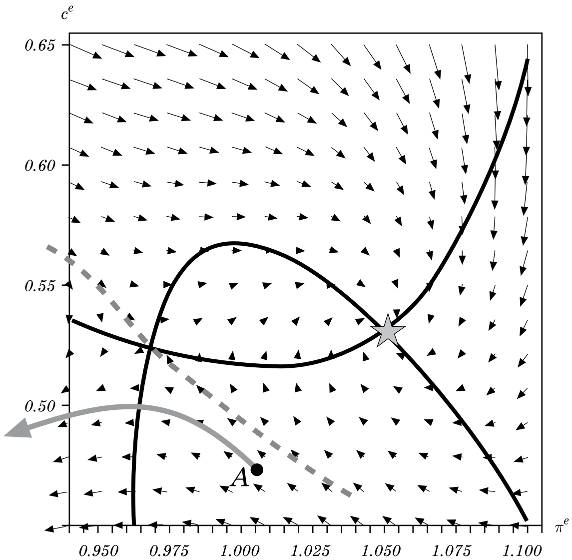
$$\pi^e_{t+1} = \pi^e_t + \phi_t(\pi_{t-1} - \pi^e_t)$$

and

$$c^e_{t+1} = c^e_t + \phi_t(c_{t-1} - c^e_t),$$

where ϕ_t is the gain sequence. The main findings are that while the intended steady state at π^* is locally stable under learning, the unintended steady state at π_L is unstable under learning. The key observation is that π_L is a saddlepoint, which implies the existence of deflationary spirals under learning. In particular, an expectational shock can lead to sufficiently pessimistic expectations, and c^e and π^e will follow paths leading to deflation and stagnation. This is illustrated in figure 7, based on E-stability dynamics.

Figure 7. The Dynamics of π^e and c^e under Normal Policy



Source: Based on Evans, Guse, and Honkapohja (2008), figure 4.

The intuition for the result can be seen by supposing that we are initially near the π_L steady state and considering a small drop in π^e. With a fixed R this would lead through the IS curve to a lower c and thus through the Phillips curve, to a lower π. A sufficient reduction in R would prevent the reductions in c and π, but this is not possible since we are close to the zero lower bound, and the global Taylor rule here dictates only small reductions in R. The falls in realized c and π then leads under learning to reductions in c^e and π^e, and this sets the deflationary spiral in motion.

Thus, under normal policy the intended steady state is not globally stable under learning. Large adverse shocks to expectations or structural changes can set in motion unstable downward paths.

Evans, Guse, and Honkapohja (2008) show that policy can be altered to avoid the deflationary spiral. The recommended policy is to set a minimum inflation threshold $\tilde{\pi}$, where $\pi_L < \tilde{\pi} < \pi^*$. For example, if the global Taylor rule is chosen so that π_L corresponds to deflation, then a convenient choice for the threshold would be zero net inflation, $\tilde{\pi} = 1$. The authorities would follow normal monetary and fiscal policy provided this delivers $\pi_t > \tilde{\pi}$. However, if π_t threatens to fall below $\tilde{\pi}$ under normal policy, then aggressive policies would be implemented to ensure that $\pi_t = \tilde{\pi}$: interest rates would be reduced, if necessary to near the zero lower bound $R = 1$, and if this is not sufficient, then government purchases, g_t, would be increased as required.

Evans, Guse, and Honkapohja (2008) show that these policies can indeed ensure $\pi_t \geq \tilde{\pi}$ always under learning, and that incorporating aggressive monetary and fiscal policies triggered by an inflation threshold $\tilde{\pi}$ leads to global stability of the intended steady state at π^*.[27] Perhaps surprisingly, they also show that it is essential to use an inflation threshold, since using an output threshold to trigger aggressive polices will not always avoid deflationary spirals.

8. Conclusions

Expectations play a large role in modern macroeconomics. While the rational expectations assumption is the natural benchmark, it is implausibly demanding. Realistically, it should be assumed that people are smart, but boundedly rational. To model bounded rationality, we recommend the principle of cognitive consistency: economic agents should be about as smart as (good) economists. When economists need to make forecasts, they do so using econometric models, so a particularly natural choice is to model agents as econometricians.

Convergence to rational expectations is possible in many economic models, with an appropriate econometric perceived law of motion. However, the stability of rational expectations equilibrium under private agent learning is not automatic. Our central message is that monetary policy must be designed to ensure both determinacy and stability under learning. This observation leads to particular choices of interest rate rules, whether we are considering standard classes of instrument rules or designing optimal monetary policy. Instrument rules that respond appropriately to "nowcasts" perform

27. For non-Ricardian economies, Bénassy (2007) develops an alternative interest rate rule that leads to global uniqueness.

well in this respect, but implementing optimal policy appears to require an appropriate response to private sector expectations about the future.

More generally, policymakers need to use policy to guide expectations, and the recent literature provides several important illustrations. If under learning there are persistent deviations from fully rational expectations, then monetary policy may need to respond more aggressively to inflation in order to stabilize expectations. The learning literature also shows how to guide the economy under extreme threats of either hyperinflation or deflationary spirals. As we have illustrated, appropriate monetary and fiscal policy design can minimize these risks.

References

Adam, K., G.W. Evans and S. Honkapohja. 2006. "Are Hyperinflation Paths Learnable?" *Journal of Economic Dynamics and Control* 30(12): 2725–48.

Arifovic, J., J. Bullard and O. Kostyshyna. 2007. "Social Learning and Monetary Policy Rules." Working paper 2007-007a. Federal Reserve Bank of St Louis.

Bank of England. 2007. *Inflation Report*. London.

Basdevant, O. 2005. "Learning Process and Rational Expectations: An Analysis Using a Small Macroeconomic Model for New Zealand." *Economic Modelling* 22(6): 1074–89.

Bénassy, J-P. 2007. *Money, Interest, and Policy: Dynamic General Equilibrium in a non-Ricardian World*. MIT Press.

Benhabib, J., S. Schmitt-Grohé and M. Uribe. 2001. "The Perils of Taylor Rules." *Journal of Economic Theory* 96(1–2): 40–69.

Berardi, M. 2008. "Monetary Policy with Heterogeneous and Misspecified Expectations." *Journal of Money, Credit, and Banking* (forthcoming).

Berardi, M. and J. Duffy. 2007. "The Value of Central Bank Transparency When Agents Are Learning." *European Journal of Political Economy* 23(1): 9–29.

Bernanke, B. 2007. "Inflation Expectations and Inflation Forecasting." Speech given at the Monetary Economics Workshop. National Bureau of Economic Research, Summer Institute, Cambridge, Massachusetts, 10 July.

Bernanke, B. and M. Woodford. 1997. "Inflation Forecasts and Monetary Policy." *Journal of Money, Credit, and Banking* 29(4): 653–84.

Branch, W.A. 2004. "The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations." *Economic Journal* 114(497): 592–621.

Branch, W.A., J. Carlson, G.W. Evans and B. McGough. 2007. "Adaptive Learning, Endogenous Inattention, and Changes in Monetary Policy." Working paper 2006-6, revised. University of Oregon.

———. 2008. "Monetary Policy, Endogenous Inattention, and the Volatility Trade-Off." *Economic Journal* (forthcoming).

Branch, W.A. and G.W. Evans. 2006a. "A Simple Recursive Forecasting Model." *Economic Letters* 91(2): 158–66.

———. 2006b. "Intrinsic Heterogeneity in Expectation Formation." *Journal of Economic Theory* 127(1): 264–95.

———. 2007. "Model Uncertainty and Endogenous Volatility." *Review of Economic Dynamics* 10(2): 207–37.

Brazier, A., R. Harrison, M. King and T. Yates. 2008. "The Danger of Inflating Expectations of Macroeconomic Stability: Heuristic Switching in an Overlapping Generations Monetary Model." *International Journal of Central Banking* 4(2): 219–54.

Brock, W.A., and C.H. Hommes. 1997. "A Rational Route to Randomness." *Econometrica* 65(5): 1059–95.

Bullard, J. 2006. "The Learnability Criterion and Monetary Policy." *Federal Reserve Bank of St. Louis Review* 88 (May): 203–17.

Bullard, J. and I-K. Cho. 2005. "Escapist Policy Rules." *Journal of Economic Dynamics and Control* 29(11): 1841–66.

Bullard, J. and S. Eusepi. 2005. "Did the Great Inflation Occur Despite Policymaker Commitment to a Taylor Rule?" *Review of Economic Dynamics* 8(2): 324–59.

Bullard, J., G.W. Evans and S. Honkapohja. 2008. "Monetary Policy, Judgment, and Near-Rational Exuberance." *American Economic Review* 98(3): 1163–77.

Bullard, J. and K. Mitra. 2002. "Learning About Monetary Policy Rules." *Journal of Monetary Economics* 49(6): 1105–29.

———. 2007. "Determinacy, Learnability, and Monetary Policy Inertia." *Journal of Money, Credit, and Banking* 39(5): 1177–212.

Bullard, J. and E. Schaling. 2006. "Monetary Policy, Determinacy, Learnability in a Two-Block World Economy." Working paper 2006-038a. Federal Reserve Bank of St Louis.

Bullard, J. and A. Singh. 2006. "Worldwide Macroeconomic Stability and Monetary Policy Rules." Working paper 2006-040b. Federal Reserve Bank of St Louis.

Calvo, G.A. 1983. "Staggered Prices in a Utility-Maximizing Framework." *Journal of Monetary Economics* 12(3): 383–98.

Carboni, G. and M. Ellison. 2007. "Learning and the Great Inflation." Discussion paper 6250. London: Centre for Economic Policy Research.

———. 2008. "The Great Inflation and the Greenbook." Frankfurt: European Central Bank.

Carlstrom, C.T. and T.S. Fuerst. 2004. "Learning and the Central Bank." *Journal of Monetary Economics* 51(2): 327–38.

Clarida, R., J. Galí and M. Gertler. 1998. "Monetary Policy Rules in Practice: Some International Evidence." *European Economic Review* 42(6): 1033–67.

———. 1999. "The Science of Monetary Policy: A New Keynesian Perspective." *Journal of Economic Literature* 37(4): 1661–707.

Cogley, T. and T.J. Sargent. 2005. "The Conquest of U.S. Inflation: Learning and Robustness to Model Uncertainty." *Review of Economic Dynamics* 8(2): 528–63.

Dennis, R. and F. Ravenna. 2008. "Learning and Optimal Monetary Policy." *Journal of Economic Dynamics and Control* 32(6): 1964–94.

Duffy, J. and W. Xiao. 2007a. "Investment and Monetary Policy: Learning and Determinacy of Equilibria." Working paper 324. University of Pittsburgh, Department of Economics.

———. 2007b. "The Value of Interest Rate Stabilization Policies When Agents are Learning." *Journal of Money, Credit, and Banking* 39(8): 2041–56.

Ellison, M. and T. Yates. 2007. "Escaping Volatile Inflation." *Journal of Money, Credit, and Banking* 39(4): 981–93.

European Central Bank. 2007. *Monthly Bulletin* (June and December). Frankfurt.

Eusepi, S. 2007. "Learnability and Monetary Policy: A Global Perspective." *Journal of Monetary Economics* 54(4): 1115–31.

Eusepi, S. and B. Preston. 2007. "Central Bank Communication and Expectations Stabilization." Working paper 13259. Cambridge, Mass.: National Bureau of Economic Research.

Evans, G.W., E. Guse and S. Honkapohja. 2008. "Liquidity Traps, Learning, and Stagnation." *European Economic Review* (forthcoming).

Evans, G.W. and S. Honkapohja. 1993. "Adaptive Forecasts, Hysteresis, and Endogenous Fluctuations." *Federal Reserve Bank of San Francisco Economic Review* 1993(1): 3–13.

———. 1995. "Adaptive Learning and Expectational Stability: An Introduction." In *Learning and Rationality in Economics*, edited by A. Kirman and M. Salmon, chap. 4. Oxford: Basil Blackwell.

———. 1999. "Learning Dynamics." In *Handbook of Macroeconomics*, vol. 1, edited by J. Taylor and M. Woodford, chap. 7. Amsterdam: Elsevier.

———. 2001. *Learning and Expectations in Macroeconomics*. Princeton University Press.

————. 2003a. "Adaptive Learning and Monetary Policy Design." *Journal of Money, Credit, and Banking* 35(6): 1045–72.

————. 2003b. "Expectational Stability of Stationary Sunspot Equilibria in a Forward-looking Model." *Journal of Economic Dynamics and Control* 28(1): 171–81.

————. 2003c. "Expectations and the Stability Problem for Optimal Monetary Policies." *Review of Economic Studies* 70(4): 807–24.

————. 2003d. "Friedman's Money Supply Rule vs. Optimal Interest Rate Policy." *Scottish Journal of Political Economy* 50(5) 550–66.

————. 2005. "Policy Interaction, Expectations, and the Liquidity Trap." *Review of Economic Dynamics* 82 303–23.

————. 2006. "Monetary Policy, Expectations, and Commitment." *Scandinavian Journal of Economics* 108(1): 15–38.

————. 2007. "Policy Interaction, Learning, and the Fiscal Theory of Prices." *Macroeconomic Dynamics* 11(5): 665–90.

————. 2008. "Robust Learning Stability with Operational Monetary Policy Rules." In *Monetary Policy Under Uncertainty and Learning* edited by K. Schmidt-Hebbel and C.E. Walsh, XX-YY. Central Bank of Chile.

Evans, G.W., S. Honkapohja and R. Marimon. 2001. "Convergence in Monetary Inflation Models with Heterogeneous Learning Rules." *Macroeconomic Dynamics* 5(1): 1–31.

————. 2007. "Stable Sunspot Equilibria in a Cash-in-Advance Economy." *The Berkeley Electronic Journal of Macroeconomics (Advances)* 7(1): 1165.

Evans, G.W. and B. McGough. 2005a. "Monetary Policy, Indeterminacy, and Learning." *Journal of Economic Dynamics and Control* 29(11): 1809–40.

————. 2005b. "Stable Sunspot Solutions in Models with Predetermined Variables." *Journal of Economic Dynamics and Control* 29(4): 601–25.

————. 2007. "Optimal Constrained Interest-Rate Rules." *Journal of Money, Credit, and Banking* 39(6): 1335–56.

Evans, G.W. and G. Ramey. 2006. "Adaptive Expectations, Underparameterization, and the Lucas Critique." *Journal of Monetary Economics* 53(2): 249–64.

Ferrero, G. 2007. "Monetary Policy, Learning, and the Speed of Convergence." *Journal of Economic Dynamics and Control* 31(9): 3006–41.

Gaspar, V., F. Smets and D. Vestin. 2005. "Optimal Monetary Policy Under Adaptive Learning." Bank of Portugal.

————. 2006. "Adaptive Learning, Persistence, and Optimal Monetary Policy." *Journal of the European Economic Association* 4(2–3): 376–85.

Giannitsarou, C. 2003. "Heterogeneous Learning." *Review of Economic Dynamics* 6(4): 885–906.

Guse, E. 2008. "Learning in a Misspecified Multivariate Self-Referential Linear Stochastic Model." *Journal of Economic Dynamics and Control* 32(5): 1517–42.

Honkapohja, S. and K. Mitra. 2003. "Learning with Bounded Memory in Stochastic Models." *Journal of Economic Dynamics and Control* 27(8): 1437–57.

————. 2004. "Are Non-Fundamental Equilibria Learnable in Models of Monetary Policy?" *Journal of Monetary Economics* 51(1): 1743–70.

————. 2005a. "Performance of Inflation Targeting Based on Constant Interest Rate Projections." *Journal of Economic Dynamics and Control* 29(11): 1867–92.

————. 2005b. "Performance of Monetary Policy with Internal Central Bank Forecasting." *Journal of Economic Dynamics and Control* 29(4): 627–58.

————. 2006. "Learning Stability in Economies with Heterogeneous Agents." *Review of Economic Dynamics* 9(2): 284–309.

Kurozumi, T. 2006. "Determinacy and Expectational Stability of Equilibrium in a Monetary Sticky-Price Model with Taylor Rule." *Journal of Monetary Economics* 53(4): 827–46.

Kurozumi, T. and W. Van Zandweghe. 2008a. "Investment, Interest Rate Policy, and Equilibrium Stability." *Journal of Economic Dynamics and Control* 32(5): 1489–516.

————. 2008b. "Labor Market Search and Interest Rate Policy." Tokyo: Bank of Japan.

Llosa, G. and V. Tuesta. 2007. "E-Stability of Monetary Policy When the Cost Channel Matters." Washington: Inter-American Development Bank.

————. 2008. "Determinacy and Learnability of Monetary Policy Rules in Small Open Economies." *Journal of Money, Credit and Banking* 40(5): 1033-63.

Lucas, R.E., Jr. 1976. "Econometric Policy Evaluation: A Critique." In *Carnegie-Rochester Conference Series on Public Policy* 1: 19–46.

Marcet, A. and J.P. Nicolini. 2003. "Recurrent Hyperinflations and Learning." *American Economic Review* 93(5): 1476–98.

Marimon, R. 1997. "Learning from Learning in Economics." In *Advances in Economics and Econometrics: Theory and Applications*, vol. 1, edited by D. Kreps and K. Wallis, chap. 9. Cambridge University Press.

McCallum, B.T. 1999. "Issues in the Design of Monetary Policy Rules." In *Handbook of Macroeconomics*, vol. 1, edited by J. Taylor and M. Woodford, chap. 23. Amsterdam: Elsevier.

————. 2003. "Is the Fiscal Theory of the Price Level Learnable?" *Scottish Journal of Political Economy* 50(5): 634–49.

————. 2007. "E-stability vis-à-vis Determinacy Results for a Broad Class of Linear Rational Expectations Models." *Journal of Economic Dynamics and Control* 31(4): 1376–91.

McCallum, B.T. and E. Nelson. 1999. "Performance of Operational Policy Rules in an Estimated Semi-Classical Model." In *Monetary Policy Rules*, edited by J. Taylor, pp. 15–45. University of Chicago Press.

————. 2004. "Timeless Perspective vs. Discretionary Monetary Policy in Forward-Looking Models." *Federal Reserve Bank of St.Louis Review* 86: 43–56.

McGough, B., G. Rudebusch and J. Williams. 2005. "Using a Long-Term Interest Rate as the Monetary Policy Instrument." *Journal of Monetary Economics* 52(5): 855–79.

Milani, F. 2005. "Adaptive Learning and Inflation Persistence." Working paper 050607. University of California at Irvine, Department of Economics.

————. 2007. "Expectations, Learning, and Macroeconomic Persistence." *Journal Of Monetary Economics* 54(7): 2065–82.

Mitra, K. 2003. "Desirability of Nominal GDP Targeting under Adaptive Learning." *Journal of Money, Credit, and Banking* 35(2): 197–220.

Molnar, K. and S. Santoro. 2006. "Optimal Monetary Policy When Agents Are Learning." Discussion paper 2006/01. Institute of Economics, Hungarian Academy of Sciences.

Murray, J. 2007. "Empirical Significance of Learning in a New Keynesian Model with Firm-Specific Capital." Working paper 2007-027. Indiana University, Center for Applied Economics and Policy Research.

Muto, I. 2008. "Monetary Policy and Learning from the Central Bank's Forecast." Discussion paper 2008-E-1. Tokyo: Bank of Japan.

Orphanides, A. and J.C. Williams. 2005a. "Imperfect Knowledge, Inflation Expectations, and Monetary Policy." In *The Inflation-*

Targeting Debate, edited by B. Bernanke and M. Woodford, chap. 5. University of Chicago Press.

————. 2005b. "Inflation Scares and Forecast-Based Monetary Policy." *Review of Economic Dynamics* 8(2): 498–527.

————. 2005c. "The Decline of Activist Stabilization Policy: Natural Rate Misperceptions, Learning, and Expectations." *Journal of Economic Dynamics and Control* 29(11): 1927–50.

————. 2007. "Robust Monetary Policy with Imperfect Knowledge." *Journal of Monetary Economics* 54(5): 1406–35.

Pfajfar, D. 2007. "Formation of Rationally Heterogeneous Expectations." University of Cambridge.

Pfajfar, D. and E. Santoro. 2007a. "Credit Market Distortions, Asset Prices, and Monetary Policy." University of Cambridge.

————. 2007b. "Heterogeneity and Learning in Inflation Expectation Formation: An Empirical Assessment." University of Cambridge.

Preston, B. 2005. "Learning about Monetary Policy Rules When Long-Horizon Expectations Matter." *International Journal of Central Banking* 1(2): 81–126.

————. 2006. "Adaptive Learning, Forecast-based Instrument Rules, and Monetary Policy." *Journal of Monetary Economics* 53(3): 507–35.

Primiceri, G.E. 2006. "Why Inflation Rose and Fell: Policymakers' Beliefs and U.S. Postwar Stabilization Policy." *Quarterly Journal of Economics* 121(3): 867–901.

Rotemberg, J.J. 1982. "Sticky Prices in the United States." *Journal of Political Economy* 90(6): 1187–211.

Rotemberg, J.J. and M. Woodford. 1999. "Interest Rate Rules in an Estimated Sticky-Price Model." In *Monetary Policy Rules*, edited by J. Taylor, pp. 57–119. University of Chicago Press.

Sargent, T.J. 1993. *Bounded Rationality in Macroeconomics*. Oxford University Press.

————. 1999. *The Conquest of American Inflation*. Princeton University Press.

Sargent, T.J. and N. Wallace. 1975. "'Rational Expectations,' the Optimal Monetary Instrument, and the Optimal Money Supply Rule." *Journal of Political Economy* 83(2): 241–54.

Sargent, T.J., N. Williams and T. Zha. 2006. "Shocks and Government Beliefs: The Rise and Fall of American Inflation." *American Economic Review* 96(4): 1193–224.

Slobodyan, S. and R. Wouters. 2007. "Learning in an Estimated DSGE Model." Prague: CERGE-EI.

Svensson, L.E.O. 1999. "Inflation Targeting as a Monetary Policy Rule." *Journal of Monetary Economics* 43(3): 607–54.

————. 2003. "What Is Wrong with Taylor Rules? Using Judgement in Monetary Policy through Targeting Rules." *Journal of Economic Literature* 41(2): 426–77.

Svensson, L.E.O. and M. Woodford. 2005. "Implementing Optimal Policy through Inflation-Forecast Targeting." In *The Inflation-Targeting Debate*, edited by B. Bernanke and M. Woodford, pp. 19–83. University of Chicago Press.

Taylor, J. 1993. "Discretion versus Policy Rules in Practice." *Carnegie-Rochester Conference Series on Public Policy* 39: 195–214.

Tesfaselassie, M., E. Schaling and S. Eijffinger. 2008. "Learning about the Term Structure and Optimal Rules for Inflation Targeting." Discussion paper 88. Tilburg University, Center for Economic Research.

Tetlow, R.J. and P. von zur Muehlen. 2008. "Robustifying Learnability." *Journal of Economic Dynamics and Control* (forthcoming).

Trichet, J-C. 2005. "Monetary Policy and Private Expectations." Zolotas Lecture at the Bank of Greece. Athens, 25 February.

Walsh, C. 2003. *Monetary Theory and Policy*, 2d ed. MIT Press.

Wang, Q. 2006. "Learning Stability for Monetary Policy Rules in a Two-Country Model." Working paper 659, University of Cambridge.

Wieland, V. 2008. "Learning, Endogenous Indexation, and Disinflation in the New Keynesian Model." Working paper 6749. London: Centre for Economic Policy Research.

Woodford, M. 1999a. "Commentary: How Should Monetary Policy Be Conducted in an Era of Price Stability?" In *New Challenges for Monetary Policy*, edited by the Federal Reserve Bank of Kansas City, pp. 277–316.

————. 1999b. "Optimal Monetary Policy Inertia." *Manchester School* 67: (supplement) 1–35.

————. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

————. 2005. "Robustly Optimal Monetary Policy with Near-Rational Expectations." Working paper 11896. Cambridge, Mass.: National Bureau of Economic Research.

Zanna, L-F. 2006. "PPP Rules, Macroeconomic (In)stability, and Learning." International finance discussion paper 814. Washington: Board of Governors of the Federal Reserve System.

Optimal Monetary Policy under Uncertainty in DSGE Models: A Markov Jump-Linear-Quadratic Approach

Lars E.O. Svensson
Sveriges Riksbank and Princeton University

Noah Williams
University of Wisconsin

Our previous work develops methods to study optimal policy in Markov jump-linear-quadratic (MJLQ) models with forward-looking variables: models with conditionally linear dynamics and conditionally quadratic preferences, where the matrices in both preferences and dynamics are random (Svensson and Williams, 2007a, 2007b). In particular, each model has multiple "modes"—a finite collection of different possible values for the matrices, whose evolution is governed by a finite-state Markov chain. In our previous work, we discuss how these modes could be structured to capture many different types of uncertainty relevant for policymakers. Here we put those suggestions into practice. We start by briefly discussing how an MJLQ model can be derived as a mode-dependent linear-quadratic approximation of an underlying nonlinear model, and we then apply our methods to a simple empirical mode-dependent New-Keynesian model of the U.S. economy, using a variant of a model by Lindé (2005).

In Svensson and Williams (2007b), we study optimal policy design in MJLQ models when policymakers can or cannot observe the current mode, but we abstract from any learning and inference about the

We thank James Bullard, Timothy Cogley, and Andrew Levin for comments on an earlier paper of ours, which helped inspire this paper, and Carl E. Walsh for comments on this paper. The views expressed in this paper are solely our responsibility and should not to be interpreted as reflecting the views of any other member of the Executive Board of Sveriges Riksbank. Financial support from the Central Bank of Chile and the National Science Foundation is gratefully acknowledged.

Monetary Policy under Uncertainty and Learning, edited by Klaus Schmidt-Hebbel and Carl E. Walsh, Santiago, Chile.

current mode. Although in many cases the optimal policy under no learning (NL) is not a normatively desirable policy, it serves as a useful benchmark for our later policy analyses. In Svensson and Williams (2007a), we focus on learning and inference in the more relevant situation, particularly for model-uncertainty applications in which the modes are not directly observable. Thus, decisionmakers must filter their observations to make inferences about the current mode. As in most Bayesian learning problems, the optimal policy typically includes an experimentation component reflecting the endogeneity of information. This class of problems has a long history in economics, and solutions are notoriously difficult to obtain. We developed algorithms to solve numerically for the optimal policy.[1] Given the curse of dimensionality, the Bayesian optimal policy (BOP) is only feasible in relatively small models. Confronted with these difficulties, we also considered adaptive optimal policy (AOP).[2] In this case, the policymaker in each period updates the probability distribution of the current mode in a Bayesian way, but the optimal policy is computed each period under the assumption that the policymaker will not learn from observations in the future. In our setting, the AOP is significantly easier to compute, and in many cases it provides a good approximation to the BOP. Moreover, the AOP analysis is of some interest in its own right, as it is closely related to specifications of adaptive learning that have been widely studied in macroeconomics.[3] The AOP specification also rules out the experimentation that some may view as objectionable in a policy context.[4]

In this paper, we apply our methodology to study optimal monetary policy design under uncertainty in dynamic stochastic

1. In addition to the classic literature (on such problems as a monopolist learning its demand curve), Wieland (2000, 2006) and Beck and Wieland (2002) examine Bayesian optimal policy and optimal experimentation in a context similar to ours but without forward-looking variables. Tesfaselassie, Schaling, and Eijffinger (2006) examine passive and active learning in a simple model with a forward-looking element in the form of a long interest rate in the aggregate demand equation. Ellison and Valla (2001) and Cogley, Colacito, and Sargent (2007) study situations like ours, but their expectational component is as in the Lucas supply curve ($E_{t-1}\pi_t$, for example) rather than our forward-looking case ($E_t\pi_{t+1}$, for example). More closely related to our present paper, Ellison (2006) analyzes active and passive learning in a New-Keynesian model with uncertainty about the slope of the Phillips curve.

2. The literature also refers to optimal policy under no learning, adaptive optimal policy, and Bayesian optimal policy as myopia, passive learning, and active learning, respectively.

3. See Evans and Honkapohja (2001) for an overview.

4. AOP is also useful for technical reasons, as it gives us a good starting point for our more intensive numerical calculations in the BOP case.

general equilibrium (DSGE) models. We begin by summarizing the main findings from our previous work, leading to implementable algorithms for analyzing policy in MJLQ models. We then turn to analyzing optimal policy in DSGE models. To quantify the gains from experimentation, we focus on a small empirical benchmark New-Keynesian model. In this model, we compare and contrast optimal policies under no learning, AOP, and BOP. We analyze whether learning is beneficial—it is not always so, a fact that at least partially reflects our assumption of symmetric information between the policymakers and the public—and then quantify the additional gains from experimentation.[5]

Since we typically find that the gains from experimentation are small, the rest of the paper focuses on our adaptive optimal policy, which shuts down the experimentation channel. As the AOP is much easier to compute, this allows us to work with much larger and more empirically relevant policy models. In the latter part of the paper, we analyze one such model, an estimated forward-looking model that is a mode-dependent variant of Lindé (2005). There, we focus on how optimal policy should respond to uncertainty about the degree to which agents are forward-looking, and we show that there are substantial gains from learning in this framework.

The paper is organized as follows. Section 1 presents the MJLQ framework and summarizes our earlier work. Section 2 presents our analysis of learning and experimentation in a simple benchmark New-

5. In addition to our own previous work, MJLQ models have been widely studied in the control-theory literature for the special case in which the model modes are observable and there are no forward-looking variables (see Costa, Fragoso, and Marques, 2005, and the references therein). Do Val and Basar (1999) provide an application of an adaptive-control MJLQ problem in economics. Zampolli (2006) uses such an MJLQ model to examine monetary policy under shifts between regimes with and without an asset-market bubble. Blake and Zampolli (2006) extend the MJLQ model with observable modes to include forward-looking variables and present an algorithm for the solution of an equilibrium resulting from optimization under discretion. Svensson and Williams (2007b) provide a more general extension of the MJLQ framework with forward-looking variables and present algorithms for the solution of an equilibrium resulting from optimization under commitment in a timeless perspective, as well as arbitrary time-varying or time-invariant policy rules, using the recursive saddlepoint method of Marcet and Marimon (1998). That paper also provides two concrete examples: an estimated backward-looking model (a three-mode variant of Rudebusch and Svensson, 1999) and an estimated forward-looking model (a three-mode variant of Lindé, 2005). Svensson and Williams (2007b) also extend the MJLQ framework to the more realistic case of unobservable modes, although without introducing learning and inference about the probability distribution of modes. Svensson and Williams (2007a) focus on learning and experimentation in the MJLQ framework.

Keynesian model, and section 3 presents our analysis in an estimated empirical New-Keynesian model. Section 4 presents some conclusions and suggestions for further work.

1. MJLQ Analysis of Optimal Policy

This section summarizes our earlier work (Svensson and Williams, 2007a, 2007b). We start by describing our MJLQ model and then briefly discuss approximate MJLQ models. Finally, we explore the three types of optimal policies considered: optimal policy with no learning, adaptive optimal policy, and Bayesian optimal policy.

1.1 An MJLQ Model

We consider an MJLQ model of an economy with forward-looking variables. The economy has a private sector and a policymaker. We let $\mathbf{X}_t$ denote an n_X vector of predetermined variables in period t, $\mathbf{x}_t$ an n_x vector of forward-looking variables, and $\mathbf{i}_t$ an n_i vector of policymaker instruments (control variables).[6] We let model uncertainty be represented by n_j possible modes and let $j_t \in N_j \equiv \{1, 2, ..., n_j\}$ denote the mode in period t. The model of the economy can then be written

$$\mathbf{X}_{t+1} = \mathbf{A}_{11j_{t+1}}\mathbf{X}_t + \mathbf{A}_{12j_{t+1}}\mathbf{x}_t + \mathbf{B}_{1j_{t+1}}\mathbf{i}_t + \mathbf{C}_{1j_{t+1}}\varepsilon_{t+1}, \tag{1}$$

$$E_t\mathbf{H}_{j_{t+1}}\mathbf{x}_{t+1} = \mathbf{A}_{21j_t}\mathbf{X}_t + \mathbf{A}_{22j_t}\mathbf{x}_t + \mathbf{B}_{2j_t}\mathbf{i}_t + \mathbf{C}_{2j_t}\varepsilon_t, \tag{2}$$

where ε_t is a multivariate normally distributed random i.i.d. n_ε vector of shocks with mean zero and contemporaneous covariance matrix $\mathbf{I}_{n_\varepsilon}$. The matrices $\mathbf{A}_{11j}$, $\mathbf{A}_{12j}$, ..., $\mathbf{C}_{2j}$ have the appropriate dimensions and depend on the mode j. Given that a structural model here is simply a collection of matrices, each mode can represent a different model of the economy. Thus, uncertainty about the prevailing mode *is* model uncertainty.[7]

The matrices on the right-hand side of equation (1) depend on the mode j_{t+1} in period $t + 1$, whereas the matrices on the right-hand side

6. The first component of $\mathbf{X}_t$ may be unity, to allow for mode-dependent intercepts in the model equations.

7. See also Svensson and Williams (2007b), where we show how many different types of uncertainty can be mapped into our MJLQ framework.

of equation (2) depend on the mode j_t in period t. Equation (1) then determines the predetermined variables in period $t + 1$ as a function of the mode and shocks in period $t + 1$ and the predetermined variables, forward-looking variables, and instruments in period t. Equation (2) determines the forward-looking variables in period t as a function of the mode and shocks in period t, the expectations in period t of the next period's mode and forward-looking variables, and the predetermined variables and instruments in period t. The matrix $\mathbf{A}_{22j}$ is nonsingular for each $j \in N_j$.

The mode j_t follows a Markov process with the transition matrix $\mathbf{P} \equiv [P_{jk}]$.[8] The shocks ε_t have mean zero and are i.i.d. with probability density φ, and we assume, without loss of generality, that ε_t is independent of j_t.[9] We also assume that $\mathbf{C}_{1j}\varepsilon_t$ and $\mathbf{C}_{2k}\varepsilon_t$ are independent for all $j, k \in N_j$. These shocks, along with the modes, are the driving forces in the model. They are not directly observed. For technical reasons, it is convenient but not necessary that they are independent. We let $\mathbf{p}_t = (p_{1t},...,p_{n_jt})'$ denote the true probability distribution of j_t in period t. We let $p_{t+\tau|t}$ denote the policymaker's and private sector's estimate in the beginning of period t of the probability distribution in period $t + \tau$. The prediction equation for the probability distribution is

$$\mathbf{p}_{t+1|t} = \mathbf{P}'\mathbf{p}_{t|t}. \tag{3}$$

We let the operator $\mathrm{E}_t[\cdot]$ in the expression $\mathrm{E}_t \mathbf{H}_{j_{t+1}}\mathbf{x}_{t+1}$ on the left-hand side of equation (2) denote expectations in period t, conditional on the policymaker's and the private sector's information in the beginning of period t, including $\mathbf{X}_t$, $\mathbf{i}_t$, and $\mathbf{p}_{t|t}$, but excluding j_t and ε_t. The maintained assumption is thus symmetric information between the policymaker and the (aggregate) private sector. Since forward-looking variables will be allowed to depend on j_t, parts of the private sector—but not the aggregate private sector—may be able to observe j_t and parts of ε_t. While we focus on the determination of the optimal policy instrument $\mathbf{i}_t$, our results also show how private sector choices as embodied in $\mathbf{x}_t$ are affected by uncertainty and learning. The precise informational assumptions and the determination of $\mathbf{p}_{t|t}$ are specified below.

8. Obvious special cases are $\mathbf{P} = \mathbf{I}_{n_j}$, when the modes are completely persistent, and $\mathbf{P}_{j\cdot} = \bar{\mathbf{p}}'$, $(j \in N_j)$, when the modes are serially i.i.d. with probability distribution $\bar{\mathbf{p}}$.

9. We can still incorporate additive mode-dependent shocks since the models allow mode-dependent intercepts (as well as mode-dependent standard deviations).

We let the policymaker's intertemporal loss function in period t be

$$E_t \sum_{\tau=0}^{\infty} \delta^{\tau} L(\mathbf{X}_{t+\tau}, \mathbf{x}_{t+\tau}, \mathbf{i}_{t+\tau}, j_{t+\tau}), \tag{4}$$

where δ is a discount factor satisfying $0 < \delta < 1$, and the period loss, $L(\mathbf{X}_t, \mathbf{x}_t, \mathbf{i}_t, j_t)$, satisfies

$$L(\mathbf{X}_t, \mathbf{x}_t, \mathbf{i}_t, j_t) \equiv \begin{bmatrix} \mathbf{X}_t \\ \mathbf{x}_t \\ \mathbf{i}_t \end{bmatrix}' \mathbf{W}_{j_t} \begin{bmatrix} \mathbf{X}_t \\ \mathbf{x}_t \\ \mathbf{i}_t \end{bmatrix}, \tag{5}$$

where the matrix $\mathbf{W}_j$ ($j \in N_j$) is positive semidefinite. We assume that the policymaker optimizes under commitment in a timeless perspective. As explained below, we then add the term

$$\Xi_{t-1} \frac{1}{\delta} E_t \mathbf{H}_{j_t} \mathbf{x}_t \tag{6}$$

to the intertemporal loss function in period t. As we show below, the n_x vector Ξ_{t-1} is the vector of Lagrange multipliers for equation (2) from the optimization problem in period $t-1$. For the special case in which there are no forward-looking variables ($n_x = 0$), the model consists of equation (1) only, without the term $\mathbf{A}_{12j_{t+1}}\mathbf{x}_t$ the period loss function depends on $\mathbf{X}_t$, $\mathbf{i}_t$, and j_t only; and there is no role for the Lagrange multipliers Ξ_{t-1} or the term in equation (6).

1.2 Approximate MJLQ Models

While in this paper we start with an MJLQ model, the usual formulations of economic models are not of this type. However, the same type of approximation methods that are widely used to convert nonlinear models into their linear counterparts can also convert nonlinear models into MJLQ models. We analyze this issue in Svensson and Williams (2007b) and present an illustration, as well. Here we briefly discuss the main ideas. Rather than analyzing local deviations from a single steady state as in conventional linearizations, for an MJLQ approximation we analyze the local deviations from (potentially) separate, mode-dependent steady states. Standard linearizations are justified as asymptotically valid for small shocks, since an increasing time is spent in the vicinity of the steady state. Our MJLQ approximations are asymptotically valid for small shocks and persistent modes, since an increasing time is spent in the vicinity of each mode-dependent steady state. Thus, for slowly varying Markov chains, our MJLQ models provide accurate approximations of nonlinear models with Markov switching.

1.3 Types of Optimal Policies

We distinguish three cases: optimal policy when there is no learning (NL), adaptive optimal policy (AOP), and Bayesian optimal policy (BOP). By NL, we refer to a situation in which the policymaker and the aggregate private sector have a probability distribution $\mathbf{p}_{t|t}$ over the modes in period t and update the probability distribution in future periods using the transition matrix only, so the updating equation is

$$\mathbf{p}_{t+1|t+1} = \mathbf{P}'\mathbf{p}_{t|t} \tag{7}$$

That is, the policymaker and the private sector do not use observations of the variables in the economy to update the probability distribution. The policymaker then determines optimal policy in period t conditional on $\mathbf{p}_{t|t}$ and equation (7). This is a variant of a case examined in Svensson and Williams (2007b).

By AOP, we refer to a situation in which the policymaker in period t determines optimal policy as in the NL case, but then uses observations of the realization of the variables in the economy to update its probability distribution according to Bayes' theorem. In this case, the instruments will generally have an effect on the updating of future probability distributions, and through this channel they separately affect the intertemporal loss. However, the policymaker does not exploit that channel in determining optimal policy. That is, the policymaker does not do any conscious experimentation. By BOP, we refer to a situation in which the policymaker acknowledges that the current instruments will affect future inference and updating of the probability distribution and takes this separate channel into account when calculating optimal policy. BOP thus includes optimal experimentation, whereby the policymaker may, for instance, pursue policy that increases losses in the short run but improves the inference of the probability distribution and therefore lowers losses in the longer run.

1.3.1 Optimal policy with no learning

We first consider the NL case. Svensson and Williams (2007b) derive the equilibrium under commitment in a timeless perspective for the case in which $\mathbf{X}_t$, $\mathbf{x}_t$, and $\mathbf{i}_t$ are observable in period t, j_t is unobservable, and the updating equation for $p_{t|t}$ is given by equation (7). Observations of $\mathbf{X}_t$, $\mathbf{x}_t$, and $\mathbf{i}_t$ are then not used to update $p_{t|t}$.

It is useful to replace equation (2) by the two equivalent equations,

$$E_t \mathbf{H}_{j_{t+1}} \mathbf{x}_{t+1} = \mathbf{z}_t \tag{8}$$

and

$$0 = \mathbf{A}_{21j_t} \mathbf{X}_t + \mathbf{A}_{22j_t} \mathbf{x}_t - \mathbf{z}_t + \mathbf{B}_{2j_t} \mathbf{i}_t + \mathbf{C}_{2j_t} \varepsilon_t, \tag{9}$$

where we introduce the n_x vector of additional forward-looking variables, $\mathbf{z}_t$. Introducing this vector is a practical way of keeping track of the expectations term on the left-hand side of equation (2). Furthermore, it is practical to use equation (9) to solve $\mathbf{x}_t$ as a function of $\mathbf{X}_t$, $\mathbf{z}_t$, $\mathbf{i}_t$, j_t, and ε_t:

$$\mathbf{x}_t = \tilde{\mathbf{x}}(\mathbf{X}_t, \mathbf{z}_t, \mathbf{i}_t, j_t, \varepsilon_t) \equiv \mathbf{A}_{22j_t}^{-1} \left(\mathbf{z}_t - \mathbf{A}_{21j_t} \mathbf{X}_t - \mathbf{B}_{2j_t} \mathbf{i}_t - \mathbf{C}_{2j_t} \varepsilon_t \right). \tag{10}$$

For a given j_t, this function is linear in $\mathbf{X}_t$, $\mathbf{z}_t$, $\mathbf{i}_t$, and ε_t.

To solve for the optimal decisions, we use the recursive saddlepoint method.[10] We thus introduce Lagrange multipliers for each forward-looking equation, the lagged values of which become state variables and reflect costs of commitment, while the current values become control variables. The dual period loss function can be written

$$E_t \tilde{L}\left(\tilde{\mathbf{X}}_t, \mathbf{z}_t, \mathbf{i}_t, \gamma_t, j_t, \varepsilon_t \right) \equiv \sum_j p_{jt|t} \int \tilde{L}\left(\tilde{\mathbf{X}}_t, \mathbf{z}_t, \mathbf{i}_t, \gamma_t, j, \varepsilon_t \right) \varphi(\varepsilon_t) d\varepsilon_t,$$

where $\tilde{\mathbf{X}}_t \equiv (\mathbf{X}_t', \Xi_{t-1}')'$ is the $(n_X + n_x)$ vector of extended predetermined variables (that is, including the n_x vector, Ξ_{t-1}), γ_t is an n_x vector of

10. See Marcet and Marimon (1998), Svensson and Williams (2007b), and Svensson (2007) for details of the recursive saddlepoint method.

Lagrange multipliers, and $\varphi(\cdot)$ denotes a generic probability density function (for ε_t, the standard normal density function), and where

$$\begin{aligned}\tilde{L}\left(\tilde{\mathbf{X}}_t,\mathbf{z}_t,\mathbf{i}_t,\gamma_t,j_t,\varepsilon_t\right) \equiv L\left(\mathbf{X}_t,\tilde{\mathbf{x}}\left(\mathbf{X}_t,\mathbf{z}_t,\mathbf{i}_t,j_t,\varepsilon_t\right),\mathbf{i}_t,j_t\right) \\ -\gamma_t'\mathbf{z}_t + \Xi_{t-1}'\frac{1}{\delta}\mathbf{H}_{j_t}\tilde{\mathbf{x}}\left(\mathbf{X}_t,\mathbf{z}_t,\mathbf{i}_t,j_t,\varepsilon_t\right).\end{aligned} \tag{11}$$

As discussed in Svensson and Williams (2007b), the failure of the law of iterated expectations leads us to introduce a collection of value functions, $\hat{V}(\mathbf{s}_t, j)$, which condition on the mode, while the value function $\hat{V}(\mathbf{s}_t)$ averages over these and represents the solution of the dual optimization problem. The somewhat unusual Bellman equation for the dual problem can be written

$$\begin{aligned}\tilde{V}(\mathbf{s}_t) &\equiv E_t\hat{V}(\mathbf{s}_t,j_t) \equiv \sum\nolimits_j p_{jt|t}\hat{V}(\mathbf{s}_t,j) \\ &= \max_{\gamma_t}\min_{(\mathbf{z}_t,\mathbf{i}_t)} E_t\left[\tilde{L}(\tilde{\mathbf{X}}_t,\mathbf{z}_t,\mathbf{i}_t,\gamma_t,j_t,\varepsilon_t)+\delta\hat{V}(g(\mathbf{s}_t,\mathbf{z}_t,\mathbf{i}_t,\gamma_t,j_t,\varepsilon_t,j_{t+1},\varepsilon_{t+1}),j_{t+1})\right] \\ &\equiv \max_{\gamma_t}\min_{(\mathbf{z}_t,\mathbf{i}_t)}\sum\nolimits_j p_{jt|t}\int\left[\begin{array}{l}\tilde{L}(\tilde{\mathbf{X}}_t,\mathbf{z}_t,\mathbf{i}_t,\gamma_t,j,\varepsilon_t) \\ +\delta\sum\nolimits_k P_{jk}\hat{V}(g(\mathbf{s}_t,\mathbf{z}_t,\mathbf{i}_t,\gamma_t,j,\varepsilon_t,k,\varepsilon_{t+1}),k)\end{array}\right]\varphi(\varepsilon_t)\varphi(\varepsilon_{t+1})d\varepsilon_t d\varepsilon_{t+1}.\end{aligned} \tag{12}$$

where $\mathbf{s}_t \equiv (\tilde{\mathbf{X}}_t', \mathbf{p}_{t|t}')'$ denotes the perceived state of the economy (it includes the perceived probability distribution, $\mathbf{p}_{t|t}$, but not the true mode) and $(\mathbf{s}_t, j_t)$ denotes the true state of the economy (it includes the true mode of the economy). As we discuss in more detail below, it is necessary to include the mode j_t in the state vector because the beliefs do not satisfy the law of iterated expectations. In the BOP case, beliefs do satisfy this property, so the state vector is simply $\mathbf{s}_t$. Also, in the Bellman equation we require that all the choice variables respect the information constraints, and they thus depend on the perceived state $\mathbf{s}_t$ but not directly on the mode j.

The optimization is subject to the transition equation for $\mathbf{X}_t$,

$$\mathbf{X}_{t+1} = \mathbf{A}_{11j_{t+1}}\mathbf{X}_t + \mathbf{A}_{12j_{t+1}}\tilde{\mathbf{x}}(\mathbf{X}_t,\mathbf{z}_t,\mathbf{i}_t,j_t,\varepsilon_t) + \mathbf{B}_{1j_{t+1}}\mathbf{i}_t + \mathbf{C}_{1j_{t+1}}\varepsilon_{t+1}, \tag{13}$$

where we have substituted $\tilde{\mathbf{x}}(\mathbf{X}_t, \mathbf{z}_t, \mathbf{i}_t, j_t, \varepsilon_t)$ for $\mathbf{x}_t$; the new dual transition equation for Ξ_t,

$$\Xi_t = \gamma_t, \tag{14}$$

and the transition equation (7) for $\mathbf{p}_{t|t}$. Combining equations, we have the transition for $\mathbf{s}_t$,

$$\begin{aligned}
\mathbf{s}_{t+1} \equiv \begin{bmatrix} \mathbf{X}_{t+1} \\ \Xi_t \\ \mathbf{p}_{t+1|t+1} \end{bmatrix} &= g(\mathbf{s}_t, \mathbf{z}_t, \mathbf{i}_t, \gamma_t, j_t, \varepsilon_t, j_{t+1}, \varepsilon_{t+1}) \\
&\equiv \begin{bmatrix} \mathbf{A}_{11j_{t+1}} \mathbf{X}_t + \mathbf{A}_{12j_{t+1}} \tilde{\mathbf{x}}(\mathbf{X}_t, \mathbf{z}_t, \mathbf{i}_t, j, \varepsilon_t) + \mathbf{B}_{1j_{t+1}} \mathbf{i}_t + \mathbf{C}_{1j_{t+1}} \varepsilon_{t+1} \\ \gamma_t \\ \mathbf{P}' \mathbf{p}_{t|t} \end{bmatrix}.
\end{aligned} \tag{15}$$

It is straightforward to see that the solution of the dual optimization problem (equation 12) is linear in $\tilde{\mathbf{X}}_t$ for given $\mathbf{p}_{t|t}$, j_t,

$$\begin{bmatrix} \mathbf{z}_t \\ \mathbf{i}_t \\ \gamma_t \end{bmatrix} = \begin{bmatrix} \mathbf{z}(\mathbf{s}_t) \\ \mathbf{i}(\mathbf{s}_t) \\ \gamma(\mathbf{s}_t) \end{bmatrix} = F(\mathbf{p}_{t|t}) \tilde{\mathbf{X}}_t \equiv \begin{bmatrix} F_{\mathbf{z}}(\mathbf{p}_{t|t}) \\ F_{\mathbf{i}}(\mathbf{p}_{t|t}) \\ F_{\gamma}(\mathbf{p}_{t|t}) \end{bmatrix} \tilde{\mathbf{X}}_t, \tag{16}$$

$$\begin{aligned}
\mathbf{x}_t = \mathbf{x}(\mathbf{s}_t, j_t, \varepsilon_t) &\equiv \tilde{\mathbf{x}}(\mathbf{X}_t, \mathbf{z}(\mathbf{s}_t), \mathbf{i}(\mathbf{s}_t), j_t, \varepsilon_t) \\
&\equiv F_{\mathbf{x}\tilde{\mathbf{X}}}(\mathbf{p}_{t|t}, j_t) \tilde{\mathbf{X}}_t + F_{\mathbf{x}\varepsilon}(\mathbf{p}_{t|t}, j_t) \varepsilon_t.
\end{aligned} \tag{17}$$

This solution is also the solution to the original primal optimization problem. We note that $\mathbf{x}_t$ is linear in ε_t for given $p_{t|t}$ and j_t. The equilibrium transition equation is then given by

$$\mathbf{s}_{t+1} = \hat{g}(\mathbf{s}_t, j_t, \varepsilon_t, j_{t+1}, \varepsilon_{t+1}) \equiv g(\mathbf{s}_t, \mathbf{z}(\mathbf{s}_t), \mathbf{i}(\mathbf{s}_t), \gamma(\mathbf{s}_t), j_t, \varepsilon_t, j_{t+1}, \varepsilon_{t+1}). \tag{18}$$

As can be easily verified, the (unconditional) dual value function $\hat{V}(\mathbf{s}_t)$ is quadratic in $\tilde{\mathbf{X}}_t$ for given $\mathbf{p}_{t|t}$, taking the form

$$\tilde{V}(\mathbf{s}_t) \equiv \tilde{\mathbf{X}}_t' \tilde{V}_{\tilde{\mathbf{X}}\tilde{\mathbf{X}}}(\mathbf{p}_{t|t}) \tilde{\mathbf{X}}_t + w(\mathbf{p}_{t|t}).$$

The conditional dual value function $\hat{V}(\mathbf{s}_t, j_t)$ gives the dual intertemporal loss conditional on the true state of the economy, $(\mathbf{s}_t, j_t)$. It follows that this function satisfies

$$\hat{V}(\mathbf{s}_t, j) \equiv \int \begin{bmatrix} \tilde{L}(\tilde{\mathbf{X}}_t, \mathbf{z}(\mathbf{s}_t), \mathbf{i}(\mathbf{s}_t), \gamma(\mathbf{s}_t), j, \varepsilon_t) \\ +\delta \sum_k P_{jk} \hat{V}(\hat{g}(\mathbf{s}_t, j, \varepsilon_t, k, \varepsilon_{t+1}), k) \end{bmatrix} \varphi(\varepsilon_t)\varphi(\varepsilon_{t+1}) d\varepsilon_t d\varepsilon_{t+1}, \quad (j \in N_j).$$

The function $\hat{V}(\mathbf{s}_t, j_t)$ is also quadratic in $\tilde{\mathbf{X}}_t$ for given $\mathbf{p}_{t|t}$ and j_t,

$$\hat{V}(\mathbf{s}_t, j_t) \equiv \tilde{\mathbf{X}}_t' \hat{V}_{\tilde{\mathbf{X}}\tilde{\mathbf{X}}}(\mathbf{p}_{t|t}, j_t)\tilde{\mathbf{X}}_t + \hat{w}(\mathbf{p}_{t|t}, j_t).$$

It follows that we have

$$\tilde{V}_{\tilde{\mathbf{X}}\tilde{\mathbf{X}}}(\mathbf{p}_{t|t}) \equiv \sum_j p_{jt|t} \hat{V}_{\tilde{\mathbf{X}}\tilde{\mathbf{X}}}(\mathbf{p}_{t|t}, j);$$

$$w(\mathbf{p}_{t|t}) \equiv \sum_j p_{jt|t} \hat{w}(\mathbf{p}_{t|t}, j).$$

Although we find the optimal policies from the dual problem, we use the value function for the primal problem (with the original, unmodified loss function) to measure true expected losses. This value function, with the period loss function $E_t L(\mathbf{X}_t, \mathbf{x}_t, \mathbf{i}_t, j_t)$ rather than $E_t \tilde{L}(\tilde{\mathbf{X}}_t, \mathbf{z}_t, \mathbf{i}_t, \gamma_t, j_t, \varepsilon_t)$, satisfies

$$\begin{aligned} V(\mathbf{s}_t) &\equiv \tilde{V}(\mathbf{s}_t) - \Xi_{t-1}' \frac{1}{\delta} \sum_j p_{jt|t} \mathbf{H}_j \int \mathbf{x}(\mathbf{s}_t, j, \varepsilon_t)\varphi(\varepsilon_t) d\varepsilon_t \\ &= \tilde{V}(\mathbf{s}_t) - \Xi_{t-1}' \frac{1}{\delta} \sum_j p_{jt|t} \mathbf{H}_j \mathbf{x}(\mathbf{s}_t, j, 0). \end{aligned} \qquad (19)$$

where the second equality follows since $\mathbf{x}(\mathbf{s}_t, j_t, \varepsilon_t)$ is linear in ε_t for given $\mathbf{s}_t$ and j_t. It is quadratic in $\tilde{\mathbf{X}}_t$ for given $\mathbf{p}_{t|t}$,

$$V(\mathbf{s}_t) \equiv \tilde{\mathbf{X}}_t' V_{\tilde{\mathbf{X}}\tilde{\mathbf{X}}}(\mathbf{p}_{t|t})\tilde{\mathbf{X}}_t + w(\mathbf{p}_{t|t}),$$

where the scalar $w(\mathbf{p}_{t|t})$ in the primal value function is identical to that in the dual value function. This is the value function conditional on $\tilde{\mathbf{X}}_t$ and $\mathbf{p}_{t|t}$ after $\mathbf{X}_t$ has been observed but before $\mathbf{x}_t$ has been observed, taking into account that j_t and ε_t are not observed. Hence, the second term on the right-hand side of equation (19) contains the expectation of $\mathbf{H}_{j_t}\mathbf{x}_t$ conditional on that information.[11]

In Svensson and Williams (2007a, 2007b), we present algorithms to compute the solution and the primal and dual value functions for the no-learning case. For future reference, we note that the value function for the primal problem also satisfies

$$V(\mathbf{s}_t) \equiv \sum_j p_{jt|t} \breve{V}(\mathbf{s}_t, j),$$

where the conditional value function, $\breve{V}(\mathbf{s}_t, j_t)$, satisfies

$$\breve{V}(\mathbf{s}_t, j) = \int \begin{bmatrix} L(\mathbf{X}_t, \mathbf{x}(\mathbf{s}_t, j, \varepsilon_t), \mathbf{i}(\mathbf{s}_t), j) \\ +\delta \sum_k P_{jk} \breve{V}(\hat{g}(\mathbf{s}_t, j, \varepsilon_t, k, \varepsilon_{t+1}), k) \end{bmatrix} \varphi(\varepsilon_t)\varphi(\varepsilon_{t+1}) d\varepsilon_t d\varepsilon_{t+1}, \quad (j \in N_j). \tag{20}$$

1.3.2 Adaptive optimal policy

Consider now the case of adaptive optimal policy, in which the policymaker uses the same policy function as in the no-learning case, but each period updates the probabilities on which this policy is conditioned. This case is thus simple to implement recursively, as we have already discussed how to solve for the optimal decisions and below we show how to update probabilities. However, the ex ante evaluation of expected loss is more complex, as we show below. In particular, we assume that $\mathbf{C}_{2j_t} \not\equiv 0$ and that both ε_t and j_t are unobservable. The estimate $\mathbf{p}_{t|t}$ is the result of Bayesian updating, using all information available, but the optimal policy in period t is computed under the perceived updating equation (7). That is, we disregard the fact that the policy choice will affect future $\mathbf{p}_{t+\tau|t+\tau}$ and that future expected loss will change when $\mathbf{p}_{t+\tau|t+\tau}$ changes. Under the assumption that the expectations on the left-hand side of equation (2) are conditional on equation (7), the variables $\mathbf{z}_t$, $\mathbf{i}_t$, γ_t, and $\mathbf{x}_t$ in period t are still determined by equations (16) and (17).

To determine the updating equation for $\mathbf{p}_{t|t}$, we specify an explicit sequence of information revelation as follows, in nine steps. The timing assumptions are necessary to spell out the appropriate conditioning for decisions and updating of beliefs.

11. To be precise, the observation of $\mathbf{X}_t$, which depends on $\mathbf{C}_{1j_t}\varepsilon_t$, allows some inference of ε_t, $\varepsilon_{t|t}$. The variable $\mathbf{x}_t$ will depend on j_t and on ε_t, but on ε_t only through $\mathbf{C}_{2j_t}\varepsilon_t$. By assumption, $\mathbf{C}_{1j}\varepsilon_t$ and $\mathbf{C}_{2k}\varepsilon_t$ are independent. Hence, any observation of $\mathbf{X}_t$ and $\mathbf{C}_{1j}\varepsilon_t$ does not convey any information about $\mathbf{C}_{2j}\varepsilon_t$, so $E_t\mathbf{C}_{2j_t}\varepsilon_t = 0$.

First, the policymaker and the private sector enter period t with the prior $\mathbf{p}_{t|t-1}$. They know $\mathbf{X}_{t-1}$, $\mathbf{x}_{t-1} = \mathbf{x}(\mathbf{s}_{t-1}, j_{t-1}, \varepsilon_{t-1})$, $\mathbf{z}_{t-1} = \mathbf{z}(\mathbf{s}_{t-1})$, $\mathbf{i}_{t-1} = \mathbf{i}(\mathbf{s}_{t-1})$, and $\Xi_{t-1} = \gamma(\mathbf{s}_{t-1})$ from the previous period.

Second, the mode j_t and the vector of shocks ε_t are realized in the beginning of period t. The vector of predetermined variables $\mathbf{X}_t$ is then realized according to equation (1).

Third, the policymaker and the private sector observe $\mathbf{X}_t$. They then know that $\tilde{\mathbf{X}}_t \equiv (\mathbf{X}_t', \Xi_{t-1}')'$. They do not observe j_t or ε_t.

Fourth, the policymaker and the private sector update the prior $\mathbf{p}_{t|t-1}$ to the posterior $\mathbf{p}_{t|t}$ according to Bayes' theorem and the updating equation

$$p_{jt|t} = \frac{\varphi(\mathbf{X}_t \mid j_t = j, \mathbf{X}_{t-1}, \mathbf{x}_{t-1}, \mathbf{i}_{t-1}, \mathbf{p}_{t|t-1})}{\varphi(\mathbf{X}_t \mid \mathbf{X}_{t-1}, \mathbf{x}_{t-1}, \mathbf{i}_{t-1}, \mathbf{p}_{t|t-1})} p_{jt|t-1}, \qquad (j \in N_j), \tag{21}$$

where again $\varphi(\cdot)$ denotes a generic density function.[12] Then the policymaker and the private sector know that $\mathbf{s}_t \equiv (\tilde{\mathbf{X}}_t', \mathbf{p}_{t|t}')'$.

Fifth, the policymaker solves the dual optimization problem, determines $\mathbf{i}_t = \mathbf{i}(\mathbf{s}_t)$, and implements or announces the instrument setting $\mathbf{i}_t$.

Sixth, the private sector and policymaker form their expectations,

$$\mathbf{z}_t = E_t \mathbf{H}_{j_{t+1}} \mathbf{x}_{t+1} \equiv E\left[\mathbf{H}_{j_{t+1}} \mathbf{x}_{t+1} | \mathbf{s}_t\right].$$

In equilibrium, these expectations will be determined by equation (16). These expectations are by assumption formed before $\mathbf{x}_t$ is observed. The private sector and the policymaker know that $\mathbf{x}_t$ will, in equilibrium, be determined in the next step according to equation (17). Hence, they can form expectations of the soon-to-be determined $\mathbf{x}_t$ conditional on $j_t = j$,[13]

$$\mathbf{x}_{jt|t} = \mathbf{x}(\mathbf{s}_t, j, 0). \tag{22}$$

12. The policymaker and private sector can also estimate the shocks $\varepsilon_{t|t}$ as $\varepsilon_{t|t} = \Sigma_j p_{jt|t} \varepsilon_{jt|t}$, where $\varepsilon_{jt|t} \equiv \mathbf{X}_t - \mathbf{A}_{11j}\mathbf{X}_{t-1} - \mathbf{A}_{12j}\mathbf{x}_{t-1} - \mathbf{B}_{1j}\mathbf{i}_{t-1}$ $(j \in N_j)$. However, because of the assumed independence of $\mathbf{C}_{1j}\varepsilon_t$ and $\mathbf{C}_{2k}\varepsilon_t$, j, $k \in N_j$, we do not need to keep track of $\varepsilon_{jt|t}$.

13. Note that 0 instead of $\varepsilon_{jt|t}$ enters above. The inference $\varepsilon_{jt|t}$ above is inference about $\mathbf{C}_{1j}\varepsilon_t$, whereas $\mathbf{x}_t$ depends on ε_t through $\mathbf{C}_{2j}\varepsilon_t$. Since we assume that $\mathbf{C}_{1j}\varepsilon_t$ and $\mathbf{C}_{2j}\varepsilon_t$ are independent, there is no inference of $\mathbf{C}_{2j}\varepsilon_t$ from observing $\mathbf{X}_t$. Hence, $E_t\mathbf{C}_{2j_t}\varepsilon_t \equiv 0$. Because of the linearity of $\mathbf{x}_t$ in ε_t, the integration of $\mathbf{x}_t$ over ε_t results in $\mathbf{x}(\mathbf{s}_t, j_t, 0_t)$.

The private sector and the policymaker can also infer Ξ_t from

$$\Xi_t = \gamma(\mathbf{s}_t). \tag{23}$$

This allows the private sector and the policymaker to form the expectations

$$\mathbf{z}_t = \mathbf{z}(\mathbf{s}_t) = E_t\left[\mathbf{H}_{j_{t+1}}\mathbf{x}_{t+1}|\mathbf{s}_t\right] = \sum\nolimits_{j,k} P_{jk}\, p_{jt|t}\mathbf{H}_k\mathbf{x}_{k,t+1|jt}, \tag{24}$$

where

$$\begin{aligned}\mathbf{x}_{k,t+1|jt} &= \int \mathbf{x}\left(\begin{bmatrix}\mathbf{A}_{11k}\mathbf{X}_t + \mathbf{A}_{12k}\mathbf{x}(\mathbf{s}_t, j, \varepsilon_t) + \mathbf{B}_{1k}\mathbf{i}(\mathbf{s}_t)\\ \Xi_t \\ \mathbf{P}'\mathbf{p}_{t|t}\end{bmatrix}, k, \varepsilon_{t+1}\right)\varphi(\varepsilon_t)\varphi(\varepsilon_{t+1})d\varepsilon_t d\varepsilon_{t+1}\\ &= \mathbf{x}\left(\begin{bmatrix}\mathbf{A}_{11k}\mathbf{X}_t + \mathbf{A}_{12k}\mathbf{x}(\mathbf{s}_t, j, 0) + \mathbf{B}_{1k}\mathbf{i}(\mathbf{s}_t)\\ \Xi_t \\ \mathbf{P}'\mathbf{p}_{t|t}\end{bmatrix}, k, 0\right),\end{aligned}$$

and where we have exploited the linearity of $\mathbf{x}_t = \mathbf{x}(\mathbf{s}_t, j_t, \varepsilon_t)$ and $\mathbf{x}_{t+1} = \mathbf{x}(\mathbf{s}_{t+1}, j_{t+1}, \varepsilon_{t+1})$ in ε_t and ε_{t+1}. Under AOP, $\mathbf{z}_t$ is formed conditional on the belief that the probability distribution in period $t + 1$ will be given by $\mathbf{p}_{t+1|t+1} = \mathbf{P}'\mathbf{p}_{t|t}$, not by the true updating equation that we are about to specify.

Seventh, after the expectations $\mathbf{z}_t$ have been formed, $\mathbf{x}_t$ is determined as a function of $\mathbf{X}_t$, $\mathbf{z}_t$, $\mathbf{i}_t$, j_t, and ε_t by equation (10).

Eighth, the policymaker and the private sector then use the observed $\mathbf{x}_t$ to update $\mathbf{p}_{t|t}$ to the new posterior $\mathbf{p}^+_{t|t}$ according to Bayes' theorem, via the updating equation

$$p^+_{jt|t} = \frac{\varphi(\mathbf{x}_t \mid j_t = j, \mathbf{X}_t, \mathbf{z}_t, \mathbf{i}_t, \mathbf{p}_{t|t})}{\varphi(\mathbf{x}_t \mid \mathbf{X}_t, \mathbf{z}_t, \mathbf{i}_t, \mathbf{p}_{t|t})}\, p_{jt|t}, \qquad (j \in N_j). \tag{25}$$

Ninth, the policymaker and the private sector then leave period t and enter period $t + 1$ with the prior $\mathbf{p}_{t+1|t}$ given by the prediction equation

$$\mathbf{p}_{t+1|t} = \mathbf{P}'\mathbf{p}_{t|t}^{+}. \tag{26}$$

In the beginning of period t + 1, the mode j_{t+1} and the vector of shocks ε_{t+1} are realized, and $\mathbf{X}_{t+1}$ is determined by equation (1) and observed by the policymaker and the private sector. The sequence of the nine steps above then repeats itself. For more detail on the explicit densities in the updating equations (21) and (25), see Svensson and Williams (2007a).

The transition equation for $\mathbf{p}_{t+1|t+1}$ can be written

$$\mathbf{p}_{t+1|t+1} = Q(\mathbf{s}_t, \mathbf{z}_t, \mathbf{i}_t, j_t, \varepsilon_t, j_{t+1}, \varepsilon_{t+1}), \tag{27}$$

where $Q(\mathbf{s}_t, \mathbf{z}_t, \mathbf{i}_t, j_t, \varepsilon_t, j_{t+1}, \varepsilon_{t+1})$ is defined by the combination of equation (21) for period t + 1 with equations (13) and (26). The equilibrium transition equation for the full state vector is then given by

$$\begin{aligned}\mathbf{s}_{t+1} &\equiv \begin{bmatrix} \mathbf{X}_{t+1} \\ \Xi_t \\ \mathbf{p}_{t+1|t+1} \end{bmatrix} = \bar{g}(\mathbf{s}_t, j_t, \varepsilon_t, j_{t+1}, \varepsilon_{t+1}) \\ &\equiv \begin{bmatrix} \mathbf{A}_{11j_{t+1}}\mathbf{X}_t + \mathbf{A}_{12j_{t+1}}\mathbf{x}(\mathbf{s}_t, j_t, \varepsilon_t) + \mathbf{B}_{1j_{t+1}}\mathbf{i}(\mathbf{s}_t) + \mathbf{C}_{1j_{t+1}}\varepsilon_{t+1} \\ \gamma(\mathbf{s}_t) \\ Q(\mathbf{s}_t, \mathbf{z}(\mathbf{s}_t), \mathbf{i}(\mathbf{s}_t), j_t, \varepsilon_t, j_{t+1}, \varepsilon_{t+1}) \end{bmatrix},\end{aligned} \tag{28}$$

where the third row is given by the true updating equation (27) together with the policy function (16). Thus, in this AOP case, there is a distinction between the "perceived" transition equation (15) and the equilibrium transition equation (18), both of which include the perceived updating equation (7) in the bottom block, and the "true" equilibrium transition equation (28), which replaces the perceived updating equation (7) with the true updating equation (27).

Note that $V(\mathbf{s}_t)$ in equation (19), which is subject to the perceived transition equation (15), does not give the true (unconditional) value function for the AOP case. This is instead given by

$$\bar{V}(\mathbf{s}_t) \equiv \sum\nolimits_j p_{jt|t}\breve{V}(\mathbf{s}_t, j),$$

where the true conditional value function, $\breve{V}(s_t, j_t)$, satisfies

$$\breve{V}(\mathbf{s}_t, j) = \int \begin{bmatrix} L(\mathbf{X}_t, \mathbf{x}(\mathbf{s}_t, j, e_t), \mathbf{i}(\mathbf{s}_t), j) \\ +\delta \sum_k P_{jk} \breve{V}(\bar{g}(\mathbf{s}_t, j, e_t, k, e_{t+1}), k) \end{bmatrix} \varphi(e_t)\varphi(e_{t+1}) de_t de_{t+1}, \quad (j \in N_j). \quad (29)$$

That is, the true value function $\bar{V}(\mathbf{s}_t)$ takes into account the true updating equation for $\mathbf{p}_{t|t}$, equation (27), whereas the optimal policy, equation (16), and the perceived value function, $V(\mathbf{s}_t)$ in equation (19), are conditional on the perceived updating equation (7) and thereby the perceived transition equation (15). Also, $\bar{V}(\mathbf{s}_t)$ is the value function after $\tilde{\mathbf{X}}_t$ has been observed but before $\mathbf{x}_t$ is observed, so it is conditional on $\mathbf{p}_{t|t}$ rather than on $\mathbf{p}^+_{t|t}$. Since the full transition equation (28) is no longer linear given the belief updating in equation (27), the true value function $\bar{V}(\mathbf{s}_t)$ is no longer quadratic in $\tilde{\mathbf{X}}_t$ for given $\mathbf{p}_{t|t}$. Thus, more complex numerical methods are required to evaluate losses in the AOP case, although policy is still determined simply as in the NL case.

As we discuss in Svensson and Williams (2007a), the difference between the true updating equation for $\mathbf{p}_{t+1|t+1}$, (27), and the perceived updating equation (7) is that in the true updating equation, $\mathbf{p}_{t+1|t+1}$ becomes a random variable from the point of view of period t, with mean equal to $\mathbf{p}_{t+1|t}$. This is because $\mathbf{p}_{t+1|t+1}$ depends on the realization of j_{t+1} and ε_{t+1}. Bayesian updating thus induces a mean-preserving spread over beliefs, which in turn sheds light on the gains from learning. If the conditional value function $\breve{V}(\mathbf{s}_t, j_t)$ under NL is concave in $\mathbf{p}_{t|t}$ for given $\tilde{\mathbf{X}}_t$ and j_t, then by Jensen's inequality the true expected future loss under AOP will be lower than the true expected future loss under NL. That is, the concavity of the value function in beliefs means that learning leads to lower losses. While it is likely that $\breve{V}$ is indeed concave, as we show in applications, it need not be globally so, and thus learning need not always reduce losses. In some cases, the losses incurred by increased variability of beliefs may offset the expected precision gains. Furthermore, under BOP, it may be possible to adjust policy so as to further increase the variance of $\mathbf{p}_{t|t}$, that is, to achieve a mean-preserving spread that might further reduce the expected future loss.[14] This amounts to optimal experimentation.

14. Kiefer (1989) examines the properties of a value function, including concavity, under Bayesian learning for a simpler model without forward-looking variables.

1.3.3 Bayesian optimal policy

Finally, we consider the BOP case, in which optimal policy is determined while taking the updating equation (27) into account. That is, we now allow the policymaker to choose $\mathbf{i}_t$ taking into account that his actions will affect $\mathbf{p}_{t+1|t+1}$, which in turn will affect future expected losses. In particular, experimentation is allowed and is optimally chosen. Hence, for the BOP case, there is no distinction between the "perceived" and "true" transition equations.

The transition equation for the BOP case is

$$\mathbf{s}_{t+1} \equiv \begin{bmatrix} \mathbf{X}_{t+1} \\ \Xi_t \\ \mathbf{p}_{t+1|t+1} \end{bmatrix} = g(\mathbf{s}_t, \mathbf{z}_t, \mathbf{i}_t, \gamma_t, j_t, \varepsilon_t, j_{t+1}, \varepsilon_{t+1})$$
$$\equiv \begin{bmatrix} \mathbf{A}_{11j_{t+1}} \mathbf{X}_t + \mathbf{A}_{12j_{t+1}} \tilde{\mathbf{x}}(\mathbf{s}_t, \mathbf{z}_t, \mathbf{i}_t, j_t, \varepsilon_t) + \mathbf{B}_{1j_{t+1}} \mathbf{i}_t + \mathbf{C}_{1j_{t+1}} \varepsilon_{t+1} \\ \gamma_t \\ Q(\mathbf{s}_t, \mathbf{z}_t, \mathbf{i}_t, j_t, \varepsilon_t, j_{t+1}, \varepsilon_{t+1}) \end{bmatrix}. \tag{30}$$

Then the dual optimization problem can be written as equation (12) subject to the above transition equation (30). Matters simplify somewhat in the Bayesian case, however, as we do not need to compute the conditional value functions $\hat{V}(\mathbf{s}_t, j_t)$, which were required in the AOP case given the failure of the law of iterated expectations. The second term on the right-hand side of equation (12) can be written as

$$E_t \hat{V}(\mathbf{s}_{t+1}, j_{t+1}) \equiv E\left[\hat{V}(\mathbf{s}_{t+1}, j_{t+1}) \middle| \mathbf{s}_t\right].$$

Since, in the Bayesian case, the beliefs do satisfy the law of iterated expectations, this is then the same as

$$E\left[\hat{V}(\mathbf{s}_{t+1}, j_{t+1}) \middle| \mathbf{s}_t\right] = E\left[\tilde{V}(\mathbf{s}_{t+1}) \middle| \mathbf{s}_t\right].$$

See Svensson and Williams (2007a) for a proof.

Thus, the dual Bellman equation for the Bayesian optimal policy is

$$\begin{aligned}\tilde{V}(\mathbf{s}_t) &= \max_{\gamma_t} \min_{(\mathbf{z}_t, \mathbf{i}_t)} E_t\left[\tilde{L}(\tilde{\mathbf{X}}_t, \mathbf{z}_t, \mathbf{i}_t, \gamma_t, j_t, \varepsilon_t) + \delta \tilde{V}(g(\mathbf{s}_t, \mathbf{z}_t, \mathbf{i}_t, \gamma_t, j_t, \varepsilon_t, j_{t+1}, \varepsilon_{t+1}))\right] \\ &\equiv \max_{\gamma_t} \min_{(\mathbf{z}_t, \mathbf{i}_t)} \sum_j p_{jt|t} \int \begin{bmatrix} \tilde{L}(\tilde{\mathbf{X}}_t, \mathbf{z}_t, \mathbf{i}_t, \gamma_t, j, \varepsilon_t) \\ +\delta \sum_k P_{jk} \tilde{V}(g(\mathbf{s}_t, \mathbf{z}_t, \mathbf{i}_t, \gamma_t, j, \varepsilon_t, k, \varepsilon_{t+1})) \end{bmatrix} \varphi(\varepsilon_t)\varphi(\varepsilon_{t+1}) d\varepsilon_t d\varepsilon_{t+1},\end{aligned} \tag{31}$$

where the transition equation is given by equation (30).

The solution to the optimization problem can be written

$$\tilde{\mathbf{i}}_t \equiv \begin{bmatrix} \mathbf{z}_t \\ \mathbf{i}_t \\ \gamma_t \end{bmatrix} = \tilde{\mathbf{i}}(\mathbf{s}_t) \equiv \begin{bmatrix} \mathbf{z}(\mathbf{s}_t) \\ \mathbf{i}(\mathbf{s}_t) \\ \gamma(\mathbf{s}_t) \end{bmatrix} = F(\tilde{\mathbf{X}}_t, \mathbf{p}_{t|t}) \equiv \begin{bmatrix} F_{\mathbf{z}}(\tilde{\mathbf{X}}_t, \mathbf{p}_{t|t}) \\ F_{\mathbf{i}}(\tilde{\mathbf{X}}_t, \mathbf{p}_{t|t}) \\ F_{\gamma}(\tilde{\mathbf{X}}_t, \mathbf{p}_{t|t}) \end{bmatrix}, \tag{32}$$

$$\mathbf{x}_t = \mathbf{x}(\mathbf{s}_t, j_t, \varepsilon_t) \equiv \tilde{\mathbf{x}}(\mathbf{X}_t, \mathbf{z}(\mathbf{s}_t), \mathbf{i}(\mathbf{s}_t), j_t, \varepsilon_t) \equiv F_{\mathbf{x}}(\tilde{\mathbf{X}}_t, \mathbf{p}_{t|t}, j_t, \varepsilon_t). \tag{33}$$

Because of the nonlinearity of equations (27) and (30), the solution is no longer linear in $\tilde{\mathbf{X}}_t$ for given $\mathbf{p}_{t|t}$. The dual value function, $\tilde{V}(\mathbf{s}_t)$, is no longer quadratic in $\tilde{\mathbf{X}}_t$ for given $\mathbf{p}_{t|t}$. The value function of the primal problem, $V(\mathbf{s}_t)$, is given by, equivalently, equation (19); equation (29) with the equilibrium transition equation (28) and with the solution (32); or

$$V(\mathbf{s}_t) = \sum_j p_{jt|t} \int \begin{bmatrix} L(\mathbf{X}_t, \mathbf{x}(\mathbf{s}_t, j, \varepsilon_t), \mathbf{i}(\mathbf{s}_t), j) \\ +\delta \sum_k P_{jk} V(\bar{g}(\mathbf{s}_t, j, \varepsilon_t, k, \varepsilon_{t+1})) \end{bmatrix} \varphi(\varepsilon_t)\varphi(\varepsilon_{t+1}) d\varepsilon_t d\varepsilon_{t+1}. \tag{34}$$

It it is also no longer quadratic in $\tilde{\mathbf{X}}_t$ for given $\mathbf{p}_{t|t}$. More complex and detailed numerical methods are thus necessary in this case to find the optimal policy and the value function. Therefore, little can be said in general about the solution of the problem. Nonetheless, in numerical analysis it is very useful to have a good starting guess at a solution, which here comes from the AOP case. In our examples below, we explain in more detail how the BOP and AOP cases differ and what drives the differences.

2. Learning and Experimentation in a Simple New-Keynesian Model

We consider the benchmark standard New-Keynesian model, consisting of a New-Keynesian Phillips curve and a consumption Euler equation:[15]

$$\pi_t = \delta E_t \pi_{t+1} + \gamma_{j_t} y_t + c_\pi \varepsilon_{\pi t}; \tag{35}$$

$$y_t = E_t y_{t+1} - \sigma_{j_t} (i_t - E_t \pi_{t+1}) + c_y \varepsilon_{yt} + c_g g_t; \tag{36}$$

$$g_{t+1} = \rho g_t + \varepsilon_{g,t+1}. \tag{37}$$

Here π_t is the inflation rate, y_t is the output gap, δ is the subjective discount factor (as above), γ_{j_t} is a composite parameter reflecting the elasticity of demand and frequency of price adjustment, and σ_{j_t} is the intertemporal elasticity of substitution. There are three shocks in the model: two unobservable shocks, $\varepsilon_{\pi t}$ and ε_{yt}, which are independent standard normal random variables, and the observable serially correlated shock, g_t. This last shock is interpretable as a demand shock coming from variation in preferences, government spending, or the underlying efficient level of output. Woodford (2003) combines and renormalizes these shocks into a composite shock representing variation in the natural rate of interest.

In the standard formulations of this model, the shocks are observable and policy responds directly to the shocks. However, some components of the shocks need to be unobservable in order for there to be a nontrivial inference problem for agents. We have assumed that both the slope of the Phillips curve, γ_{j_t}, and the interest sensitivity, σ_{j_t}, vary with the mode, j_t. For the former, this could reflect changes in the degree of monopolistic competition (which also lead to varying markups) or changes in the degree of price stickiness. The interest sensitivity shift is purely a change in the preferences of the agents in the economy, although it could also result from nonhomothetic preferences coupled with shifts in output (in which case the preferences themselves would not shift, but the intertemporal elasticity would vary with the level of output). Unlike our illustration above, there are no switches in the steady-state levels of the variables of interest here,

15. See Woodford (2003) for an exposition.

as we consider the usual approximations around a zero inflation rate and an efficient level of output.

2.1 Optimal Policy: No Learning, Adaptive Optimal Policy, and Bayesian Optimal Policy

Here we examine value functions and optimal policies for this simple New-Keynesian model under no learning (NL), adaptive optimal policy (AOP), and Bayesian optimal policy (BOP). We use the following loss function:

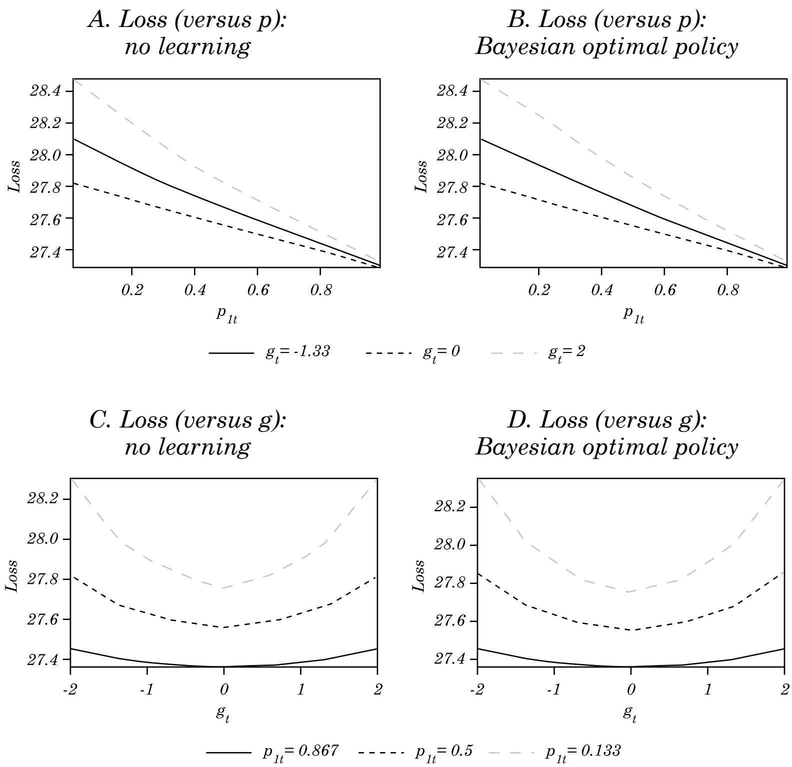
$$L_t = \pi_t^2 + \lambda_j y_t^2 + \mu i_t^2. \tag{38}$$

We set the following parameters, mostly following Woodford's (2003) calibration, as follows: $\gamma_1 = 0.024$, $\gamma_2 = 0.075$, $\sigma_1 = 1.000/0.157 = 6.370$, $\sigma_2 = 1.0$, $c_\pi = c_y = c_g = 0.5$, and $\rho = 0.5$. We set the loss function parameters as: $\delta = 0.99$, $\lambda_j = 2\gamma_j$, and $\mu = 0.236$. Most of the structural parameters are taken from Woodford (2003), while the two modes represent reasonable alternatives. Mode 1 is Woodford's benchmark case; mode 2 has a substantially smaller interest rate sensitivity (one consistent with logarithmic preferences) and a larger response, γ, of inflation to output. We set the transition matrix to

$$\mathbf{P} = \begin{bmatrix} 0.98 & 0.02 \\ 0.02 & 0.98 \end{bmatrix}.$$

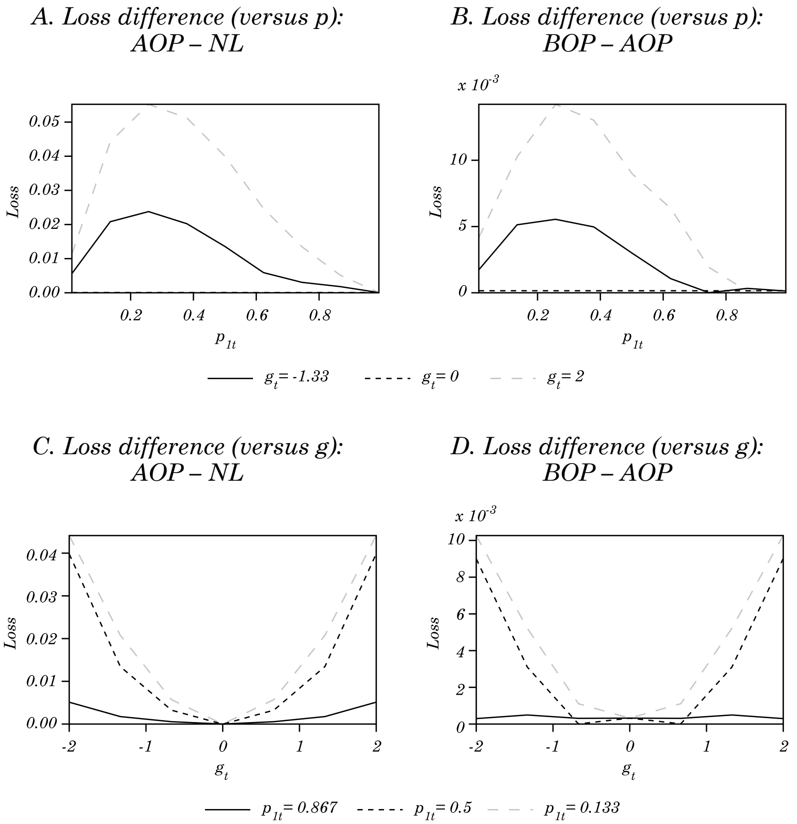
We have two forward-looking variables, $\mathbf{x}_t \equiv (\pi_t, y_t)'$, and consequently two Lagrange multipliers, $\mathbf{\Xi}_{t-1} \equiv (\Xi_{\pi,t-1}, \Xi_{y,t-1})'$. We have one predetermined variable ($\mathbf{X}_t \equiv g_t$) and the estimated mode probabilities, $\mathbf{p}_{t|t} \equiv (p_{1t|t}, p_{2t|t})'$ (of which we only need keep track of one, $p_{1t|t}$). Thus, the value and policy functions, $V(\mathbf{s}_t)$ and $i(\mathbf{s}_t)$, are all four dimensional: $\mathbf{s}_t = (g_t, \mathbf{\Xi}'_{t-1}, p_{1t|t})'$. We are therefore forced for computational reasons to restrict attention to relatively sparse grids with few points. The following plots show two-dimensional slices of the value and policy functions, focusing on the dependence on g_t and $p_{1t|t}$ (which we for simplicity denote by p_{1t} in the figures). In particular, all of the plots are for $\mathbf{\Xi}_{t-1} = (0, 0)'$. Figure 1 shows losses under NL and BOP as functions of p_{1t} and g_t. Figure 2 shows the difference between losses under NL, AOP, and BOP. Figures 3 and 4 show the corresponding policy functions and their differences.

Figure 1. Losses from No Learning and Bayesian Optimal Policy



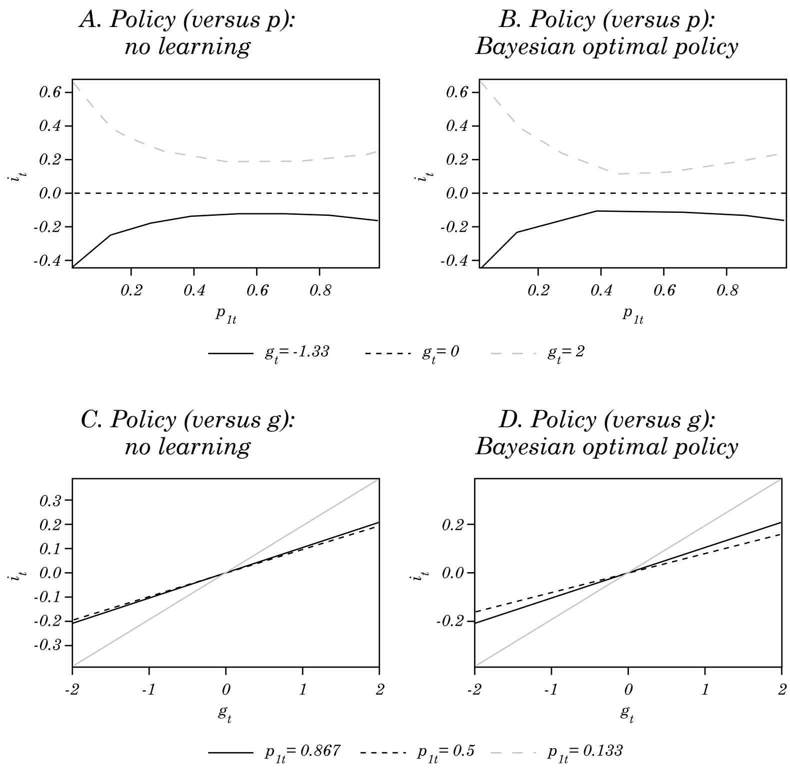
Source: Authors' calculations.

Figure 2. Differences in Losses from No Learning, Adaptive Optimal Policy, and Bayesian Optimal Policy



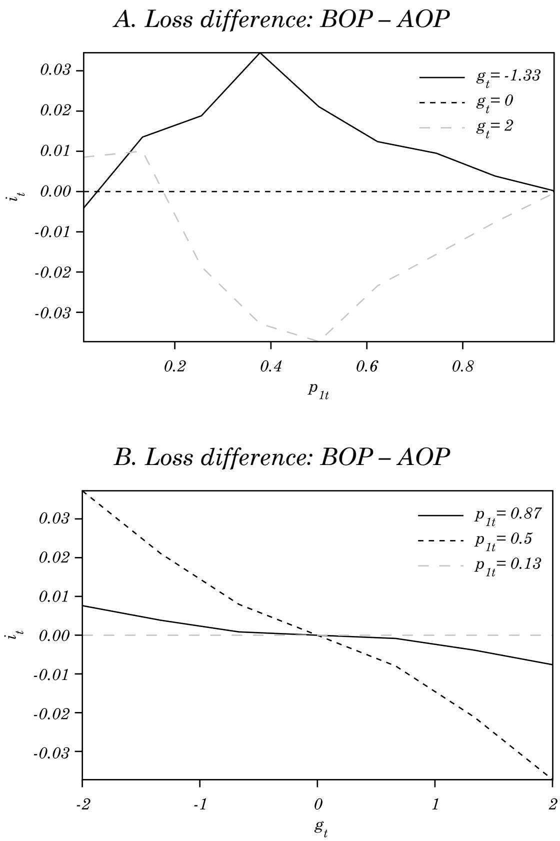
Source: Authors' calculations.

Figure 3. Optimal Policies under No Learning and Bayesian Optimal Policy



Source: Authors' calculations.

Figure 4. Differences in Policies under No Learning and Bayesian Optimal Policy



Source: Authors' calculations.

In Svensson and Williams (2007a) we show that learning implies a mean-preserving spread of the random variable $\mathbf{p}_{t+1|t+1}$ (which under learning is a random variable from the vantage point of period t). Hence, concavity of the value function under NL in p_{1t} implies that learning is beneficial, since then a mean-preserving spread reduces the expected future loss. However, figure 1 illustrates that the value function is actually slightly convex in p_{1t}, so learning is not beneficial here. In contrast, the value function is concave and learning is beneficial in a backward-looking example in Svensson and Williams (2007a).

Consequently, AOP gives higher losses than NL, as shown in figure 2. Furthermore, somewhat surprisingly, BOP gives higher losses than AOP (although the difference is very small). This is all counter to an example with a backward-looking model in Svensson and Williams (2007a).

Why is this different in a model with forward-looking variables? It may at least partially be a remnant of our assumption of symmetric beliefs and information between the private sector and the policymaker. Backward-looking models generally find that learning is beneficial. Moreover, with backward-looking models, the BOP is always weakly better than the AOP, as acknowledging the endogeneity of information in the BOP case need not mean that policy must change. (That is, the AOP policy is always feasible in the BOP problem.) Neither of these conclusions holds with forward-looking models. Under our assumption of symmetric information and beliefs between the private sector and the policymaker, both the private sector and the policymaker learn. The difference then comes from the way that private sector beliefs also respond to learning and to the experimentation motive. Having more reactive private sector beliefs may add volatility and make it more difficult for the policymaker to stabilize the economy. Acknowledging the endogeneity of information in the BOP case then need not be beneficial either, as it may induce further volatility in agents' beliefs.[16]

3. Learning in an Estimated Empirical New-Keynesian Model

The previous section focused on a simple small model to explore the impacts of learning and experimentation. Since computing

16. In the forward-looking case, we solve saddlepoint problems, and moving from AOP to BOP expands the feasible set for both the minimizing and maximizing choices.

BOP is computationally intensive, there are limits to the degree of empirical realism of the models we can address in that framework. In this section, we focus on a more empirically plausible model, using a version of Lindé's (2005) model that we estimated in Svensson and Williams (2007b). This model includes richer dynamics for inflation and the output gap, which both have backward- and forward-looking components. However, these additional dynamics increase the dimension of the state space, which implies that it is not very feasible to consider the BOP. We therefore focus on the impact of learning on policy and compare NL and AOP. In Svensson and Williams (2007b), we computed the optimal policy under no learning, and here we see how inference on the mode affects the dynamics of output, inflation, and interest rates.

3.1 The Model

The structural model is a mode-dependent simplification of Lindé's (2005) model of the U.S. economy and is given by

$$\begin{aligned} \pi_t &= \omega_{fj} E_t \pi_{t+1} + (1-\omega_{fj})\pi_{t-1} + \gamma_j y_t + c_{\pi j}\varepsilon_{\pi t}; \\ y_t &= \beta_{fj} E_t y_{t+1} + (1-\beta_{fj})\left[\beta_{yj} y_{t-1} + (1-\beta_{yj}) y_{t-2}\right] \\ &\quad -\beta_{rj}(i_t - E_t \pi_{t+1}) + c_{yj}\varepsilon_{yt}. \end{aligned} \tag{39}$$

Here $j \in \{1, 2\}$ indexes the mode, and the shocks, $\varepsilon_{\pi t}$, ε_{yt}, and ε_{it}, are independent standard normal random variables. In particular, we consider a two-mode MJLQ model in which one mode has forward- and backward-looking elements and the other is backward-looking only. Thus we specify that mode 1 is unrestricted, while in mode 2 we restrict $\omega_f = \beta_f = 0$, so that the mode is backward-looking. For estimation, we also impose a particular instrument rule for i_t, but we do not include that here since our focus is on optimal policy.

In Svensson and Williams (2007b), we estimate the model on U.S. data using Bayesian methods. The maximum posterior estimates are given in table 1, with the unconditional expectation of the coefficients for comparison. Apart from the forward-looking terms (which are restricted), the variation in the other parameters across the modes is relatively minor. There are some differences in the estimated policy functions (not reported here), but relatively little change across modes in the other structural coefficients. The estimated transition matrix **P** and its implied stationary distribution $\bar{\mathbf{p}}$ are given by

$$\mathbf{P} = \begin{bmatrix} 0.9579 & 0.0421 \\ 0.0169 & 0.9831 \end{bmatrix}, \qquad \bar{\mathbf{p}} = \begin{bmatrix} 0.2869 \\ 0.7131 \end{bmatrix}.$$

Table 1. Estimates of the Constant-Coefficient Model and a Restricted Two-Mode Lindé Model

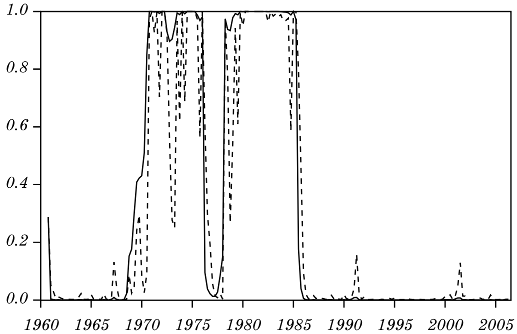
Parameter	*Mean*	*Mode 1*	*Mode 2*
ω_f	0.0938	0.3272	0.0000
γ	0.0474	0.0580	0.0432
β_f	0.1375	0.4801	0.0000
β_r	0.0304	0.0114	0.0380
β_y	1.3331	1.5308	1.2538
c_π	0.8966	1.0621	0.8301
c_y	0.5572	0.5080	0.5769

Source: Authors' calculations.

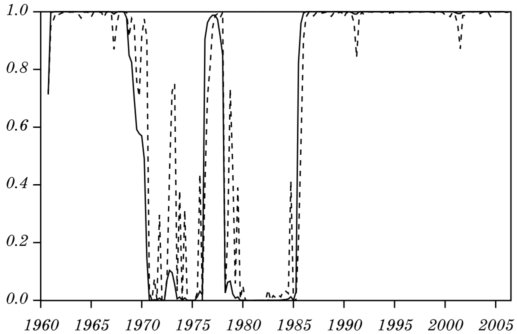
Mode 2 is thus the most persistent and has the largest mass in the invariant distribution. This is consistent with our estimation of the modes, as shown in figure 5. Again, the plots show both the smoothed and filtered estimates. Mode 2, the backward-looking model mode, was experienced the most throughout much of the sample, holding for 1961–68 and then, with near certainty, continually since 1985. The forward-looking model held in periods of rapid changes in inflation, holding for both the run-ups in inflation in the early and late 1970s and the disinflationary period of the early 1980s. During periods of relative tranquility, such as the Greenspan era, the backward-looking model fits the data the best.

Figure 5. Estimated Probabilities of Being the Different Modes[a]

A. Probabiliy in mode 1



B. Probabiliy in mode 2



Source: Authors' calculations.

a. In the figure, solid lines graph the smoothed (full-sample) inference, while dashed lines represent the filtered (one-sided) inference.

3.2 Optimal Policy: No Learning and Adaptive Optimal Policy

Using the methods described above, we solve for the optimal policy functions

$$i_t = F_i(\mathbf{p}_{t|t})\tilde{\mathbf{X}}_t,$$

where now $\tilde{\mathbf{X}}_t \equiv (\pi_{t-1}, y_{t-1}, y_{t-2}, i_{t-1}, \Xi_{\pi,t-1}, \Xi_{y,t-1})'$. In Svensson and Williams (2007b), we focus on the observable and no-learning cases, and we assume that the shocks $\varepsilon_{\pi t}$ and ε_{yt} are observable. We thus set $C_2 \equiv 0$ and treat the shocks as additional predetermined variables. To focus on the role of learning, we now assume that those shocks are unobservable. If they were observable, then agents would be able to infer the mode from their observations of the forward-looking variables. We use the following loss function:

$$L_t = \pi_t^2 + \lambda y_t^2 + \nu(i_t - i_{t-1})^2, \tag{40}$$

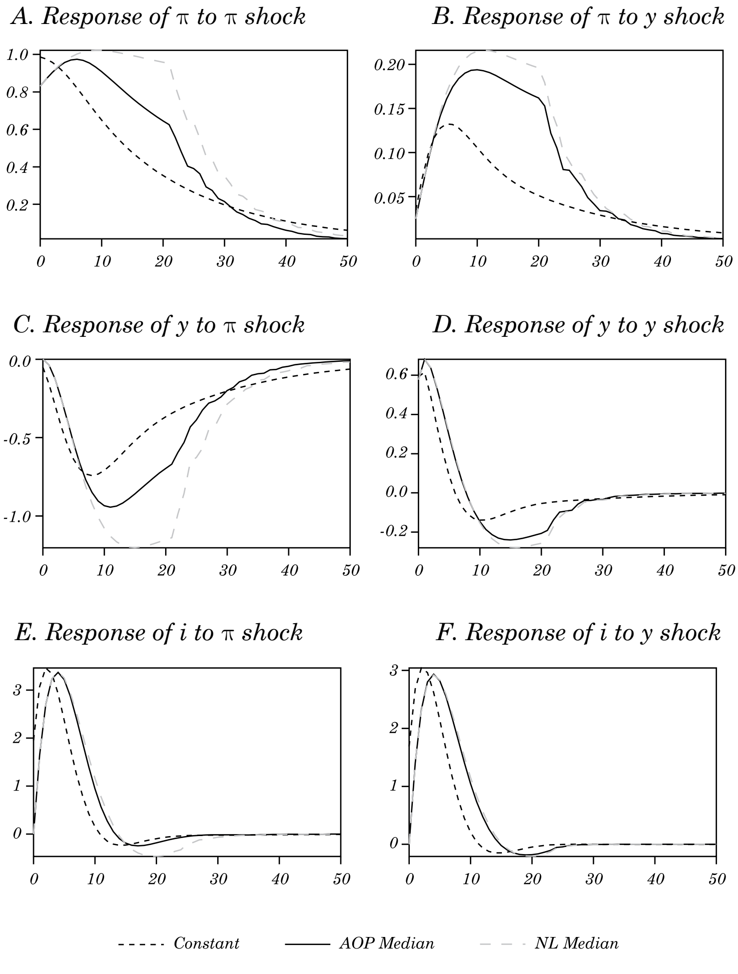
which is a common central bank loss function. We set the weights to $\lambda = 1$ and $\nu = 0.2$, and fix the discount factor in the intertemporal loss function to $\delta = 1$.

For ease of interpretation, we plot the distribution of the impulse responses of inflation, the output gap, and the instrument rate to the two structural shocks in figure 6. We consider 10,000 simulations of fifty periods, and we plot the median responses for the optimal policy under NL and AOP and the corresponding optimal responses for the constant-coefficient model.[17]

Compared with the constant-coefficient case, the mean impulse responses are consistent with larger effects of the shocks that are also longer lasting. In terms of the optimal policy responses, the AOP and NL cases are quite similar, and in both cases the peak response to shocks is nearly the same as in the constant-coefficient case, but it comes with a delay. Again compared with the constant-coefficient case, the responses of inflation and the output gap are larger and more sustained when there is model uncertainty.

17. The shocks are $\varepsilon_{\pi 0} = 1$ and $\varepsilon_{y0} = 1$, respectively, so the shocks to the inflation and output-gap equations in period 0 are mode dependent and equal to $c_{\pi j}$ and c_{yj} (j = 1, 2, 3), respectively. The distribution of modes in period 0 (and thereby all periods) is again the stationary distribution.

Figure 6. Unconditional Impulse Responses to Shocks under the Optimal Policy for the Two-Mode Version of the Lindé Model[a]



Source: Authors' calculations.
a. In the figure, solid lines represent the median responses under AOP, dashed lines represent the median responses under NL, and dot-dashed lines represent the constant-coefficient responses.

Learning can be beneficial, however, as the optimal policy under AOP dampens the responses to shocks, particularly for shocks to inflation. Since the optimal policy responses are nearly identical, this seems to be largely due to more accurate forecasts by the public, which lead to more rapid stabilization.

While these impulse responses are revealing, they do not capture the full benefits of learning, as by definition they simply provide the responses to a single shock. To gain a better understanding of the role of learning, we simulated our model under the NL and AOP policies to compare the realized economic performance. Table 2 summarizes various statistics resulting from a thousand simulations of a thousand periods each. Thus, for example, the entry for the average π_t is the average across the thousand simulations of the sample average (over the thousand periods) of inflation, while the standard deviation of π_t is the average across simulations of the standard deviation (in each time series) of inflation. In the table, the average period loss (L_t) under AOP is less than half that under NL. Figure 7 plots the distribution across samples of the key components of the loss function. There we plot a kernel smoothed estimate of the distribution from the thousand simulations. The figure shows that the distribution of sample losses is much more favorable under AOP than under NL.

In figure 8 we show one representative simulation to illustrate the differences. The figure reveals that the stabilization of inflation and the output gap are more effective under AOP than NL for very similar instrument rate settings.

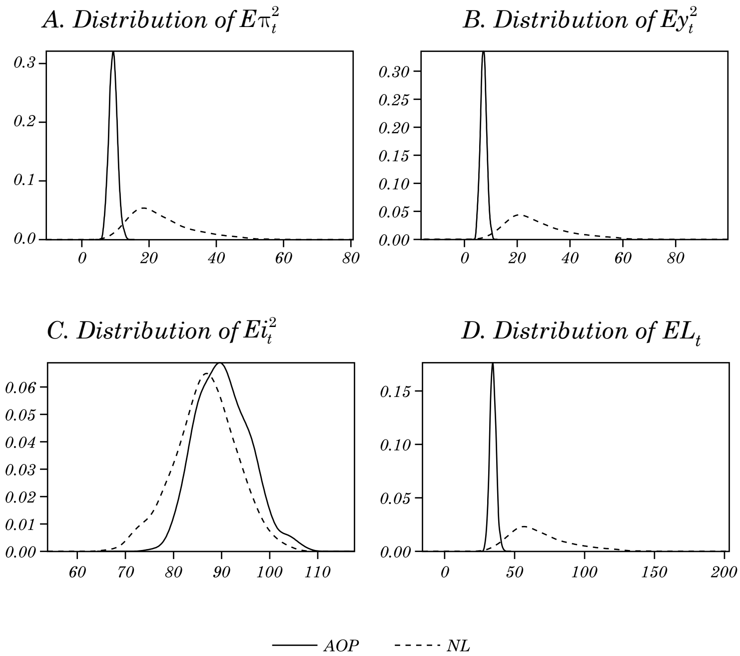
Table 2. Average of Different Statistics under No Learning and the Adaptive Optimal Policy[a]

	π_t		y_t		i_t		L_t
Policy	*Average*	*Std. dev.*	*Average*	*Std. dev.*	*Average*	*Std. dev.*	*Average*
NL	−0.1165	5.2057	0.1303	5.6003	0.0073	10.0239	88.4867
AOP	−0.0300	3.1696	0.0299	2.7698	0.0011	9.9989	38.8710

Source: Authors' calculations.

a. The table presents the average value and average standard deviation of each variable from one thousand simulations of one thousand periods each of our estimated model under the no learning (NL) and under the adaptive optimal policy (AOP).

Figure 7. Distribution across Samples of Various Statistics under the Optimal Policy for the Two-Mode Version of the Lindé Model



Source: Authors' calculations.

Figure 8. Simulated Time Series under the Optimal Policy for the Two-Mode Version of the Lindé Model[a]

A. Inflation

B. Output gap

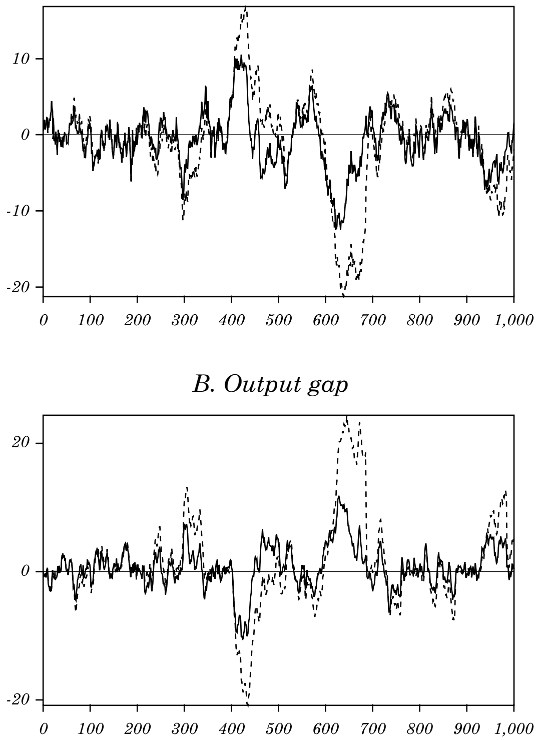
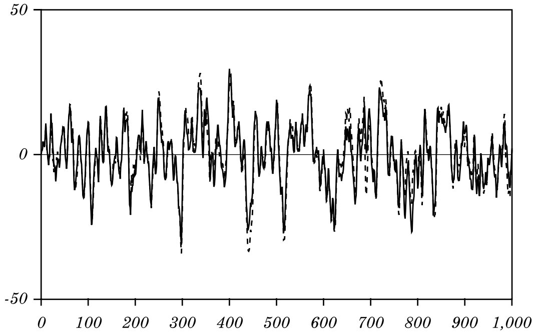
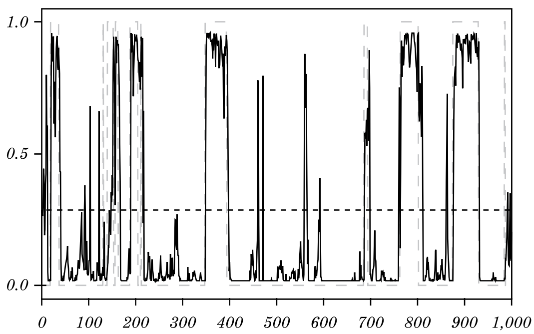


Figure 8. (continued)

C. Interest rate



D. Probability in mode 1



Source: Authors' calculations.

a. In panels A, B, and C, solid lines denote AOP, while dashed lines graph NL. In panel D, the solid line represents the probability of mode 1, the dotted line represents the true mode, and the dashed gray line represents the unconditional probability of mode 1.

4. Conclusions

In this paper, we have presented a relatively general framework for analyzing model uncertainty and the interactions between learning and optimization. While this is a classic issue, very little has been done to date for systems with forward-looking variables, which are essential elements of modern models for policy analysis. Our specification is general enough to cover many practical cases of interest, yet remains relatively tractable in implementation. This is definitely true when decisionmakers do not learn from the data they observe (our case of no learning, NL) or when they do learn but do not account for learning in optimization (our case of adaptive optimal policy, AOP). In both of these cases, we have developed efficient algorithms for solving for the optimal policy, which can handle relatively large models with multiple modes and many state variables. However, in the case of the Bayesian optimal policy (BOP), which takes the experimentation motive into account, we must solve more complex numerical dynamic programming problems. Thus to fully examine optimal experimentation, we are haunted by the curse of dimensionality, forcing us to study relatively small and simple models.

An issue of much practical importance is the size of the experimentation component of policy and the losses entailed in abstracting from it. While our results in this paper are far from comprehensive, they suggest that the experimentation motive may not be a concern in practical settings. The above and similar examples that we have considered indicate that the benefits of learning (moving from NL to AOP) may be substantial, whereas the benefits from experimentation (moving from AOP to BOP) are modest or even insignificant. If this preliminary finding stands up to scrutiny, experimentation in economic policy in general and monetary policy in particular may not be very beneficial, in which case there is little need to face the difficult ethical and other issues involved in conscious experimentation in economic policy. Furthermore, the AOP is much easier to compute and implement than the BOP. More simulations and cases need to be examined for this to truly be a robust implication.

References

Beck, G.W. and V. Wieland. 2002. "Learning and Control in a Changing Economic Environment." *Journal of Economic Dynamics and Control* 26(9–10): 1359–77.

Blake, A.P. and F. Zampolli. 2006. "Time Consistent Policy in Markov Switching Models." Working paper 298. London: Bank of England.

Cogley, T., R. Colacito and T.J. Sargent. 2007. "Benefits from U.S. Monetary Policy Experimentation in the Days of Samuelson and Solow and Lucas." *Journal of Money, Credit, and Banking* 39(s1): 67–99.

Costa, O.L.V., M.D. Fragoso and R.P. Marques. 2005. *Discrete-Time Markov Jump Linear Systems***.** London: Springer.

Do Val, J.B.R. and T. Basar. 1999. "Receding Horizon Control of Jump Linear Systems and a Macroeconomic Policy Problem." *Journal of Economic Dynamics and Control* 23(8): 1099–131.

Ellison, M. 2006. "The Learning Cost of Interest Rate Reversals." *Journal of Monetary Economics* 53(8): 1895–907.

Ellison, M. and N. Valla. 2001. "Learning, Uncertainty, and Central Bank Activism in an Economy with Strategic Interactions." *Journal of Monetary Economics* 48(1): 153–71.

Evans, G. and S. Honkapohja. 2001. *Learning and Expectations in Macroeconomics*. Princeton University Press.

Kiefer, N.M. 1989. "A Value Function Arising in the Economics of Information." *Journal of Economic Dynamics and Control* 13(2): 201–23.

Lindé, J. 2005. "Estimating New-Keynesian Phillips Curves: A Full Information Maximum Likelihood Approach." *Journal of Monetary Economics* 52(6): 1135–49.

Marcet, A. and R. Marimon. 1998. "Recursive Contracts." Economics working paper 337. Universitat Pompeu Fabra, Department of Economics and Business.

Rudebusch, G.D. and L.E.O. Svensson. 1999. "Policy Rules for Inflation Targeting." In *Monetary Policy Rules*, edited by J.B. Taylor. University of Chicago Press.

Svensson, L.E.O. 2007. "Optimization under Commitment and Discretion, the Recursive Saddlepoint Method, and Targeting Rules and Instrument Rules: Lecture Notes." Available at www.princeton.edu/svensson.

Svensson, L.E.O. and N. Williams. 2007a. "Bayesian and Adaptive Optimal Policy Under Model Uncertainty." Working paper 13414. Cambridge, Mass.: National Bureau of Economic Research.

————. 2007b. "Monetary Policy with Model Uncertainty: Distribution Forecast Targeting." Discussion paper 6331. London: Centre for Economic Policy Research.

Tesfaselassie, M.F., E. Schaling and S.C.W. Eijffinger. 2006. "Learning about the Term Structure and Optimal Rules for Inflation Targeting." Discussion paper 5896. London: Centre for Economic Policy Research.

Wieland, V. 2000. "Learning by Doing and the Value of Optimal Experimentation." *Journal of Economic Dynamics and Control* 24(4): 501–34.

————. 2006. "Monetary Policy and Uncertainty about the Natural Unemployment Rate: Brainard-Style Conservatism versus Experimental Activism." *Advances in Macroeconomics* 6(1): 1288–97.

Woodford, M. 2003. *Interest and Prices*: *Foundations of a Theory of Monetary Policy*. Princeton University Press.

Zampolli, F. 2006. "Optimal Monetary Policy in a Regime-Switching Economy: The Response to Abrupt Shifts in Exchange-Rate Dynamics." Working paper 297. London: Bank of England.

Imperfect Knowledge and the Pitfalls of Optimal Control Monetary Policy

Athanasios Orphanides
Central Bank of Cyprus

John C. Williams
Federal Reserve Bank of San Francisco

Sixty years ago, Milton Friedman questioned the usefulness of the optimal control approach because of policymakers' imperfect knowledge of the economy and favored instead a simple rule approach to monetary policy (1947, 1948). These are still live issues, despite the development of powerful techniques to derive and analyze optimal control policies, which central banks use in their large-scale models (see Svensson and Woodford, 2003; Woodford, 2003; Giannoni and Woodford, 2005; Svensson and Tetlow, 2005). Although the optimal control approach provides valuable insights, it also presents problems. In particular, because it assumes a single correctly specified reference model, it ignores important sources of uncertainty about the economy that monetary policymakers face. Robust control methods of the type analyzed by Hansen and Sargent (2007) extend the standard optimal control approach to allow for unspecified model uncertainty; however, these methods are designed for relatively modest deviations from the reference model.[1] In practice, policymakers

At the time of the conference, Noah Williams was affiliated with Princeton University.

We thank Jim Bullard, Richard Dennis, Bob Hall, Sheryl Kennedy, Andy Levin, Julio Rotemberg, Klaus Schmidt-Hebbel, David Stockman, Carl E. Walsh, Volker Wieland, Mike Woodford, and Hakan Yilmazkuday for helpful comments and suggestions. The opinions expressed are those of the authors and do not necessarily reflect views of the Central Bank of Cyprus, the Governing Council of the European Central Bank, the management of the Federal Reserve Bank of San Francisco, or the Board of Governors of the Federal Reserve System.

1. Svensson and Williams (2007) propose a method to compute optimal policy under model uncertainty using a Markov-switching framework. Computing optimal policies under model uncertainty with this method is extremely computationally intensive, and its application to real-world problems remains infeasible.

Monetary Policy under Uncertainty and Learning, edited by Klaus Schmidt-Hebbel and Carl E. Walsh, Santiago, Chile.

are concerned with more fundamental sources of model uncertainty, and the robustness of monetary policy strategies to uncertainty is generally viewed as important (McCallum, 1988; Taylor, 1993). Thus, a key question is whether our understanding of the macroeconomic environment has improved enough to make the optimal control approach to monetary policy preferable to well-designed simple rules.

Relatively little research to date explores the robustness properties of optimal control policies to moderate or large degrees of model misspecification.[2] Gianonni and Woodford (2005) show that optimal control policies are robust to misspecification of the shock processes as long as the central bank forecasts are optimal. In contrast, Levin and Williams (2003) show that optimal control policies can perform very poorly if the structural equations of the central bank's reference model are badly misspecified. Orphanides and Williams (2008) examine the robustness of optimal control policies if the reference model misspecifies the way private agents form expectations. That paper finds that if private agents are uncertain of the true model and form expectations based on an estimated forecasting model, then optimal control policies designed under the assumption of rational expectations can perform poorly. The paper also shows that optimal control policies can be made more robust to this type of model uncertainty by placing less weight on stabilizing economic activity and interest rates in the central bank objective used in deriving the optimal control policy.

This paper extends the analysis in Orphanides and Williams (2008) to include uncertainty about the natural rates of interest and unemployment. We allow for exogenous time variation in the natural rates of interest and unemployment that the central bank may measure with error. There is considerable evidence of significant time variation in these natural rates and the difficulties of their real-time estimation (see, for example, Staiger, Stock, and Watson, 1997; Laubach, 2001; Orphanides and Williams, 2002; Laubach and Williams, 2003).[3] We assume that the central bank has a good understanding of the process describing the evolution of these natural rates, but may not observe

2. In contrast, there has been considerable research on the robustness of simple monetary policy rules to model uncertainty, including Taylor (1999), Levin, Wieland, and Williams (1999, 2003), Orphanides and Williams (2002, 2007), and Brock, Durlauf, and West (2007).

3. The natural rate of output is prone to considerable real-time mismeasurement, causing problems for monetary policy similar to the mismeasurement of the natural rate of unemployment, as discussed in Orphanides and others (2000), Orphanides and van Norden (2002), and Cukierman and Lippi (2005).

them directly, in which case it must estimate the natural rates using available data. We consider both the case in which the central bank uses the optimal statistical filter—the Kalman filter in the model of this paper—to estimate the natural rates, and the case in which the central bank's estimate of the key gain parameter of the filter is misspecified. Laubach and Williams (2003) and Clark and Kozicki (2005) document the imprecision in estimates of the gain parameter in the Kalman filter, making uncertainty about this key parameter a real-world problem for central bank estimates of natural rates.

We find that the optimal control policy derived assuming rational expectations and known natural rates performs relatively poorly in our estimated model of the U.S. economy when agents have imperfect knowledge of the structure of the economy, but instead must learn and the central bank must estimate movements in natural rates. The key shortcoming of the optimal control policy derived under the assumption of perfect knowledge is that it is overly fine-tuned to the assumptions in the benchmark model. As a result, the optimal control policy works extremely well when private and central bank knowledge are perfect. When agents learn, however, and the central bank may make mistakes due to misperceptions of natural rates, expectations can deviate from those implied under perfect knowledge, and the finely-tuned optimal control policy can go awry. In particular, by implicitly assuming that inflation expectations are always well anchored, the optimal control policy responds insufficiently strongly to movements in inflation, which results in excessive variability of inflation.

We then seek to construct policies that take advantage of the optimal control approach, but are robust to the forms of imperfect knowledge that we study.[4] Specifically, following the approach in Orphanides and Williams (2008), we look for weights in the central bank objective function such that an optimal control policy derived using these "biased" weights performs well under imperfect knowledge about the structure of the economy. We find that optimal policies derived assuming much lower weights on stabilizing economic activity and interest rates than in the true central bank objective perform well in the presence of both private agent learning and natural rate uncertainty. Relative to our earlier results, the incorporation of natural

4. An alternative approach, followed by Gaspar, Smets, and Vestin (2006), is to derive optimal monetary policy under learning. Because the model with learning is nonlinear, they apply dynamic programming techniques that are infeasible for the type of models studied in this paper and used in central banks for monetary policy analysis.

rate uncertainty further reduces the optimal weights on economic activity and interest rates in the objective function used in deriving optimal policies that are robust to imperfect knowledge.

Finally, we compare the performance of optimal control policies to two types of simple monetary policy rules that have been found to be robust to various types of model uncertainty in the literature. The first is a forward-looking version of a Taylor-type policy rule, similar to the rule that Levin, Wieland, and Williams (2003) found to perform very well in a number of estimated rational expectations models of the U.S. economy. The second is the rule proposed by Orphanides and Williams (2007), which differs from the first rule in that policy responds to the change in the measure of economic activity, rather than the level. This type of rule has been shown to be robust to mismeasurement of natural rates in the economy (Orphanides and Williams, 2002, 2007) and to perform very well in a counterfactual analysis of monetary policy in 1996–2003 (Tetlow, 2006). Under rational expectations, these rules perform somewhat worse than the optimal control policy.

The two simple monetary policy rules perform very well under learning and with natural rate mismeasurement. These rules clearly outperform the optimal control policy when knowledge is imperfect and generally perform about as well as the optimal control policies derived to be robust to imperfect knowledge by using a biased objective function. The relatively small advantage that the optimal control policy has over these robust rules when the model is correctly specified implies that the "insurance" payment required to gain the sizable robustness benefits found here is quite small.

The remainder of the paper is organized as follows. Section 1 describes the model and its estimation. Section 2 describes the central bank objective and the optimal control policy. Section 3 describes the models of expectation formation and the simulation methods. Section 4 examines the performance of the optimal control policy under imperfect knowledge. Section 5 analyzes the optimal weights in the central bank objective that yield robust optimal control policies that perform well under imperfect knowledge. Section 6 compares the performance of the simple rules to optimal control policies. Section 7 concludes.

1. An Estimated Model of the U.S. Economy

Our analysis is conducted using an estimated quarterly model of the U.S. economy. The basic structure of the model is the same as in Orphanides and Williams (2008), but it is extended to incorporate time variation in the natural rates of interest and unemployment. The model consists of equations that describe the dynamic behavior of the unemployment rate and the inflation rate and equations describing the natural rates of interest and unemployment and the shocks. To close the model, the short-term interest rate is set by the central bank, as described in the next section.

1.1 The Model

The IS curve equation is motivated by the Euler equation for consumption with adjustment costs or habit:

$$u_t = \phi_u u_{t+1}^e + (1-\phi_u)u_{t-1} + \alpha_u(i_t^e - \pi_{t+1}^e - r_t^*) + v_t; \tag{1}$$

$$v_t = \rho_v \, v_{t-1} + e_{v,t}, \qquad e_v \sim N\left(0, \sigma_{e_v}^2\right). \tag{2}$$

We specify the IS equation in terms of the unemployment rate rather than output to facilitate the estimation of the equation using real-time data. This equation relates the unemployment rate, u_t, to the unemployment rate expected in the next period, one lag of the unemployment rate, and the difference between the expected ex ante real interest rate (equal to the difference between the nominal short-term interest rate, i_t, and the expected inflation rate in the following period, π_{t+1}) and the natural rate of interest, r_t^*. The unemployment rate is subject to a shock, v_t, that is assumed to follow a first-order autoregressive, or AR(1), process with innovation variance $\sigma_{e_v}^2$. The AR(1) specification for the shocks is based on the evidence of serial correlation in the residuals of the estimated unemployment equation, as discussed below.

The Phillips curve equation is motivated by the New-Keynesian Phillips curve with indexation:

$$\pi_t = \phi_\pi \pi_{t+1}^e + (1-\phi_\pi)\pi_{t-1} + \alpha_\pi\left(u_t - u_t^*\right) + e_{\pi,t}, \qquad e_\pi \sim N\left(0, \sigma_{e_\pi}^2\right). \tag{3}$$

It relates inflation, π_t, (measured as the annualized percent change in the GNP or GDP price index, depending on the period) during quarter t to lagged inflation, expected future inflation, denoted by π_{t+1}^e, and the difference between the unemployment rate, u_t, and the natural rate of unemployment, u_t^*, in the current quarter. The parameter ϕ_π measures the importance of expected inflation on the determination of inflation, while $(1 - \phi_\pi)$ captures the effects of inflation indexation. The mark-up shock, $e_{\pi,t}$, is assumed to be a white noise disturbance with variance $\sigma^2_{e_\pi}$.

We model the low-frequency behavior of the natural rates of unemployment and interest as exogenous AR(1) processes independent of all other variables:

$$u_t^* = (1 - \rho_{u^*})\bar{u}^* + \rho_{r^*} u_{t-1}^* + e_{u^*,t}, \qquad e_{u^*} \sim N\left(0, \sigma^2_{e_u^*}\right); \tag{4}$$

$$r_t^* = (1 - \rho_{r^*})\bar{r}^* + \rho_{u^*} r_{t-1}^* + e_{r^*,t}, \qquad e_{r^*} \sim N\left(0, \sigma^2_{e_r^*}\right). \tag{5}$$

We assume these processes are stationary based on the finding using the standard augmented Dickey-Fuller (ADF) test that one can reject the null of nonstationarity of both the unemployment rate and real federal funds rate over 1950–2003 at the 5 percent level. The unconditional mean values of the natural rates are irrelevant to the policy analysis, so we set them both to zero.[5]

1.2 Model Estimation and Calibration

The details of the estimation method for the inflation and unemployment rate equations are described in Orphanides and Williams (2008). The estimation results are reported below, with standard errors indicated in parentheses.

Unrestricted estimation of the IS curve equation yields a point estimate for ϕ_u of 0.39, with a standard error of 0.15. This estimate is below the lower bound of 0.5 implied by theory; however, the null hypothesis of a value of 0.5 is not rejected by the data.[6] We therefore impose $\phi_u = 0.5$ in estimating the remaining parameters of

5. Because we ignore the zero lower bound on nominal interest rates, as well as any other potential source of nonlinear behavior in the structural model, the unconditional means of variables are irrelevant. Inclusion of the zero bound would severely complicate the analysis and is left for future work.

6. This finding is consistent with the results reported by Giannoni and Woodford (2005), who find, in a similar model, that the corresponding coefficient is constrained to be at its theoretical lower bound.

the equation. The estimated equation also includes a constant term (not shown) that provides an estimate of the natural real interest rate, which is assumed to be constant for the purpose of estimating this equation.

$$u_t = 0.5u^e_{t+1} + 0.5u_{t-1} + \underset{(0.022)}{0.056}(\tilde{r}^e_t - r^*) + v_t, \tag{6}$$

$$v_t = \underset{(0.085)}{0.513}v_{t-1} + e_{v,t}, \qquad \hat{\sigma}_{e_v} = 0.30, \tag{7}$$

$$\pi_t = 0.5\pi^e_{t+1} + 0.5\pi_{t-1} - \underset{(0.087)}{0.294}(u_t - u^*_t) + e_{\pi,t}, \qquad \hat{\sigma}_{e_\pi} = 1.35. \tag{8}$$

Unrestricted estimation of the Phillips curve equation yields a point estimate for ϕ_π of 0.51, just barely above the lower bound implied by theory.[7] For symmetry with our treatment of the IS curve, we impose $\phi_\pi = 0.5$ and estimate the remaining parameters using ordinary least squares (OLS). The estimated residuals for this equation show no signs of serial correlation in the price equation (Durbin-Watson = 2.09), consistent with the assumption of the model.

There is considerable uncertainty regarding the magnitude and persistence of low-frequency fluctuations in the natural rates of unemployment and interest (see Staiger, Stock, and Watson, 1997; Laubach, 2001; Orphanides and Williams, 2002; Laubach and Williams, 2003; Clark and Kozicki, 2005.) We do not estimate a model of natural rates; instead, we calibrate the parameters of the AR(1) processes based on estimates found elsewhere in the literature. To capture the highly persistent movements in natural rates, we set the autocorrelation parameters, ρ_{u^*} and ρ_{r^*}, to 0.99. In our benchmark calibration, we set the innovation standard deviation of the natural rate of unemployment to 0.07 and that of the natural rate of interest to 0.085. These values imply an unconditional standard deviation of the natural rate of unemployment (interest) of 0.50 (0.60), in the low end of the range of standard deviations of smoothed estimates of these natural rates suggested by various estimation methods. We also consider an alternative calibration in which the standard deviations of the natural rate innovations are twice as large, consistent with the upper end of the range of estimates of natural rate variation.

7. For comparison, Giannoni and Woodford (2005) find that the corresponding coefficient is constrained to be at its theoretical lower bound of 0.5.

2. Optimal Control Monetary Policy

In this section, we describe the optimal control monetary policy. The policy instrument is the nominal short-term interest rate. We assume that the central bank observes all variables from previous periods when making the current-period policy decision. We further assume that the central bank has access to a commitment technology; that is, we study policy under commitment.

The central bank's objective is to minimize a loss equal to the weighted sum of the unconditional variances of the inflation rate, the difference between the unemployment rate and the natural rate of unemployment, and the first-difference of the nominal federal funds rate:

$$L = \mathrm{var}(\pi - \pi^*) + \lambda\, \mathrm{var}(u - u^*) + \nu\, \mathrm{var}(\Delta(i)), \tag{9}$$

where var(x) denotes the unconditional variance of variable x. We assume an inflation target of zero percent. As a benchmark for our analysis, we assume $\lambda = 4$ and $\nu = 1$. Based on an Okun's Law relationship, the variance of the unemployment gap is about one-quarter that of the output gap, so this choice of λ corresponds to equal weights on inflation and output gap variability.

The optimal control monetary policy is that which minimizes the loss subject to the equations describing the economy. We construct the optimal control policy, as is typical in the literature, assuming that the policymaker knows the true parameters of the structural model and assumes all agents use rational expectations and the central bank knows the natural rates of unemployment and interest.[8] For the optimal control policy, as well as the simple monetary policy rules described below, we use lagged information in the determination of the interest rate, reflecting the lag in data releases. The optimal control policy is described by a set of equations representing the first-order optimality condition for policy and the behavior of the Lagrange multipliers associated with the constraints on the optimization problem implied by the structural equations of the model economy.

Because we are interested in describing the setting of interest rates in a potentially misspecified model, it is useful to represent the

8. See, for example, Sargent's (2007) description of the optimal policy approach.

optimal control policy in an equation that relates the policy instrument to macroeconomic variables, rather than in terms of Lagrange multipliers that depend on the model. There are infinitely many such representations. In the following, we focus on one representation of the optimal control (OC) policy. In the OC policy, the current interest rate depends on three lags of the following variables: the inflation rate, the difference between the unemployment rate and the central bank's estimate of the natural rate of unemployment, and the difference between the nominal interest rate and the estimate of the natural rate of interest. The OC representation yields a determinate rational expectations equilibrium. We find that including three lags of these variables is sufficient to very closely mimic the optimal control outcome assuming the central bank observes natural rates.

2.1 Central Bank Estimation of Natural Rates

As noted above, we compute the OC policy assuming the central bank observes the true values of the natural rates of interest and unemployment. In our policy evaluation exercises, we consider the possibility that the central bank must estimate natural rates in real time. In such cases, we assume that the central bank knows the true structure of the model, including the model parameters (and the unconditional means of the natural rates), and observes all other variables including private forecasts, but does not observe the shocks directly. Given our model, the Kalman filter is the optimal method for estimating the natural rates, and we assume that the central bank uses the appropriate specification of the Kalman filters to estimate natural rates. These assumptions represent a best case for the central bank with respect to its ability to estimate natural rates. In other work, we examine the implications of model uncertainty regarding the data-generating processes for natural rates (Orphanides and Williams, 2005, 2007).

The central bank's real-time estimate of the natural rate of unemployment, $\hat{u}_t^*$, is given by

$$\hat{u}_t^* = a_1 \hat{u}_{t-1}^* + a_2 \left(u_t^* - \frac{e_{\pi,t}}{\alpha_\pi} \right), \tag{10}$$

where a_1 and a_2 are the Kalman gain parameters and the term within the parentheses is the current-period shock to inflation, which incorporates the effects of the transitory inflation disturbance and the deviation of the natural rate of unemployment from its unconditional mean, scaled in units of the unemployment rate. The central bank only observes this surprise and not the decomposition into its two components.

The central bank estimate of the natural rate of interest, $\hat{r}_t^*$, is given by

$$\hat{r}_t^* = b_1\hat{r}_{t-1}^* + b_2\left(r_t^* - \frac{v_t}{\alpha_u}\right) + b_3\left(r_{t-1}^* - \frac{v_{t-1}}{\alpha_u}\right), \tag{11}$$

where the first term in parentheses is the current-period unemployment rate shock and the final term is the lagged shock. The final term appears in the equation due to the assumption of an AR(1) process for the shocks to the unemployment rate equation.

The optimal values of the gain parameters depend on the variances of the four shocks. In our policy evaluation exercises, we consider alternative assumptions regarding the parameter values that the central bank uses in implementing the Kalman filters. In one case, we assume that the central bank uses the optimal values implied by the variances in our baseline calibration of the model. These values are as follows: $a_1 = 0.982$, $a_2 = 0.008$, $b_1 = 0.987$, $b_2 = 0.006$, and $b_3 = -0.003$. As noted above, there is a great deal of uncertainty regarding the values of the gain parameters, and real-world estimates tend to be very imprecise. We therefore examine two cases in which the central bank uses incorrect gain parameters. In one, the central bank assumes that the natural rates are constant, so the gain parameters are zero. In the other, we assume that the central bank uses the appropriate gain parameters for our baseline model calibration, but in fact the standard deviations of the natural rate shocks are twice as large as in the baseline calibration.

3. Expectations and Simulation Methods

As noted above, we are interested in studying the performance of the optimal control monetary policy derived under a misspecified model of expectations formation. We assume that private agents and, in some cases, the central bank, form expectations using an estimated reduced-form forecasting model. Specifically, following Orphanides and Williams (2005), we posit that private agents engage in perpetual learning, that is, they reestimate their forecasting model using a constant-gain least squares algorithm that weights recent data more heavily than past data.[9] This approach to modeling learning allows for the possible presence of time variation in the economy, including the natural rates of interest and unemployment. It also implies that agents' estimates are always subject to sampling variation, in that the estimates do not eventually converge to fixed values.

We assume that private agents forecast inflation, the unemployment rate, and the short-term interest rate using an unrestricted vector autoregression model (VAR) containing three lags of these three variables and a constant. We further assume that private agents do not observe or estimate the natural rates of unemployment and interest directly in forming expectations. The effects of time variation in natural rates on forecasts are reflected in the forecasting VAR by the lags of the interest rate, inflation rate, and unemployment rate. First, variants of VARs are commonly used in real-world macroeconomic forecasting, making this a reasonable choice on realism grounds. Second, the rational expectations equilibrium of our model with known natural rates is very well approximated by a VAR of this form. As discussed in Orphanides and Williams (2008), this VAR forecasting model provides accurate forecasts in model simulations.

At the end of each period, agents update their estimates of their forecasting model using data through the current period. To fix notation, let $\mathbf{Y}_t$ denote the 1×3 vector consisting of the inflation rate, the unemployment rate, and the interest rate, each measured at time t: $\mathbf{Y}_t = (\pi_t, u_t, i_t)$. Let $\mathbf{X}_t$ be the 10×1 vector of regressors in the forecast model: $\mathbf{X}_t = (1, \pi_{t-1}, u_{t-1}, i_{t-1}, \ldots, \pi_{t-3}, u_{t-3}, i_{t-3})$. Let $\mathbf{c}_t$ be the 10×3 vector of coefficients of the forecasting model. Using data through period t, the coefficients of the forecasting model can be written in recursive form:

$$\mathbf{c}_t = \mathbf{c}_{t-1} + \kappa \mathbf{R}_t^{-1} \mathbf{X}_t (\mathbf{Y}_t - \mathbf{X}_t' \mathbf{c}_{t-1}), \tag{12}$$

$$\mathbf{R}_t = \mathbf{R}_{t-1} + \kappa (\mathbf{X}_t \mathbf{X}_t' - \mathbf{R}_{t-1}), \tag{13}$$

where κ is the gain. Agents construct the multi-period forecasts that appear in the inflation and unemployment equations in the model using the estimated VAR.

9. See Sargent (1999), Cogley and Sargent (2002), and Evans and Honkapohja (2001) for related treatments of learning.

For some specifications of the VAR, the matrix $\mathbf{R}_t$ may not be full rank. To circumvent this problem, in each period of the model simulations, we check the rank of $\mathbf{R}_t$. If it is less than full rank, we assume that agents apply a standard Ridge regression (Hoerl and Kennard, 1970), where $\mathbf{R}_t$ is replaced by $\mathbf{R}_t + 0.00001^* \mathbf{I}(10)$, where $\mathbf{I}(10)$ is a 10×10 identity matrix.

3.1 Calibrating the Learning Rate

A key parameter in the learning model is the private agent updating parameter, κ. Estimates of this parameter tend to be imprecise and sensitive to model specification, but they generally lie between 0.00 and 0.04.[10] We take 0.02 to be a reasonable benchmark value for κ, a value that implies that the mean age of the weighted sample is about the same as for standard least squares with a sample of twenty-five years. Given the uncertainty about this parameter, we report results for values of κ between 0.01 (equivalent in mean sample age to a sample of about fifty years) and 0.03 (equivalent in mean sample age to a sample of about sixteen years).

3.2 Simulation Methods

In the case of rational expectations with constant and known natural rates, we compute model unconditional moments numerically as described in Levin, Wieland, and Williams (1999). In the case of learning, we compute approximations of the unconditional moments using stochastic simulations of the model.

For the stochastic simulations, we initialize all model variables to their respective steady-state values, which we assume to be zero. The initial conditions of $\mathbf{C}$ and $\mathbf{R}$ are set to the steady-state values implied by the forecasting perceived law of motion (PLM) in the rational expectations equilibrium with known natural rates. Each period, innovations are generated from independent Gaussian distributions with variances reported above. The private agent's forecasting model is updated each period and a new set of forecasts computed, as are the central bank's natural rate estimates. We simulate the model for 44,000 periods and discard the first 4,000 periods to eliminate the effects of initial conditions. We compute the unconditional moments from the remaining 40,000 periods (10,000 years) of simulated data.

10. See Sheridan (2003), Orphanides and Williams (2005), Branch and Evans (2006), and Milani (2007).

Learning introduces nonlinear dynamics into the model that may cause the model to display explosive behavior in a simulation. In simulations where the model is beginning to display signs of explosive behavior, we follow Marcet and Sargent (1989) and stipulate modifications to the model that curtail the explosive behavior. One potential source of explosive behavior is that the forecasting model itself may become explosive. We take the view that in practice private forecasters reject explosive models. Therefore, in each period of the simulation, we compute the maximum root of the forecasting VAR (excluding the constants). If this root falls below the critical value of 1, the forecast model is updated as described above; if not, we assume that the forecast model is not updated and the matrices **C** and **R** are held at their respective values from the previous period.[11] This constraint is encountered relatively rarely with the policies analyzed in this paper.

This constraint on the forecasting model is insufficient to ensure that the model economy does not exhibit explosive behavior in all simulations. We therefore impose a second condition that eliminates explosive behavior. In particular, the inflation rate, the nominal interest rate, and the unemployment gap are not allowed to exceed (in absolute value) six times their respective unconditional standard deviations (computed under the assumption of rational expectations and known natural rates) from their respective steady-state values. This constraint on the model is invoked extremely rarely in the simulations.

4. Performance of the Optimal Control Policy

In this section, we examine the performance of the optimal control policy derived under the assumption of rational expectations and known natural rates to deviations from this reference model. We start by considering the case in which private agents learn and natural rates are known by the central bank. We then turn to the case of natural rate uncertainty.

4.1 Known Natural Rates

The OC policy, derived for $\lambda = 4$ and $\nu = 1$, is given by the following equation:

11. We chose this critical value so that the test would have a small effect on model simulation behavior while eliminating explosive behavior in the forecasting model.

$$
\begin{aligned}
i_t = {} & 1.13(i_{t-1} - \hat{r}^*_{t-1}) + 0.02(i_{t-2} - \hat{r}^*_{t-1}) - 0.26(i_{t-3} - \hat{r}^*_{t-1}) + 0.18\pi_{t-1} \\
& + 0.03\pi_{t-2} + 0.04\pi_{t-3} - 2.48(u_{t-1} - \hat{u}^*_{t-1}) + 2.03(u_{t-2} - \hat{u}^*_{t-1}) \\
& - 0.34(u_{t-3} - \hat{u}^*_{t-1}).
\end{aligned} \tag{14}
$$

The first line of table 1 reports the outcomes for the OC policy under rational expectations and known natural rates. These outcomes serve as a benchmark against which the results under imperfect knowledge can be compared. The OC policy is characterized by a high degree of policy inertia, as measured by the sum of the coefficients on the lagged interest rates of 0.89. The sum of the coefficients on lagged inflation equals 0.25 and that on the lagged differences between the unemployment rates equals –0.89. As discussed in Orphanides and Williams (2008), the optimal control policy is characterized by a muted interest rate response to deviations of inflation from target. Following a shock to inflation, the OC policy only gradually brings inflation back to target and thus restrains the magnitude of deviations of unemployment from its natural rate and that of changes in the interest rate.

Table 1. Performance of Alternative Monetary Policies under Rational Expectations and Known Natural Rates[a]

	Standard deviation			*Loss*
Policy	π	$u - u^*$	Δi	L
Optimal control	1.83	0.68	1.20	6.64
Levin, Wieland, and Williams (2003)	1.87	0.70	1.24	6.98
Orphanides and Williams (2008)	1.83	0.73	1.39	7.45

Source: Authors' calculations.
a. The policies are derived for $\lambda = 4$ and $\nu = 1$.

Macroeconomic performance under the OC policy deteriorates under private agent learning, with the magnitude in fluctuations in all three objective variables increasing in the updating rate, κ. The upper panel of table 2 reports the results when private agents learn assuming constant natural rates. These results are very similar to those reported in Orphanides and Williams (2008), where natural rates are assumed to be constant and known. Thus, the incorporation of known time-varying natural rates does not have notable additional implications for the design of optimal monetary policy under imperfect knowledge. With learning, agents are never certain of the structure of the economy or the behavior of the central bank. As discussed in Orphanides and Williams (2005), particularly large shocks or a "bad run" of one-sided shocks can be misinterpreted by agents as reflecting a monetary policy regime that places less weight on inflation stabilization or has a different long-run inflation target than is actually the case. This confusion adds persistent noise to the economy, which worsens macroeconomic performance relative to the rational expectations benchmark.

Table 2. Performance of OC Policy under Learning and Time-Varying Natural Rates[a]

	Standard deviation			*Loss*
κ	π	$u - u^*$	Δi	L
Known natural rates				
0.01	2.28	0.80	1.33	9.52
0.02	2.77	0.93	1.55	13.59
0.03	3.23	1.09	1.80	18.46
Natural rate estimates with optimal Kalman filters				
0.01	2.26	0.88	1.33	9.99
0.02	3.16	1.10	1.82	17.79
0.03	3.59	1.23	1.99	22.94
Natural rates assumed constant				
0.01	2.81	0.92	1.44	13.39
0.02	3.68	1.12	1.82	21.89
0.03	4.11	1.25	2.09	27.53

Source: Authors' calculations.
a. The policies are derived for $\lambda = 4$ and $\nu = 1$.

4.2 Estimated Natural Rates

We now analyze the performance of the OC policy designed assuming rational expectations and known natural rates when private agents learn and natural rates are not directly observable. The middle section of table 2 reports the results assuming that the central bank uses the optimal Kalman filters to estimate both natural rates. As noted above, this case assumes that the central bank has precise knowledge of the structure of the IS and Phillips curve equations, observes private expectations that appear in those equations, and knows the covariance matrix of the shocks (which is used in determining the coefficients of the Kalman filter).

If expectations are close to the rational expectations benchmark and the central bank efficiently estimates natural rates, then natural rate uncertainty by itself has little additional effect on macroeconomic performance under the OC policy. For example, in the case of $\kappa = 0.01$, the standard deviations of inflation and the first difference of interest rates are about the same whether natural rates are known or optimally estimated. Not surprisingly, the standard deviation of the difference between the unemployment rate and its natural rate is somewhat higher if natural rates are not directly observed, since in that case the central bank will sometimes aim for the "wrong" unemployment rate target. These errors do not spill over into increased variability of other variables, however.

If the learning rate is 0.02 or above, the interaction of natural rate misperceptions and learning leads to a much greater deterioration of macroeconomic performance. Natural rate misperceptions introduce serially correlated errors into monetary policy. When agents are learning, these policy errors interfere with the public's understanding of the monetary policy rule. As a result, the variability of all three target variables increases. If the central bank uses the incorrect gains in the Kalman filters, macroeconomic performance worsens even further. The effects of using the wrong Kalman gains are illustrated in the lower panel of table 1. In this example, the central bank incorrectly assumes Kalman gains of zero in estimating natural rates (that is, it assumes that the variances of the shocks to the natural rates are zero). The resulting outcomes under the OC policy are significantly worse for all three learning rates shown in the table. The deterioration in performance is primarily due to a rise in the variability of inflation. Evidently, the combination of private agent learning and policy mistakes associated with poor measurement of natural rates significantly worsens the anchoring of inflation expectations and the stabilization of inflation.

5. Robust Optimal Control Policies

The preceding analysis shows that the optimal control policy derived under rational expectations and known natural rates may not be robust to imperfect knowledge. We now consider an approach to deriving policies that take advantage of the optimal control methodology but are robust to imperfect knowledge. Specifically, following Orphanides and Williams (2008), we search for the "biased" central bank loss function for which the implied OC policy derived with rational expectations and known natural rates performs best under imperfect knowledge for the true social loss function. This approach applies existing methods of computing optimal policies under rational expectations and is therefore feasible in practice.

For a given value of κ and assumptions regarding natural rates and natural rate measurement, we search for the values of $\tilde{\lambda}$ and $\tilde{\nu}$ such that the OC policy derived using the loss,

$$\tilde{L} = \text{var}(\pi - \pi^*) + \tilde{\lambda}\,\text{var}(u - u^*) + \tilde{\nu}\,\text{var}(\Delta(i)),$$

minimizes the true social loss, which we assume to be given by the benchmark values of $\lambda = 4$ and $\nu = 1$.[12] We use a grid search to find the optimal weights (up to one decimal place) for the biased central bank loss and refer to the resulting policy as the robust optimal control (ROC) policy.

5.1 Known Natural Rates

With known natural rates, the optimal weights for the central bank loss on unemployment and interest rate variability are significantly smaller than the true weights in the social loss, and this downward bias is increasing in the learning rate κ. The results from this exercise are reported in the upper panel of table 3, which considers the same set of assumptions regarding natural rate measurement as in table 2. For comparison, the losses under the OC policy, denoted L^*, are reported in the final column of the table. The results with known natural rates are similar to that in Orphanides and Williams (2008), where natural rates are assumed to be constant. The presence of learning makes it optimal to assign the central bank a loss that places much greater relative weight on inflation stabilization than the true social loss—that is, to employ a conservative central banker, in the terminology of Rogoff (1985).[13] The ROC policies yield significantly lower losses than the OC policy.

12. This approach can be generalized to allow the inclusion of additional variables in the loss function. We leave this to future research.

13. Orphanides and Williams (2005), using a very simple theoretical model, similarly find that a central bank loss function biased toward stabilizing inflation (relative to output) is optimal when private agents learn.

Table 3. Optimal Weights in Central Bank Loss under Imperfect Knowledge[a]

	Weights		*Standard deviation*			*Loss*	
κ	$\tilde{\lambda}$	$\tilde{\upsilon}$	π	$u - u^*$	Δi	$\tilde{L}$	L^*
Known natural rates							
0.01	1.4	0.3	1.95	0.80	1.55	8.76	9.52
0.02	1.2	0.2	2.14	0.89	1.85	11.13	13.59
0.03	0.2	0.1	2.03	1.04	2.20	13.31	18.46
Natural rate estimates with optimal Kalman filters							
0.01	1.2	0.4	1.95	0.91	1.45	9.29	9.99
0.02	0.5	0.1	2.06	1.03	1.80	11.70	17.79
0.03	0.4	0.2	2.10	1.06	2.23	13.85	22.94
Natural rates assumed constant							
0.01	0.5	0.2	1.97	0.91	1.64	9.90	13.39
0.02	0.5	0.2	2.22	0.98	1.78	11.97	21.89
0.03	0.2	0.1	2.18	1.07	2.21	14.18	27.53

Source: Authors' calculations.

a. The policies are derived for $\lambda = 4$ and $\nu = 1$. $\tilde{L}$ denotes the loss under the OC policy under the optimized values of $\tilde{\lambda}$ and $\tilde{\upsilon}$, evaluated using $\lambda = 4$ and $\nu = 1$. L^* denotes the loss under the OC policy derived for $\lambda = 4$ and $\nu = 1$.

5.2 Estimated Natural Rates

With estimated natural rates, the optimal weights for the central bank loss on unemployment and interest rate variability are generally smaller than in the case of known natural rates. Thus, the combination of learning and natural rate mismeasurement strengthens the case for placing much greater relative weight on inflation stabilization than the true social loss. For example, in the case of $\kappa = 0.02$ and optimally estimated natural rates, the optimal central bank objective weights are about one-half as large as in the case of known natural rates. In that case, the ROC policy for $\kappa = 0.02$ and optimally estimated natural rates is given by the following equation:

$$\begin{aligned} i_t &= 1.11(i_{t-1} - \hat{r}^*_{t-1}) - 0.12(i_{t-2} - \hat{r}^*_{t-1}) - 0.15(i_{t-3} - \hat{r}^*_{t-1}) + 0.51\pi_{t-1} \\ &\quad +0.28\pi_{t-2} + 0.00\pi_{t-3} - 3.32(u_{t-1} - \hat{u}^*_{t-1}) + 2.40(u_{t-2} - \hat{u}^*_{t-1}) \\ &\quad -0.43(u_{t-3} - \hat{u}^*_{t-1}). \end{aligned} \tag{15}$$

This ROC policy is characterized by a much larger direct response to the inflation rate than the OC policy derived for the benchmark loss (and reported in equation 14), reflecting the greater relative weight on inflation stabilization for the biased central bank loss. The ROC policy responds somewhat more to lags of the difference between the unemployment rate and the perceived natural rate of interest, with a sum of coefficients of –1.35 (versus –0.89 in the OC policy). It also exhibits less intrinsic policy inertia, with the sum of the coefficients on the lagged interest rate of 0.84 (versus 0.89 in the OC policy), reflecting the much smaller weight on interest rate variability underlying the ROC policy.

When the central bank incorrectly assumes that natural rates are constant, the optimal weights for the central bank loss on unemployment and on interest rate variability are at most one-fifth as large as the true values. The differences in the losses under the OC and ROC policies are much larger than in the case of known natural rates. The central bank loss under imperfect knowledge tends to be relatively insensitive to small differences in the weights used in deriving the robust optimal policies. As a result, the precise choice of optimal weights is not crucial. What is crucial is that the weights on unemployment and the change in interest rates are small relative to the weight on inflation.

5.3 Greater Natural Rate Variability

Thus far, we have assumed a relatively low degree of natural rate variability. We now explore the implications of more variable natural rates, consistent with some estimates in the literature.[14] In the following discussion, we assume that the standard deviation of the natural rate innovations is twice that assumed in our benchmark calibration. The results for these experiments are reported in table 4. The final column of the table reports the loss, denoted L^*, under the standard OC policy derived assuming rational expectations and known natural rates with the benchmark calibration of innovation variances.

If the central bank is assumed to observe the true values of the natural rates, then the greater degree of natural rate variability does not significantly affect the optimal choices of weights in the objective function used to derive the ROC policy. Comparing the upper panels of tables 3 and 4 shows that the optimal values of $\tilde{\lambda}$ and $\tilde{\nu}$ are similar for the two calibrations of natural rate variability. The losses associated with the OC policy are much larger when natural rates are more variable. In contrast, the losses under the appropriate ROC policies are not that different in the two cases.

If, however, the central bank underestimates the degree of natural rate variability in estimating natural rates, the optimal values of $\tilde{\lambda}$ and $\tilde{\nu}$ are very small, implying that the central bank should focus almost entirely on inflation stabilization in deriving optimal control policies. The lower panel of the table reports outcomes for the case in which the central bank uses the Kalman filter gains appropriate for the benchmark calibration of natural rate variability, but in fact the natural rates are twice as variable (in terms of standard deviations). In this case, the OC policy performs very badly, and the benefits of following the ROC policy rather than the OC are dramatic.

6. Simple Rules

We now compare the performance of two alternative monetary policies that have been recommended in the literature for being robust to various forms of model uncertainty to the optimal control policies

14. The case of zero variability of natural rates is analyzed in Orphanides and Williams (2008).

Table 4. Optimal Weights in Central Bank Loss under Imperfect Knowledge and High Natural Rate Variability[a]

	Weights		*Standard deviation*			*Loss*	
κ	$\tilde{\lambda}$	$\tilde{\upsilon}$	π	$u - u^*$	Δi	$\tilde{L}$	L^*
Known natural rates							
0.01	1.1	0.3	1.96	0.85	1.53	9.06	10.26
0.02	0.9	0.2	2.16	0.94	1.81	11.49	15.38
0.03	0.4	0.1	2.13	1.05	2.22	13.89	20.93
Natural rate estimates with baseline Kalman filters							
0.01	0.1	0.2	2.14	1.12	1.66	12.31	24.18
0.02	0.1	0.1	2.18	1.18	2.11	14.74	36.34
0.03	0.1	0.1	2.32	1.25	2.27	16.81	53.36

Source: Authors' calculations.
a. $\tilde{L}$ denotes the loss under the OC policy under the optimized values of $\tilde{\lambda}$ and $\tilde{\upsilon}$, evaluated using $\lambda = 4$ and $\nu = 1$. L^* denotes the loss under the OC policy derived for $\lambda = 4$ and $\nu = 1$.

analyzed above. The first rule is a version of the forecast-based policy rule proposed by Levin, Wieland, and Williams (2003). According to this rule, the short-term interest rate is determined as follows:

$$i_t = i_{t-1} + \theta_\pi(\pi^e_{t+3} - \pi^*) + \theta_u(u_{t-1} - \hat{u}^*_{t-1}), \tag{16}$$

where π^e_{t+3} is the forecast of the four-quarter change in the price level and u^* is the natural rate of unemployment which we take to be constant and known. Because this policy rule characterizes policy in terms of the first difference of the interest rate, it does not rely on estimates of the natural rate of interest, as does the standard Taylor rule (1993). The second rule we consider is proposed by Orphanides and Williams (2007) for its robustness properties in the face of natural rate uncertainty:

$$i_t = i_{t-1} + \theta_\pi(\pi^e_{t+3} - \pi^*) + \theta_{\Delta u}(u_{t-1} - u_{t-2}). \tag{17}$$

A key feature of this policy is the absence of any measures of natural rates in the determination of policy.[15]

We choose the parameters of these simple rules to minimize the loss under rational expectations and constant natural rates using a hill-climbing routine.[16] The resulting optimized Levin-Wieland-Williams rule is given by

$$i_t = i_{t-1} + 1.05(\pi^e_{t+3} - \pi^*) - 1.39(u_{t-1} - \hat{u}^*_{t-1}). \tag{18}$$

The optimized Orphanides-Williams rule is given by

$$i_t = i_{t-1} + 1.74(\pi^e_{t+3} - \pi^*) - 1.19(u_{t-1} - u_{t-2}). \tag{19}$$

15. This policy rule is related to the elastic price standard proposed by Hall (1984), whereby the central bank aims to maintain a stipulated relationship between the forecast of the unemployment rate and the price level. It is also closely related to the first difference of a modified Taylor-type policy rule in which the forecast of the price level is substituted for the forecast of the inflation rate.

16. If we allow for time-varying natural rates that are known by all agents, the optimized parameters of the Levin-Wieland-Williams and Orphanides-Williams rules under rational expectations are nearly unchanged. The relative performance of the different policies is also unaffected.

In the following, we refer to these specific parameterizations of these two rules simply as the Levin-Wieland-Williams and Orphanides-Williams rules.[17]

The lower part of table 1 reports the outcomes for the Levin-Wieland-Williams rule and the Orphanides-Williams rule under rational expectations and known natural rates. Under rational expectations and known natural rates, the OC policy yields a modestly lower loss than the Levin-Wieland-Williams and Orphanides-Williams rules, which is consistent with the findings in Williams (2003) and Levin and Williams (2003) about the relative performance of simple rules for other models.

In contrast to the OC policy, the Levin-Wieland-Williams and Orphanides-Williams rules perform very well under imperfect knowledge. Table 5 compares the performance of these rules to that of the OC policy derived under the true central bank loss and the ROC policies. (Because the Orphanides-Williams rule does not respond to natural rate estimates, outcomes are invariant to the assumption regarding central bank natural rate estimation.) In all cases reported in the table, the Levin-Wieland-Williams rule performs as well as or better than the OC policy, with the performance advantage larger the higher the learning rate and the greater the degree of natural rate misperceptions. As discussed in detail in Orphanides and Williams (2008), the Levin-Wieland-Williams rule consistently brings inflation back to target quickly following a shock to inflation, and it contains the response of inflation to the unemployment shock. The Orphanides-Williams rule does even better than the Levin-Wieland-Williams rule at containing the inflation responses to shocks, but at the cost of greater variability in the difference between the unemployment rate and its natural rate and the change in the interest rate. Consequently, the Levin-Wieland-Williams rule performs somewhat better than the Orphanides-Williams rule in terms of the stipulated central bank loss for all the cases that we consider here.

The outcomes under the Levin-Wieland-Williams and Orphanides-Williams rules are generally similar to those under the ROC policies. The first column of table 5 reports the losses under the ROC policies (repeated from table 3). The Levin-Wieland-Williams rule does slightly worse than the ROC policy in the cases closest to the perfect knowledge benchmark (that is, a low κ and modest natural rate misperceptions) and performs better as the degree of model misspecification increases.

17. These are the same rules analyzed in Orphanides and Williams (2008).

Table 5. Performance of Simple Rules under Imperfect Knowledge[a]

	OC policy		*Levin-Wieland-Williams rule*				*Orphanides-Williams rule*			
	Loss		*Standard deviation*			*Loss*	*Standard deviation*			*Loss*
κ	$\tilde{L}$	L^*	π	$u - u^*$	Δi	L	π	$u - u^*$	Δi	L
Known natural rates										
0.01	8.76	9.52	1.94	0.86	1.39	8.68	1.93	0.98	1.60	10.16
0.02	11.13	13.59	2.00	0.99	1.57	10.39	1.99	1.08	1.78	11.84
0.03	13.31	18.46	2.09	1.10	1.80	12.43	2.09	1.18	2.03	14.03
Natural rate estimates with optimal Kalman filters										
0.01	9.29	9.99	1.94	0.93	1.41	9.17	1.93	0.98	1.60	10.16
0.02	11.70	17.79	2.00	1.03	1.57	10.65	1.99	1.08	1.78	11.84
0.03	13.85	22.94	2.07	1.12	1.78	12.47	2.09	1.18	2.03	14.03
Natural rates assumed constant										
0.01	9.90	13.39	2.03	0.94	1.44	9.72	1.93	0.98	1.60	10.16
0.02	11.97	21.89	2.08	1.01	1.60	10.98	1.99	1.08	1.78	11.84
0.03	14.18	27.53	2.17	1.12	1.86	13.18	2.09	1.18	2.03	14.03

Source: Authors' calculations.

a. The policies are derived for $\lambda = 4$ and $\nu = 1$. $\tilde{L}$ denotes the loss under the OC policy under the optimized values of $\tilde{\lambda}$ and $\tilde{\upsilon}$, evaluated using $\lambda = 4$ and $\nu = 1$. L^* denotes the loss under the OC policy derived for $\lambda = 4$ and $\nu = 1$.

The Orphanides-Williams policy performs about the same or slightly worse than the ROC policies, except in the case of known natural rates, when the ROC policy performs much better. Evidently, the extra fine-tuning in the ROC policy compared to the simple rules is of little value in an environment characterized by learning and natural rate misperceptions. The results are qualitatively similar with greater natural rate variability, as seen in table 6. In this case, however, if the central bank uses the Kalman gains based on the benchmark calibration, the Orphanides-Williams rule outperforms the ROC policies.

The strong performance of the Levin-Wieland-Williams and Orphanides-Williams rules in the presence of natural rate mismeasurement reflects the fact that these rules do not rely on natural rate estimates as much as the OC policy. Indeed, the Orphanides-Williams rule does not respond to natural rates at all, while the Levin-Wieland-Williams rule responds only to estimates of the natural rate of unemployment. Importantly, these rules respond aggressively to movements in inflation. In the case of the Levin-Wieland-Williams rule, policy errors stemming from misperceptions of the natural rate of unemployment cause some deterioration in macroeconomic performance, but the consequences of these errors are limited by the countervailing effect of the strong response to resulting deviations of inflation from target.

7. Conclusion

Current methods of deriving optimal control policies ignore important sources of model uncertainty. This paper has examined the robustness of optimal control policies to uncertainty regarding the formation of expectations and natural rates and analyzed monetary policy strategies designed to be robust to these sources of imperfect knowledge. Our analysis shows that standard approaches to optimal policy yield policies that are not robust to imperfect knowledge. More positively, this analysis helps us identify and highlight key features of policies that are robust to these sources of model uncertainty.

The main finding is that a reorientation of policy toward stabilizing inflation relative to economic activity and interest rates is crucial for good economic performance in the presence of imperfect knowledge. Indeed, focusing on price stability in this manner is the policy that should be pursued even when the central bank cares greatly about stabilizing economic activity and interest rates. Although following

Table 6. Performance of Simple Rules under Imperfect Knowledge and High Natural Rate Variability[a]

	OC policy		*Levin-Wieland-Williams rule*				*Orphanides-Williams rule*			
	Loss		*Standard deviation*			*Loss*	*Standard deviation*			*Loss*
κ	$\tilde{L}$	L^*	π	$u - u^*$	Δi	L	π	$u - u^*$	Δi	L
Known natural rates										
0.01	9.06	10.26	1.94	0.90	1.42	9.00	1.99	1.17	1.65	12.17
0.02	11.49	15.38	2.00	1.03	1.57	10.73	2.06	1.28	1.86	14.23
0.03	13.89	20.93	2.09	1.18	1.81	13.23	2.13	1.36	2.07	16.19
Natural rate estimates with baseline Kalman filters										
0.01	12.31	24.18	2.09	1.12	1.48	11.57	1.99	1.17	1.65	12.17
0.02	14.74	36.34	2.14	1.20	1.64	12.99	2.06	1.28	1.86	14.23
0.03	16.81	53.36	2.24	1.33	1.91	15.70	2.13	1.36	2.07	16.19

Source: Authors' calculations.

a. $\tilde{L}$ denotes the loss under the OC policy under the optimized values of $\tilde{\lambda}$ and $\tilde{\upsilon}$, evaluated using $\lambda = 4$ and $\nu = 1$. L^* denotes the loss under the OC policy derived for $\lambda = 4$ and $\nu = 1$.

policies that place greater weight on economic stability may appear desirable in an environment of perfect knowledge, doing so is counterproductive and leads to greater instability when knowledge is imperfect. Moreover, in an environment of imperfect knowledge, well-designed robust simple rules perform about as well as optimal control policies designed to be robust to imperfect knowledge. This raises further doubts about the wisdom of relying on the optimal control approach in lieu of simple rules for policy design. Given the many other sources of model uncertainty, further research should be directed at analyzing robust monetary policy with a full array of sources of model uncertainty.

References

Branch, W.A. and G.W. Evans. 2006. "A Simple Recursive Forecasting Model." *Economics Letters* 91(2): 158–66.

Brock, W.A., S. Durlauf and K. West. 2007. "Model Uncertainty and Policy Evaluation: Some Theory and Empirics." *Journal of Econometrics* 127(2): 629–64.

Clark, T. and S. Kozicki. 2005. "Estimating Equilibrium Interest Rates in Real Time." *North American Journal of Economics and Finance* 16(3): 395–413.

Cogley, T. and T. Sargent. 2002. "Evolving Post-World War II U.S. Inflation Dynamics." In *NBER Macroeconomics Annual 2001*, edited by B. S. Bernanke and K. Rogoff, 331–73. MIT Press.

Cukierman, A. and F. Lippi. 2005. "Endogenous Monetary Policy with Unobserved Potential Output." *Journal of Economic Dynamics and Control* 29(11): 1951–83.

Evans, G.W., and S. Honkapohja. 2001. *Learning and Expectations in Macroeconomics*. Princeton University Press.

Friedman, M. 1947. "Lerner on the Economics of Control." *Journal of Political Economy* 55(5): 405–16.

————. 1948. "A Monetary and Fiscal Framework for Economic Stability." *American Economic Review* 38(3): 245–64.

Gaspar, V., F. Smets and D. Vestin. 2006. "Adaptive Learning, Persistence, and Optimal Monetary Policy." Working paper 644. Frankfurt: European Central Bank.

Giannoni, M.P. and M. Woodford. 2005. "Optimal Inflation Targeting Rules." In *The Inflation Targeting Debate*, edited by B. S. Bernanke and M. Woodford, 93–162. University of Chicago Press.

Hall, R.E. 1984. "Monetary Strategy with an Elastic Price Standard." *Proceedings of the Symposium on Price Stability and Public Policy,* 137–59. Federal Reserve Bank of Kansas City.

Hansen, L.P. and T.J. Sargent. 2007. *Robustness*. Princeton University Press.

Hoerl, A.E. and R.W. Kennard. 1970. "Ridge Regression: Biased Estimation of Nonorthogonal Problems." *Tecnometrics* 12(1): 69–82.

Laubach, T. 2001. "Measuring the NAIRU: Evidence from Seven Economies." *Review of Economics and Statistics* 83(2): 218–31.

Laubach, T. and J.C. Williams. 2003. "Measuring the Natural Rate of Interest." *Review of Economics and Statistics* 85(4): 1063–70.

Levin, A.T., V. Wieland and J.C. Williams. 1999. "Robustness of Simple Monetary Policy Rules under Model Uncertainty." In *Monetary Policy Rules*, edited by J. B. Taylor, 263–99. University of Chicago Press.

————. 2003. "The Performance of Forecast-Based Policy Rules under Model Uncertainty." *American Economic Review* 93(3): 622–45.

Levin, A.T. and J.C. Williams. 2003. "Robust Monetary Policy with Competing Reference Models." *Journal of Monetary Economics* 50(5): 945–75.

Marcet, A. and T.J. Sargent. 1989. "Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models." *Journal of Economic Theory* 48(2): 337–68.

McCallum, B.T. 1988. "Robustness Properties of a Rule for Monetary Policy." *Carnegie-Rochester Conference Series on Public Policy* 29: 173–203.

Milani, F. 2007. "Expectations, Learning, and Macroeconomic Persistence." *Journal of Monetary Economics* 54(7): 2065–82.

Orphanides, A., R. Porter, D. Reifschneider, R. Tetlow and F. Finan. 2000. "Errors in the Measurement of the Output Gap and the Design of Monetary Policy." *Journal of Economics and Business* 52(1–2): 117–41.

Orphanides, A. and S. van Norden. 2002. "The Unreliability of Output Gap Estimates in Real Time." *Review of Economics and Statistics* 84(4): 569–83.

Orphanides, A. and J.C. Williams. 2002. "Robust Monetary Policy Rules with Unknown Natural Rates." *Brookings Papers on Economic Activity* 2: 63–118.

————. 2005. "Imperfect Knowledge, Inflation Expectations and Monetary Policy." In *The Inflation Targeting Debate*, edited by B. S. Bernanke and M. Woodford, 201–34. University of Chicago Press.

————. 2007. "Inflation Targeting Under Imperfect Knowledge." In *Monetary Policy under Inflation Targeting*, edited by F. Mishkin and K. Schmidt-Hebbel, 77–123. Santiago: Central Bank of Chile.

————. 2008. "Learning, Expectations Formation, and the Pitfalls of Optimal Control Monetary Policy." *Journal of Monetary Economics* 55(Supplement 1): S80-S96.

Rogoff, K. 1985. "The Optimal Degree of Commitment to an Intermediate Monetary Target." *Quarterly Journal of Economics* 100(4): 1169–90.

Sargent, T.J. 1999. *The Conquest of American Inflation*. Princeton University Press.

————. 2007. "Evolution and Intelligent Design." New York University.

Sheridan, N. 2003. "Forming Inflation Expectations." Johns Hopkins University.

Staiger, D., J.H. Stock, and M.W. Watson. 1997. "How Precise Are Estimates of the Natural Rate of Unemployment?" In *Reducing Inflation: Motivation and Strategy*, edited by C. D. Romer and D. H. Romer, 195–246. University of Chicago Press.

Svensson, L.E.O. and R. Tetlow. 2005. "Optimal Policy Projections." *International Journal of Central Banking* 1(3): 177–207.

Svensson, L.E.O. and N. Williams. 2007. "Monetary Policy with Model Uncertainty: Distribution Forecast Targeting." Discussion paper 6331. London: Centre for Economic Policy Research.

Svensson, L.E.O. and M. Woodford. 2003. "Indicator Variables for Optimal Policy." *Journal of Monetary Economics* 50(3): 691–720.

Taylor, J.B. 1993. "Discretion versus Policy Rules in Practice." *Carnegie-Rochester Conference Series on Public Policy* 39 (December): 195–214.

————. 1999. "The Robustness and Efficiency of Monetary Policy Rules as Guidelines for Interest Rate Setting by the European Central Bank." *Journal of Monetary Economics* 43(3): 655–79.

Tetlow, R.J. 2006. "Real-time Model Uncertainty in the United States: 'Robust' Policies Put to the Test." Washington: Board of Governors of the Federal Reserve System.

Williams, J.C. 2003. "Simple Rules for Monetary Policy." *Federal Reserve Bank of San Francisco Economic Review* 1–12.

Woodford, M. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Robust Learning Stability with Operational Monetary Policy Rules

George W. Evans
University of Oregon and University of St Andrews

Seppo Honkapohja
Bank of Finland and University of Cambridge

The recent literature examines the conduct of monetary policy in terms of interest rate rules from the viewpoint of imperfect knowledge and learning by economic agents. The stability of the rational expectations equilibrium is taken as a key desideratum for good monetary policy design.[1] Most of this literature postulates that agents use least squares or related learning algorithms to carry out real-time estimations of the parameters of their forecast functions as new data become available. Moreover, it is usually assumed that the learning algorithms have a decreasing gain; in the most common case, the gain is the inverse of the sample size so that all data points have equal weights. Use of such a decreasing-gain algorithm makes it possible for learning to converge exactly to the rational expectations equilibrium in environments without structural change. Convergence requires that the equilibrium satisfies a stability condition, known as E-stability.

Decreasing-gain algorithms do not perform well, however, when occasional unobservable structural changes take place. So-called constant-gain algorithms are a natural alternative for estimating parameters in a way that is alert to possible structural changes. If agents use a constant-gain algorithm, then parameter estimates of the forecast functions do not fully converge to the rational expectations

We thank our discussant Claes Berg, Carl E. Walsh, and the conference participants for valuable comments. Support from National Science Foundation Grant No. SES-0617859 and ESRC grant RES-000-23-1152 is gratefully acknowledged.

1. For surveys, see Evans and Honkapohja (2003a), Bullard (2006), and Evans and Honkapohja (in this volume).

Monetary Policy under Uncertainty and Learning, edited by Klaus Schmidt-Hebbel and Carl E. Walsh, Santiago, Chile.

equilibrium values. Instead, they remain random, even asymptotically. For small values of the gain parameter, the estimates mostly remain in a small neighborhood of the rational expectations equilibrium, provided that the equilibrium is E-stable.[2] Constant-gain algorithms have recently been employed in empirical work, such as Milani (2005, 2007a), Orphanides and Williams (2005a, 2005b), and Branch and Evans (2006).

The connection between convergence of constant-gain learning and E-stability noted above is a limiting result for sufficiently small gain parameters. For finite values of the gain parameter, the stability condition for constant-gain learning is more stringent than E-stability. In this paper we examine the stability implications of various interest rate rules when agents use constant-gain learning rules with plausible positive values of the gain. We say that an interest rate rule yields robust learning stability of the economy if stability under constant-gain learning obtains for all values of the gain parameter in the range suggested by the empirical literature.[3] We focus on interest rate rules that are operational in the sense discussed by McCallum (1999), who holds that monetary policy cannot be conditioned on current values of endogenous aggregate variables. The rules we consider therefore assume that policy responds to expectations of contemporaneous (or future) values of inflation and output, but not on their actual values in the current period.

We consider robust learning stability for a variety of operational interest rate rules that have been suggested in the recent literature. These include Taylor rules and optimal reaction functions under discretion and commitment when central bank policy aims for interest rate stabilization in addition to the usual motives for flexible inflation targeting. The reaction function may be expectations-based in the spirit of Evans and Honkapohja (2003b, 2006) or of the Taylor-type form suggested by Duffy and Xiao (2007). We also analyze two interest

2. See Evans and Honkapohja (2001, chaps. 3 and 7) for the basic theoretical results on constant-gain learning. See also Evans, Honkapohja, and Williams (forthcoming) for references on recent papers on constant-gain learning. The possibility of divergence resulting from constant gain learning was noted in Slobodyan, Bogomolov, and Kolyuzhnov (2006).

3. Numerous concepts of robustness are relevant to policymaking, reflecting, for example, uncertainty about the structure of the economy and a desire by both private agents and policymakers to guard against the risk of large losses. We do not mean to downplay the importance of such factors, but we abstract from them here to focus on the importance of setting policy in such a way as to ensure stability in the face of constant-gain learning.

rate rules that approximate optimal policy under commitment, as suggested by Svensson and Woodford (2005) and McCallum and Nelson (2004). Our results show that expectations-based rules deliver robust learning stability, whereas the proposed alternatives often become unstable under learning even at quite small values of the constant-gain parameter.

1. Constant-Gain Steady-State Learning

In this paper we employ multivariate linear models. In this simplest case, in which the shocks are white noise and there are no lagged endogenous variables, the rational expectations equilibrium takes the form of a stochastic steady state. We now briefly review the basics of steady-state learning in linear models and then apply the results to Taylor rules.[4]

1.1 Theoretical Results

The steady state can be computed by postulating that agents' beliefs, called the perceived law of motion (PLM), take the form

$$\mathbf{y}_t = \mathbf{a} + \mathbf{e}_t$$

for a vector $\mathbf{y}_t$, where $\mathbf{e}_t \sim$ i.i.d.$(0, \sigma^2)$. Using the model, one then computes the actual law of motion (ALM), which describes the temporary equilibrium in the current period, given the PLM. We write the ALM using a linear operator $\mathbf{T}$ as

$$\mathbf{y}_t = \alpha + \mathbf{Ta} + \mathbf{e}_t,$$

where the matrix $\mathbf{T}$ depends on the structural parameters of the model. Examples of the $\mathbf{T}$ map are provided below. A rational expectations equilibrium is a fixed point, $\bar{\mathbf{a}}$, of the $\mathbf{T}$ map, that is,

$$\bar{\mathbf{a}} = \alpha + \mathbf{T}\bar{\mathbf{a}}.$$

We assume that $\mathbf{I} - \mathbf{T}$ is nonsingular, so that there is a unique solution $\bar{\mathbf{a}} = (\mathbf{I} - \mathbf{T})^{-1}\alpha$. For convenience, and without loss of generality, we

4. See Evans and Honkapohja (2001, chaps. 8 and 10) for a detailed discussion of adaptive learning in linear models.

now assume that the model has been written in deviation-from-the-mean form, so that $\boldsymbol{\alpha} = \mathbf{0}$. Thus the rational expectations equilibrium corresponds to $\bar{\mathbf{a}} = \mathbf{0}$ in our analysis. Under learning, agents attempt to learn the value of $\bar{\mathbf{a}}$, and hence in deviation-from-the-mean form we are examining whether agents' estimates of the mean converge to $\mathbf{a} = \mathbf{0}$.

Steady-state learning under decreasing gain is given by the recursive algorithm,

$$\mathbf{a}_t = \mathbf{a}_{t-1} + \gamma_t(\mathbf{y}_t - \mathbf{a}_{t-1}), \tag{1}$$

where the gain γ_t is a sequence of small decreasing numbers, such as $\gamma_t = 1/t$. Assuming that $\mathbf{y}_t = \mathbf{T}\mathbf{a}_{t-1} + \mathbf{e}_t$, that is, that expectations are formed using the estimate $\mathbf{a}_{t-1}$ based on data through time $t-1$, the convergence condition of algorithm (1) is given by the conditions for local asymptotic stability of $\bar{\mathbf{a}}$ under an associated differential equation:

$$\frac{d\mathbf{a}}{d\tau} = \mathbf{T}\mathbf{a} - \mathbf{a},$$

which is known as the E-stability differential equation. Here τ denotes notional or virtual time. The E-stability condition holds if and only if all eigenvalues of the matrix $\mathbf{T}$ have real parts less than one.[5]

Under constant-gain learning, the estimate $\mathbf{a}_t$ of $\mathbf{a}$ is updated according to

$$\mathbf{a}_t = \mathbf{a}_{t-1} + \gamma(\mathbf{y}_t - \mathbf{a}_{t-1}), \tag{2}$$

where $0 < \gamma \leq 1$ is the constant-gain parameter. The only difference between equation (2) and equation (1) is the constancy of the gain sequence. We now have

$$\mathbf{a}_t = \mathbf{a}_{t-1} + \gamma(\mathbf{T}\mathbf{a}_{t-1} + \mathbf{e}_t - \mathbf{a}_{t-1}),$$

or

$$\mathbf{a}_t = [\gamma\mathbf{T} + (1-\gamma)\mathbf{I}]\mathbf{a}_{t-1} + \gamma\mathbf{e}_t.$$

5. Throughout, we rule out boundary cases in which the real part of some eigenvalue of the $\mathbf{T}$ map is one.

This converges to a stationary stochastic process around the rational expectations equilibrium value (in deviation-from-the-mean form) provided all roots of the matrix $\gamma\mathbf{T} + (1 - \gamma)\mathbf{I}$ lie inside the unit circle.

Stability under constant-gain learning depends on the value of γ, and we have the following result.

Proposition 1. For a given $0 < \gamma \leq 1$, the stability condition is that the eigenvalues of $\mathbf{T}$ lie inside a circle of radius $1/\gamma$ and origin at $(1 - 1/\gamma, 0)$. This condition is therefore stricter for larger values of γ.

Proof. The stability condition is that the roots of $\gamma[\mathbf{T} + \gamma^{-1}(1 - \gamma)\mathbf{I}]$ lie inside the unit circle centered at the origin. Equivalently, the roots of $[\mathbf{T} + \gamma^{-1}(1 - \gamma)\mathbf{I}]$ must lie inside a circle of radius $1/\gamma$ centered at the origin. Since the roots of $\mathbf{T} + \gamma^{-1}(1 - \gamma)\mathbf{I}$ are the same as the roots of $\mathbf{T}$ plus $\gamma^{-1}(1 - \gamma)$, this is equivalent to the condition given.

The right edge of the circle is at (1, 0) in the complex plane, and as $\gamma \to 0$ we obtain the standard (decreasing-gain) E-stability condition that the real parts of all roots of $\mathbf{T}$ are less than one. Looking at the other extreme, $\gamma = 1$, gives the following corollary of proposition 1:

Proposition 2. We have stability for all $0 < \gamma \leq 1$ if and only if all eigenvalues of $\mathbf{T}$ lie inside the unit circle.

Stability for all constant gains, $0 < \gamma \leq 1$, is equivalent to a condition known as iterative E-stability, sometimes called IE-stability. Iterative E-stability is said to hold when $\mathbf{T}^j \to \mathbf{0}$ as $j \to \infty$.[6]

When the stability condition holds, the parameter $\mathbf{a}_t$ converges to a stationary stochastic process that we can fully describe. This, in turn, induces a stationary stochastic process for $\mathbf{y}_t = \mathbf{T}\mathbf{a}_{t-1} + \mathbf{e}_t$.

1.2 Application to Taylor Rules

Consider the standard forward-looking New-Keynesian model,

$$x_t = -\varphi(i_t - \pi^e_{t+1}) + x^e_{t+1} + g_t; \tag{3}$$

$$\pi_t = \lambda x_t + \beta\pi^e_{t+1} + u_t. \tag{4}$$

For convenience we assume that $(g_t, u_t)'$ are independent and identically distributed (i.i.d.), so that the preceding technical results

6. In many models, iterative E-stability is known to be a necessary condition for the stability of eductive learning; see, for example, Evans and Guesnerie (1993).

can be applied. Later we consider cases with first-order autoregressive, or AR(1), shocks. We use x^e_{t+1} and π^e_{t+1} to denote expectations of π_{t+1} and x_{t+1}. Below we specify the information sets available to agents when they are forming expectations, and throughout the paper we explore the implications of alternative assumptions.

Bullard and Mitra (2002) consider Taylor rules of various forms, including the contemporaneous data rule,

$$i_t = \chi_\pi \pi_t + \chi_x x_t, \tag{5}$$

and the "contemporaneous expectations" rule,

$$i_t = \chi_\pi \pi^e_t + \chi_x x^e_t. \tag{6}$$

In this section, our analysis of the contemporaneous expectations rule follows Bullard and Mitra (2002) in assuming that all expectations are based on information at time $t-1$, that is, $\pi^e_t = \hat{E}_{t-1}\pi_t$, $x^e_t = \hat{E}_{t-1}x_t$, $\pi^e_{t+1} = \hat{E}_{t-1}\pi_{t+1}$, and $x^e_{t+1} = \hat{E}_{t-1}x_{t+1}$. Since we have i.i.d. shocks, forecasts are based purely on the estimated intercept.

Bullard and Mitra (2002) show that the determinacy and E-stability conditions are the same and are identical for both interest rate rules. They are given by

$$\lambda(\chi_\pi - 1) + (1-\beta)\chi_x > 0. \tag{7}$$

Bullard and Mitra consider this finding important because of McCallum's (1999) argument that interest rate rules cannot plausibly be conditioned on contemporaneous observations of endogenous aggregate variables like inflation and output, whereas they could plausibly be conditioned on central bank forecasts or "nowcasts" $\hat{E}_{t-1}\pi_t$, $\hat{E}_{t-1}x_t$.

We reconsider this issue from the vantage point of constant-gain learning. For the interest rate rule (6), the model takes the form

$$\mathbf{y}_t = \mathbf{M}_0\mathbf{y}^e_t + \mathbf{M}_1\mathbf{y}^e_{t+1} + \mathbf{P}\mathbf{v}_t, \tag{8}$$

where $\mathbf{y}_t' = (x_t, \pi_t)$ and $\mathbf{v}_t' = (g_t, u_t)$ and where

$$\mathbf{M}_0 = \begin{pmatrix} -\chi_x\varphi & -\chi_\pi\varphi \\ -\chi_x\varphi\lambda & -\chi_\pi\varphi\lambda \end{pmatrix} \text{ and } \mathbf{M}_1 = \begin{pmatrix} 1 & \varphi \\ \lambda & \beta+\varphi\lambda \end{pmatrix}, \tag{9}$$

and

$$\mathbf{P} = \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix}.$$

Since our shocks are i.i.d., the PLM is simply $\mathbf{y}_t = \mathbf{a} + \mathbf{e}_t$, and the corresponding ALM is $\mathbf{y}_t = (\mathbf{M}_0 + \mathbf{M}_1)\mathbf{a} + \mathbf{e}_t$, where $\mathbf{e}_t = \mathbf{P}\mathbf{v}_t$. The usual E-stability condition is that the eigenvalues of $\mathbf{M}_0 + \mathbf{M}_1$ have real parts less than one, which leads to condition (7). According to proposition 2, for convergence of constant-gain learning for all gains $0 < \gamma \leq 1$, both eigenvalues of $\mathbf{M}_0 + \mathbf{M}_1$ must lie inside the unit circle.

We investigate the stability of constant-gain learning numerically, using the Woodford calibration of $\varphi^{-1} = 0.157$, $\lambda = 0.024$, $\beta = 0.99$. Setting $\chi_\pi = 1.5$, eigenvalues with real parts less than -1 arise for $\chi_x > 0.31$ and eigenvalues with real parts less than -9 arise for $\chi_x > 1.57$. This implies that when $\chi_\pi = 1.5$ and $\chi_x > 1.57$, the equilibrium is unstable under learning for constant gains $\gamma \geq 0.10$. This is perhaps not a significant practical concern since Taylor's recommended parameters are $\chi_\pi = 1.5$ and (based on the quarterly calibration of Woodford) $\chi_x = (0.5)/4 = 0.125$. However, it does show a previously unrecognized danger that arises under constant-gain learning if the Taylor rule has too strong a response to $\hat{E}_{t-1}x_t$, and this finding foreshadows instability problems that arise in more sophisticated rules discussed below.

Finally, the potential for instability under constant-gain learning arises specifically because of the need to use forecasts $\hat{E}_{t-1}\mathbf{y}_t$. For the current-data Taylor rule (5), it can be shown that condition (7) guarantees stability under learning for all constant gains $0 < \gamma \leq 1$.[7]

2. Optimal Discretionary Monetary Policy

We now consider optimal policy under constant-gain learning, starting with optimal discretionary policy. We focus on homogeneous learning by private agents and the policymaker. We initially restrict attention to the case of i.i.d. exogenous shocks, so that steady-state

7. The model now takes the form $\mathbf{y}_t = \mathbf{M}_1 \hat{E}_t \mathbf{y}_{t+1} + \mathbf{P}\mathbf{v}_t$, and the required condition is the same as the determinacy condition.

learning is appropriate. We also analyze the more general case, in which the observable shocks follow AR(1) processes.

Consider the loss function

$$E_0 \sum_{t=0}^{\infty} [(\pi_t - \pi^*)^2 + \alpha_x (x_t - x^*)^2 + \alpha_i (i_t - i^*)^2], \tag{10}$$

where π^*, x^*, and i^* represent target values. For simplicity, we set $\pi^* = x^* = 0$. The weights α_x, $\alpha_i > 0$ represent relative weights given by policymakers to squared deviations of x_t and i_t from their targets, compared with squared deviations of π_t from its target.

The first-order condition for discretionary optimal policy is

$$\lambda\pi_t + \alpha_x x_t - \alpha_i \varphi^{-1}(i_t - i^*) = 0. \tag{11}$$

We first consider a Taylor-Type Rule proposed by Duffy and Xiao (2007) and then discuss the expectations-based rule recommended by Evans and Honkapohja (2003b).

2.1 Taylor-Type Optimal Rules

Duffy and Xiao (2007) propose using the equation (11) directly to obtain a Taylor-Type Rule that implements optimal discretionary policy. Solving the first-order condition for i_t yields the rule

$$i_t = \frac{\varphi\lambda}{\alpha_i} \pi_t + \frac{\varphi\alpha_x}{\alpha_i} x_t,$$

where at this point we drop the term i^* since for brevity we are suppressing all intercepts. As Duffy and Xiao (2007) discuss, this is formally a contemporaneous-data Taylor rule. They show that for calibrated values of structural parameters and policy weights, this leads to a determinate and E-stable equilibrium.

The central bank's observing contemporaneous output and inflation is problematic. We therefore examine the rule

$$i_t = \frac{\varphi\lambda}{\alpha_i} \hat{E}_{t-1}\pi_t + \frac{\varphi\alpha_x}{\alpha_i} \hat{E}_{t-1}x_t, \tag{12}$$

where the information set for the nowcasts $\pi_t^e = \hat{E}_{t-1}\pi_t$, $x_t^e = \hat{E}_{t-1}x_t$ is past endogenous variables and exogenous variables.[8] This again leads to a model of the form (8) with coefficients (9), where $\chi_\pi = \varphi\lambda/\alpha_i$ and $\chi_x = \varphi\alpha_x/\alpha_i$. We assume that private agents and central banks estimate the same PLM. Since we are here assuming steady-state learning, we also have $\hat{E}_{t-1}\pi_{t+1} = \hat{E}_{t-1}\pi_t$ and $\hat{E}_{t-1}x_{t+1} = \hat{E}_{t-1}x_t$.

For a sufficiently large α_i, the model under this Taylor-Type Rule will suffer from indeterminacy. This follows from the Bullard-Mitra result that the determinacy condition is equation (7), from which the critical value of α_i can be deduced. The condition for determinacy is

$$\alpha_i < \bar{\alpha}_i \equiv \varphi\lambda + (1-\beta)\lambda^{-1}\varphi\alpha_x. \tag{13}$$

If the central bank's desire to stabilize the interest rate is too strong—that is, if condition (13) is not met—then the central bank fails to adjust the interest rate sufficiently to ensure that the generalized Taylor principle (7) is satisfied. To assess this point numerically, we use the calibrated parameter values of Woodford (2003, table 6.1), with $\alpha_x = 0.048$, $\varphi = 1/0.157$, $\lambda = 0.024$, and $\beta = 0.99$, which yields approximately $\bar{\alpha}_i = 0.28$. Woodford's calibrated values of α_i are 0.077 or 0.233 (the latter value is from Woodford, 1999). Thus the condition for determinacy does hold for these calibrations.

We next consider stability under learning. For the PLM $\mathbf{y}_t = \mathbf{a} + \mathbf{e}_t$, we again get the ALM $\mathbf{y}_t = (\mathbf{M}_0 + \mathbf{M}_1)\mathbf{a} + \mathbf{e}_t$ and

$$\mathbf{T} \equiv \mathbf{M}_0 + \mathbf{M}_1 = \begin{pmatrix} 1-\alpha_i^{-1}\alpha_x\varphi^2 & \varphi-\alpha_i^{-1}\lambda\varphi^2 \\ \lambda-\alpha_i^{-1}\lambda\alpha_x\varphi^2 & \beta+\lambda\varphi-\alpha_i^{-1}\lambda^2\varphi^2 \end{pmatrix}.$$

It can be shown that

$$\det(\mathbf{T}) = \beta(1-\alpha_i^{-1}\alpha_x\varphi^2).$$

Stability under all values $0 < \gamma \leq 1$ requires that

$$\left|\beta(1-\alpha_i^{-1}\alpha_x\varphi^2)\right| < 1,$$

8. An alternative would be to assume that agents and the policymaker see the contemporaneous value of the exogenous shocks but not the contemporaneous values of x_t and π_t. This would not alter our results.

and it is clear that for given β, α_x, φ this condition will not be satisfied for a sufficiently small $\alpha_i > 0$. This leads to our next proposition:

Proposition 3. Let $\hat{\alpha}_i = \beta(1+\beta)^{-1}\alpha_x\varphi^2$. For $0 < \alpha_i < \hat{\alpha}_i$, there exists $0 < \hat{\gamma}(\beta, \varphi, \alpha_i, \alpha_x) < 1$ such that the optimal discretionary Taylor-Type Rule (12) renders the rational expectations equilibrium unstable under learning for $\hat{\gamma} < \gamma \leq 1$.

Thus, in addition to the indeterminacy problem for large values of α_i, the Taylor-type optimal rule suffers from a more serious problem of instability under constant-gain learning for small values of α_i. The source of this difficulty is the interaction of strong policy responses seen in equation (12) and a large gain parameter. This combination leads to cyclical overshooting of inflation and the output gap. This is particularly evident as α_i tends to zero, since in this case, a positive change in inflation expectations $\hat{E}_{t-1}\pi_t$ leads to a large increase in i_t, which in turn leads to large negative changes in x_t and π_t via equations (3) and (4). The severity of this problem depends on the value of $\hat{\gamma}$ in proposition 3. Ideally, stability would hold for all $0 < \gamma \leq 1$, but the problem might not be a major concern if $\hat{\gamma}$ is high.

We investigate the magnitude of $\hat{\gamma}$ numerically by computing the eigenvalues of $\gamma\mathbf{T} + (1 - \gamma)\mathbf{I}$. As an example, for the Woodford calibration $\beta = 0.99$, $\varphi = 1/0.157$, and $\lambda = 0.024$, we find that with $\alpha_x = 0.048$ and $\alpha_i = 0.077$, the critical value $\hat{\gamma} \approx 0.04$. Since estimates in the macroeconomic literature suggest gains in the range 0.02 to 0.06, this indicates that optimal Taylor-Type Rules may not be stable under learning.[9] The source of the problem is that with low α_i the implied weights on $\hat{E}_{t-1}\pi_t$ and especially $\hat{E}_{t-1}x_t$ are very high. Under constant-gain learning, this can lead to instability unless the gain parameter is very low. As we demonstrate later, this problem can be avoided by using a suitable expectations-based optimal rule.

We next consider the case in which the exogenous shocks are AR(1) processes. The literature uses various information assumptions in this setting. Perhaps the most common assumption is that agents see current and lagged exogenous variables and lagged, but not current, endogenous variables. Expectations under this assumption are denoted $\hat{E}_t\pi_t$, $\hat{E}_t x_t$, $\hat{E}_t\pi_{t+1}$, $\hat{E}_t x_{t+1}$. An alternative would be to replace these with $\hat{E}_{t-1}\pi_t$, $\hat{E}_{t-1}x_t$, $\hat{E}_{t-1}\pi_{t+1}$, $\hat{E}_{t-1}x_{t+1}$, indicating that

9. Milani (2007b) considers a setting in which agents switch between decreasing-gain and constant-gain estimators, depending on recent average mean-square errors. The estimated gains are even higher in the constant-gain regime, at around 0.07 to 0.08.

agents only see lagged information.[10] Whether agents see current or only lagged exogenous shocks is not particularly crucial and does not affect our main results. We therefore follow the most common assumption that expectations are specified as $\hat{E}_t\pi_t$, $\hat{E}_t x_t$, $\hat{E}_t\pi_{t+1}$, and $\hat{E}_t x_{t+1}$.[11] In contrast, whether agents and policymakers are able to see current endogenous variables is an important issue for stability under learning, as we have already seen. This is why we use the term operationality to indicate an interest rate rule that does not depend on current endogenous variables.

We now assume that the exogenous shocks g_t and u_t follow AR(1) processes, that is,

$$g_t = \mu g_{t-1} + \tilde{g}_t$$

and

$$u_t = \rho u_{t-1} + \tilde{u}_t,$$

where $0 < |\mu|,\ |\rho| < 1$, and $\tilde{g}_t \sim \text{i.i.d.}(0,\sigma_g^2)$, $\tilde{u}_t \sim \text{i.i.d.}(0,\sigma_u^2)$ are independent white noise processes. We write this in vector form as

$$\mathbf{v}_t = \mathbf{F}\mathbf{v}_t + \tilde{\mathbf{v}}_t.$$

Under the current assumptions, the PLM of the agents is

$$\mathbf{y}_t = \mathbf{a} + \mathbf{c}\mathbf{v}_t,$$

and the forecasts are now $\hat{E}_t\mathbf{y}_t = \mathbf{a} + \mathbf{c}\mathbf{v}_t$ and $\hat{E}_t\mathbf{y}_{t+1} = \mathbf{a} + \mathbf{c}\mathbf{F}\mathbf{v}_t$. Using the general model (8), the ALM is

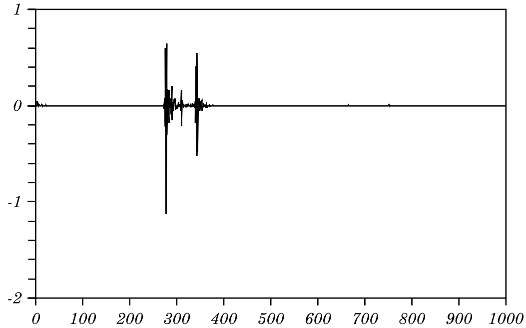
$$\mathbf{y}_t = (\mathbf{M}_0 + \mathbf{M}_1)\mathbf{a} + (\mathbf{M}_0\mathbf{c} + \mathbf{M}_1\mathbf{c}\mathbf{F} + \mathbf{P})\mathbf{v}_t,$$

10. A third alternative, which is occasionally used in the literature, allows agents to see the contemporaneous values of endogenous variables. However, this assumption runs against the requirement of operationality that we want to emphasize here.

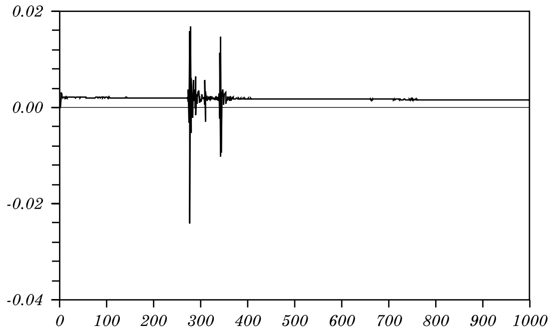
11. The standard assumption under rational expectations is that agents have contemporaneous information. Our information assumption takes account of the operationality critique, but nonetheless allows for the possibility of convergence under learning to the rational expectations equilibrium.

Figure 1. Stability of Optimal Taylor-Type Rule with γ = 0.02.

A. Deviation of x from Rational Expectation



B. Deviation of π from Rational Expectation



Source: Authors' calculations.

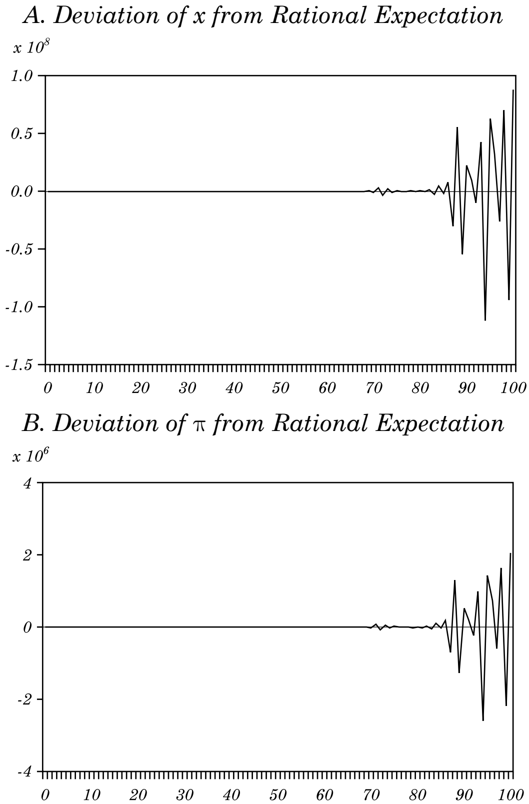
and the E-stability conditions are that all eigenvalues of the matrices $\mathbf{M}_0 + \mathbf{M}_1$ and $\mathbf{I} \otimes \mathbf{M}_0 + \mathbf{F}' \otimes \mathbf{M}_1$ have real parts less than one. Here, $\otimes$ denotes the Kronecker product of two matrices.[12]

To examine stability under constant-gain learning, we simulate the model under constant-gain recursive least squares (RLS) estimation of the PLM parameters **a** and **c**.[13] Under constant-gain RLS, agents discount old data geometrically at the rate $1 - \gamma$. Let $\mathbf{a}_t$, $\mathbf{c}_t$ denote the estimates based on data through $t - 1$. Given these estimates, expectations are formed as $\mathbf{y}_t^e = \hat{E}_t \mathbf{y}_t = \mathbf{a}_t + \mathbf{c}_t \mathbf{v}_t$ and

12. In the case of lagged information, the PLM is specified as $\mathbf{y}_t = \mathbf{a} + \mathbf{c}\mathbf{v}_{t-1} + \eta_t$, and the ALM is then $\mathbf{y}_t = (\mathbf{M}_0 + \mathbf{M}_1)\mathbf{a} + (\mathbf{M}_0\mathbf{c} + \mathbf{M}_1\mathbf{c}\mathbf{F} + \mathbf{P}\mathbf{F})\mathbf{v}_{t-1} + \tilde{\mathbf{v}}_t$.

13. See the appendix for the recursive formulation of constant-gain least squares.

Figure 2. Instability of Optimal Taylor-Type Rule with $\gamma = 0.04$.



Source: Authors' calculations.

$\mathbf{y}_{t+1}^{e} = \hat{E}_t \mathbf{y}_{t+1} = \mathbf{a}_t + \mathbf{c}_t \mathbf{F} \mathbf{v}_t$, and the temporary equilibrium is then given by equation (8) with these expectations.

We use the previous values for the structural parameters and also set $\mu = \rho = 0.8$. Simulations of the system indicate instability under constant-gain RLS learning for gain parameters at or in excess of 0.024. Thus, with regressors that include exogenous AR(1) observables, instability arises at even lower gain values than in the case of steady-state learning. Figures 1 and 2 illustrate the evolution of parameters over time under constant-gain RLS learning with the Taylor-Type Rule (12) in stable and unstable cases.[14]

14. In the stable case, the small deviation of π from rational expectations, seen in figure 1, gradually vanishes as the simulation length increases.

2.2 Expectations-Based Optimal Rules

Assume now that at time t the exogenous shocks g_t, u_t and private-sector expectations $\hat{E}_t\pi_{t+1}$, $\hat{E}_t x_{t+1}$ are observed by the central bank. The expectations-based rule is constructed so that it exactly implements equation (11), the first-order condition under discretion, even outside a rational expectations equilibrium for given expectations, as suggested by Evans and Honkapohja (2003b). To obtain the rule, we combine equations (3), (4), and (11) and solve for i_t in terms of the exogenous shocks and the expectations. The resulting expectations-based rule is

$$i_t = \frac{(\alpha_x + \lambda^2)\varphi}{\alpha_i + (\alpha_x + \lambda^2)\varphi^2}\hat{E}_t x_{t+1} + \frac{\beta\lambda\varphi + (\alpha_x + \lambda^2)\varphi^2}{\alpha_i + (\alpha_x + \lambda^2)\varphi^2}\hat{E}_t \pi_{t+1}$$
$$+ \frac{(\alpha_x + \lambda^2)\varphi}{\alpha_i + (\alpha_x + \lambda^2)\varphi^2} g_t + \frac{\lambda\varphi}{\alpha_i + (\alpha_x + \lambda^2)\varphi^2} u_t.$$

This leads to a reduced form,

$$\mathbf{y}_t = \mathbf{M}\hat{E}_t\mathbf{y}_{t+1} + \mathbf{P}\mathbf{v}_t. \tag{14}$$

Determinacy of the rational expectations equilibrium corresponding to optimal discretionary monetary policy requires that $\mathbf{M}$ has both eigenvalues inside the unit circle.[15] We again have the condition $\alpha_i < \bar{\alpha}_i$, where $\bar{\alpha}_i$ is given by equation (13).

For stability under learning, first consider the case in which the exogenous shocks $\mathbf{v}_t$ are i.i.d. and agents use steady-state learning under constant gain. For this reduced form, the PLM $\mathbf{y}_t = \mathbf{a} + \mathbf{e}_t$ gives the ALM $\mathbf{y}_t = \mathbf{Ma} + \mathbf{e}_t$ (where $\mathbf{e}_t = \mathbf{P}\mathbf{v}_t$), as discussed in section 1.1. Thus $\mathbf{T} = \mathbf{M}$, and there is a very close connection between determinacy and stability under learning. This leads to proposition 4:

Proposition 4. Assume that $\alpha_i < \bar{\alpha}_i$ and that the shocks are i.i.d. Then the expectations-based rule, which implements the first-order condition, yields a reduced form that is stable under steady-state learning for all constant-gain rules $0 < \gamma \leq 1$.

Provided $\alpha_i < \bar{\alpha}_i$, so that determinate optimal policy is possible, the expectations-based optimal rule will successfully implement the

15. Equivalently, we need $|\mathrm{tr}(\mathbf{M})| < 1 + \det(\mathbf{M})$ and $|\det(\mathbf{M})| < 1$.

optimal rational expectations equilibrium: under decreasing-gain learning there will be convergence to the equilibrium, and under small constant-gain learning, it will converge to a stochastic process near the optimal equilibrium. Furthermore, for all constant gains $0 < \gamma \leq 1$, there will be convergence to a stationary process centered at the optimal equilibrium.

Second, we examine numerically the case of AR(1) shocks with (constant-gain) RLS learning. For the Woodford calibration $\beta = 0.99$, $\varphi = 1/0.157$, $\lambda = 0.024$, $\alpha_x = 0.048$, and $\alpha_i = 0.077$ (and $\rho = \mu = 0.8$), we find that learning converges for gain values at or below $\gamma = 0.925$. In other words, the expectations-based optimal discretionary rule is quite robustly stable under learning. When the agents have to run genuine regressions, as in the current case, then the IE-stability condition does not imply convergence of constant-gain learning for all $0 < \gamma \leq 1$. However, we see that stability does hold even for γ quite close to one.

3. Optimal Policy with Commitment

For brevity, in the remainder of the paper we assume that $\alpha_i = 0$, that is, that the central bank does not have an interest rate stabilization objective.[16] Given the model described in equations (3) and (4) and the loss function (10) with $\alpha_i = 0$, optimal monetary policy under commitment (from a timeless perspective) is characterized by the condition[17]

$$\lambda\pi_t = -\alpha_x(x_t - x_{t-1}), \tag{15}$$

which is often called the optimal targeting rule. The optimal rational expectations equilibrium of interest has the form

$$x_t = b_x x_{t-1} + c_x u_t$$

and

$$\pi_t = b_\pi x_{t-1} + c_\pi u_t,$$

16. See Duffy and Xiao (2007) for an extension to the case in which the central bank also has an interest rate stabilization motive.

17. See, for example, Clarida, Galí, and Gertler (1999) and Woodford (1999). For the exposition, we follow Evans and Honkapohja (2006).

where we choose the unique $0 < b_x < 1$ that solves the equation $\beta b_x^2 - (1 + \beta + \lambda^2/\alpha_x)\, b_x + 1 = 0$ and $b_\pi = \alpha_x/\lambda(1 - b_x)$, $c_x = -[\lambda + \beta b_\pi + (1 - \beta\rho)(\alpha_x/\lambda)]^{-1}$, and $c_\pi = -(\alpha_x/\lambda)\, c_x$.

The literature proposes a number of optimal reaction functions that implement the optimal targeting rule (15). Under rational expectations, one obtains the fundamentals-based reaction function

$$i_t = \psi_x x_{t-1} + \psi_g g_t + \psi_u u_t, \tag{16}$$

where

$$\psi_x = b_x[\varphi^{-1}(b_x - 1) + b_\pi],$$

$$\psi_g = \varphi^{-1},$$

and

$$\psi_u = [b_\pi + \varphi^{-1}(b_x + \rho - 1)]\, c_x + c_\pi\rho.$$

Evans and Honkapohja (2006) show that the reaction function (16) often leads to indeterminacy and always leads to expectational instability. They propose instead the expectations-based reaction function

$$i_t = \delta_L x_{t-1} + \delta_\pi \hat{E}_t \pi_{t+1} + \delta_x \hat{E}_t x_{t+1} + \delta_g g_t + \delta_u u_t, \tag{17}$$

where the coefficients are[18]

$$\delta_L = \frac{-\alpha_x}{\varphi(\alpha_x + \lambda^2)}, \; \delta_\pi = 1 + \frac{\lambda\beta}{\varphi(\alpha_x + \lambda^2)}, \; \delta_x = \delta_g = \varphi^{-1}, \text{ and}$$

$$\delta_u = \frac{\lambda}{\varphi(\alpha_x + \lambda^2)}.$$

Under the interest rate reaction rule (17), the reduced-form model is of the form

$$\mathbf{y}_t = \mathbf{M}_1 \hat{E}_t \mathbf{y}_{t+1} + \mathbf{N}\mathbf{y}_{t-1} + \mathbf{P}\mathbf{v}_t,$$

18. In the discretionary case with $\alpha_i = 0$, the same coefficients would obtain, except that $\delta_L = 0$.

with $\mathbf{y}_t' = (x_t, \pi_t)$ and $\mathbf{v}_t' = (g_t, u_t)$. The corresponding rational expectations equilibrium takes the form $\mathbf{y}_t = \bar{\mathbf{b}}\mathbf{y}_{t-1} + \bar{\mathbf{c}}\mathbf{v}_t$. Evans and Honkapohja (2006) show that the optimal expectations-based reaction function (17) delivers a determinate and E-stable optimal equilibrium for all values of the parameters. It is therefore clearly preferred to the fundamentals-based rule (16).

In connection with constant-gain learning we have the following partial result:[19]

Proposition 5. The expectations-based rule under commitment (17) yields a reduced form for which the eigenvalues of the derivative of the **T** map, at the rational expectations equilibrium, are inside the unit circle for all values of the structural parameters.

This result is partial in the sense that the eigenvalues condition is no longer sufficient for stability of constant-gain learning for all $0 < \gamma \leq 1$. This is because in the model the regressors include exogenous and lagged endogenous variables.

We now examine numerically the performance of constant-gain RLS learning under the expectations-based optimal rule with commitment. Using Woodford's parameter values (but with $\alpha_i = 0$), we find that constant-gain RLS learning converges for values of the gain parameter below $\hat{\gamma} \approx 0.25$. The inclusion of a lagged variable among the regressors appears to have a significant effect on learning stability for large gains. However, the rule is still robust for all plausible values of the gain parameter.

As noted above, the Duffy and Xiao (2007) formulation under commitment breaks down when $\alpha_i = 0$ (as it does in the discretionary case). One might investigate numerically the performance of the Duffy-Xiao rule under constant-gain RLS for calibrated values of α_i. Based on the results in the discretionary case, we are not optimistic about robust learning stability of the Duffy-Xiao rule with commitment.

4. Alternative Rules for Optimal Policy under Commitment

This section explores two alternative rules for optimal policy under commitment: the Svensson-Woodford rule and the McCallum-Nelson rule.

19. See the appendix for a proof.

4.1 Svensson-Woodford Rule

Given that the fundamentals-based optimal rules (without interest rate stabilization) lead to problems of indeterminacy and learning instability, Svensson and Woodford (2005) suggest a modification in which the fundamentals-based rule (16) is complemented with a term based on the commitment optimality condition. We again assume that contemporaneous data are not available to the policymaker, so that current values of inflation π_t and the output gap x_t are replaced by their nowcasts $\hat{E}_t\pi_t$ and $\hat{E}_t x_t$. This results in the interest rate rule

$$i_t = \psi_x x_{t-1} + \psi_g g_t + \psi_u u_t + \theta[\hat{E}_t\pi_t + \frac{\alpha_x}{\lambda}(\hat{E}_t x_t - x_{t-1})], \tag{18}$$

where $\theta > 0$.

The full model is now given by equations (3), (4), and (18). By substituting equation (18) into equation (3), we can reduce this model to a bivariate model of the form

$$\mathbf{y}_t = \mathbf{M}_0\hat{E}_t\mathbf{y}_t + \mathbf{M}_1\hat{E}_t\mathbf{y}_{t+1} + \mathbf{N}\mathbf{y}_{t-1} + \mathbf{P}\mathbf{v}_t, \tag{19}$$

where the information set in the forecasts and nowcasts includes current values of the exogenous shocks but not of the endogenous variables. We also assume for convenience that $\mathbf{v}_t = \mathbf{F}\mathbf{v}_{t-1} + \tilde{\mathbf{v}}_t$ is a known, stationary process. The coefficient matrices are

$$\mathbf{M}_0 = \begin{pmatrix} -\varphi\alpha_x\theta\lambda^{-1} & -\varphi\theta \\ -\varphi\alpha_x\theta & -\varphi\theta\lambda \end{pmatrix},$$

$$\mathbf{M}_1 = \begin{pmatrix} 1 & \varphi \\ \lambda & \beta + \lambda\varphi \end{pmatrix},$$

$$\mathbf{N} = \begin{pmatrix} -\varphi\psi_x + \varphi\alpha_x\theta\lambda^{-1} & 0 \\ -\lambda\varphi\psi_x + \varphi\alpha_x\theta & 0 \end{pmatrix},$$

and

$$\mathbf{P} = \begin{pmatrix} 0 & -\varphi\psi_u \\ 0 & 1 - \lambda\varphi\psi_u \end{pmatrix}.$$

The PLM has the form

$$\mathbf{y}_t = \mathbf{a} + \mathbf{b}\mathbf{y}_{t-1} + \mathbf{c}\mathbf{v}_t,$$

and the $\mathbf{T}$ mapping is

$$\mathbf{T}(\mathbf{a},\mathbf{b},\mathbf{c}) = \left\{ \begin{matrix} [\mathbf{M}_0 + \mathbf{M}_1(\mathbf{I}+\mathbf{b})]\mathbf{a}, \mathbf{M}_1\mathbf{b}^2 + \mathbf{M}_0\mathbf{b} + \mathbf{N}, \\ \mathbf{M}_0\mathbf{c} + \mathbf{M}_1(\mathbf{bc} + \mathbf{cF}) + \mathbf{P} \end{matrix} \right\}.$$

The usual E-stability conditions are stated in terms of the eigenvalues of the derivative matrices,

$$\mathbf{DT_a} = \mathbf{M}_0 + \mathbf{M}_1(\mathbf{I} + \bar{\mathbf{b}}),$$

$$\mathbf{DT_b} = \bar{\mathbf{b}}' \otimes \mathbf{M}_1 + \mathbf{I} \otimes \mathbf{M}_1\bar{\mathbf{b}} + \mathbf{I} \otimes \mathbf{M}_0,$$

and

$$\mathbf{DT_c} = \mathbf{F}' \otimes \mathbf{M}_1 + \mathbf{I} \otimes \mathbf{M}_1\bar{\mathbf{b}} + \mathbf{I} \otimes \mathbf{M}_0,$$

where $\otimes$ is the Kronecker product and $\bar{\mathbf{b}}$ is the rational expectations value of $\mathbf{b}$.

We compute numerically the E-stability eigenvalues for the Woodford calibration with $\alpha_x = 0.048$ and $\theta = 1.0$. For this case the eigenvalues of $\mathbf{DT_a}$ are –9.570 and 0.990, while the eigenvalues of $\mathbf{DT_b}$ are –10.605, –9.672, 0.878, and –0.0118. However, $\theta = 1.0$ is very close to the lower bound on θ needed for E-stability (since one root of $\mathbf{DT_a}$ is almost one), and the eigenvalues are sensitive to the value of θ. For example, for $\theta = 1.5$, the eigenvalues of $\mathbf{DT_a}$ are –15.975 and 0.949, while the eigenvalues of $\mathbf{DT_b}$ are –17.059, –16.082, 0.842 and –0.011. Thus, large negative eigenvalues appear.[20]

The calculation of the E-stability eigenvalues suggests that the interest rate rule (18) can be subject to instability if learning is based on constant gain. We now examine numerically the performance of rule (18) under different values of the constant gain using the Woodford calibrated values of the model parameters and $\theta = 1.5$. Numerical

20. The eigenvalues of the same model, but with contemporaneous data available, would not deliver large negative eigenvalues in the E-stability calculation for this parameterization.

simulations show that under the interest rate rule (18), constant-gain RLS learning becomes unstable for values of γ at 0.019 or higher.

We also examine numerically the sensitivity of the stability upper bound on γ for different values of α_x, that is, the degree of flexibility of inflation targeting. Table 1 gives the approximate highest value, $\hat{\gamma}$, of the gain for which stability under constant-gain learning obtains. The table shows that robust learning stability of the Svensson-Woodford hybrid rule is very sensitive to the degree of flexibility in inflation targeting. Robust stability obtains only when the central bank is an inflation hawk.

Table 1. Critical Values of γ for Stability: Svensson-Woodford Rule

α_x	0.01	0.02	0.03	0.04	0.05	0.06	0.08	0.10
$\hat{\gamma}$	0.185	0.060	0.035	0.020	0.018	0.014	0.009	0.007

Source: Authors' calculations.

4.2 McCallum-Nelson Rule

McCallum and Nelson (2004) propose a different rule that approximates optimal interest rate policy from a timeless perspective. They suggest that the interest rate be raised above inflation whenever the timeless-perspective optimality condition is above zero. Their rule performs well if $\mathbf{y}_t$ is observable, but as McCallum and Nelson (2004) themselves point out, such a rule would be subject to the operationality problem that we have encountered several times: it presupposes that contemporaneous data on inflation and the output gap are available. One way to overcome this problem is to replace unknown contemporaneous data by nowcasts of the variables. In this case, the interest rate rule becomes

$$i_t = \hat{E}_t\pi_t + \theta[\hat{E}_t\pi_t + \frac{\alpha_x}{\lambda}(\hat{E}_t x_t - x_{t-1})]. \tag{20}$$

Under rational expectations, this rule approximates optimal policy under (timeless-perspective) commitment, provided $\theta > 0$ is large.

The model is then given by equations (3), (4), and (20). The model can be reduced to a bivariate model of the form (19), where the coefficient matrices are

$$\mathbf{M}_0 = \begin{pmatrix} -\theta\varphi\alpha_x\lambda^{-1} & -\varphi(1+\theta) \\ -\theta\varphi\alpha_x & -\varphi\lambda(1+\theta) \end{pmatrix},$$

$$\mathbf{M}_1 = \begin{pmatrix} 1 & \varphi \\ \lambda & \beta+\lambda\varphi \end{pmatrix},$$

$$\mathbf{N} = \begin{pmatrix} -\theta\varphi\alpha_x\lambda^{-1} & 0 \\ \theta\varphi\alpha_x & 0 \end{pmatrix},$$

and

$$\mathbf{P} = \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix}.$$

Using the same parameter values as in the case of the Svensson-Woodford hybrid rule, with $\alpha_x = 0.048$, we obtain that for $\theta = 1.0$, the eigenvalues of $\mathbf{DT_a}$ are –9.719 and 0.869, while the eigenvalues of $\mathbf{DT_b}$ are –10.780, –9.833, 0.750, and –0.213. For $\theta = 1.5$ the eigenvalues of $\mathbf{DT_a}$ are –16.130 and 0.873, while the eigenvalues of $\mathbf{DT_b}$ are –17.228, –16.245, 0.762 and –0.172. The results are very sensitive to α_x. For $\alpha_x = 0.100$, we obtain that for $\theta = 1.0$ the eigenvalues of $\mathbf{DT_a}$ are –22.954 and 0.912, while the eigenvalues of $\mathbf{DT_b}$ are –24.042, –23.033, 0.835 and –0.143. The large negative eigenvalues indicate the potential for instability under constant-gain learning. Using the Woodford calibration (including $\alpha_x = 0.048$) and choosing $\theta = 1.5$, we find that constant-gain RLS learning becomes unstable for values of the gain at or above 0.017.

We again examine numerically the sensitivity of the stability upper bound on γ for different values of α_x, that is, the degree of flexibility of inflation targeting. Table 2 gives the approximate highest value $\hat{\gamma}$ of the gain for which stability under constant-gain learning obtains. Comparing the two tables reveals that the stability performance of the McCallum-Nelson rule (20) is about the same as that of the hybrid rule (18) for the same parameter values. Neither rule is robust for many plausible values of the gain parameter.

Table 2. Critical Values of γ for Stability: McCallum-Nelson Rule

α_x	0.01	0.02	0.03	0.04	0.05	0.06	0.08	0.10
$\hat{\gamma}$	0.174	0.057	0.031	0.020	0.017	0.014	0.009	0.007

Source: Authors' calculations.

McCallum and Nelson (2004) suggest that a preferable alternative to equation (20) is to use forward expectations instead of nowcasts, since this delivers superior results under rational expectations. In this case, the model has no lagged endogenous variables, that is, $\mathbf{N} = \mathbf{0}$ in equation (19). We analyze this case numerically in Evans and Honkapohja (2003a, 2006). Large negative eigenvalues no longer arise in this formulation. However, determinacy and E-stability require a small value of the parameter θ, which can result in significant welfare losses for optimal policy.

5. Conclusions

A lot of recent applied research on learning and monetary policy emphasizes discounted (constant-gain) least-squares learning by private agents. We have examined the stability performance of various operational interest rate rules under constant-gain learning for different values of the gain parameter. Since estimates of the gain parameter tend to be in the range of 0.02 to 0.06 for quarterly macroeconomic data, ideally there should be convergence of learning for gain parameters up to 0.1. Based on this criterion, we have found that many proposed interest rate rules are not robustly stable under learning in this sense. An exception to this finding is the class of expectations-based optimal rules in which the interest rate depends on private expectations in an appropriate way.

Appendix
Constant-Gain RLS Algorithm

Suppose the economy is described in terms of a multivariate linear model, which includes possible dependence on lagged endogenous variables. Under least-squares learning, agents have the PLM

$$\mathbf{y}_t = \mathbf{a} + \mathbf{b}\mathbf{y}_{t-1} + \mathbf{c}\mathbf{v}_t + \mathbf{e}_t, \tag{21}$$

where $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ denote parameters to be estimated. Here $\mathbf{y}_t$ is a $p \times 1$ vector of endogenous variables. $\mathbf{v}_t$ is $k \times 1$ vector of observable exogenous variables, and $\mathbf{e}_t$ is a vector of white noise shocks. If the model does not have lagged endogenous variables, then the term $\mathbf{b}\mathbf{y}_{t-1}$ is omitted.

At time t agents compute their forecasts using equation (21) with the estimated values $(\mathbf{a}_t, \mathbf{b}_t, \mathbf{c}_t)$ based on data up to period $t - 1$. Constant-gain RLS takes the form

$$\xi_t = \xi_{t-1} + \gamma \mathbf{R}_t^{-1} \mathbf{Z}_{t-1}(\mathbf{y}_{t-1} - \xi_{t-1}' \mathbf{Z}_{t-1})',$$

$$\mathbf{R}_t = \mathbf{R}_{t-1} + \gamma(\mathbf{Z}_{t-1}\mathbf{Z}_{t-1}' - \mathbf{R}_{t-1}),$$

where $\xi_t' = (\mathbf{a}_t, \mathbf{b}_t, \mathbf{c}_t)$, $\mathbf{Z}_t' = (1, \mathbf{y}_{t-1}', \mathbf{v}_t')$, and $1 > \gamma > 0$. The algorithm starts at $t = 1$ with a complement of initial conditions. The only difference from standard RLS is that the latter assumes a decreasing gain $\gamma_t = 1/t$.[21]

Proof of proposition 5

We now sketch a proof of proposition 5. We examine the formulas given in equations (A7) through (A9) of Evans and Honkapohja (2006, p. 36). Two of the eigenvalues of $\mathbf{DT_b}$ are 0, while the remaining eigenvalues are those of the matrix

21. The formal analysis of recursive least squares (RLS) learning in linear multivariate models is developed, for example, in Evans and Honkapohja (1998; 2001, chap. 10).

$$\mathbf{K_b} = \begin{pmatrix} \dfrac{-\lambda\beta b_\pi}{\alpha_x + \lambda^2} & \dfrac{-\lambda\beta b_x}{\alpha_x + \lambda^2} \\ \dfrac{\alpha_x\beta b_\pi}{\alpha_x + \lambda^2} & \dfrac{\alpha_x\beta b_x}{\alpha_x + \lambda^2} \end{pmatrix}.$$

The eigenvalues of $\mathbf{K_b}$ are 0 and $-1 < \alpha_x\beta(2b_x - 1)/(\alpha_x + \lambda^2) < 1$. Likewise, two of the eigenvalues of $\mathbf{DT_c}$ are 0, while the other two eigenvalues are those of the matrix

$$\mathbf{K_c} = \begin{pmatrix} \dfrac{-\lambda\beta b_\pi}{\alpha_x + \lambda^2} & \dfrac{-\lambda\beta\rho}{\alpha_x + \lambda^2} \\ \dfrac{\alpha_x\beta b_\pi}{\alpha_x + \lambda^2} & \dfrac{\alpha_x\beta\rho}{\alpha_x + \lambda^2} \end{pmatrix}.$$

The eigenvalues of $\mathbf{K_c}$ are 0 and $\alpha_x\beta(b_x - 1 + \rho)/(\alpha_x + \lambda^2)$, which is inside the unit circle unless ρ is negative and large in magnitude. Finally,

$$\mathbf{DT_a} = \begin{pmatrix} \dfrac{-\lambda\beta b_\pi}{\alpha_x + \lambda^2} & \dfrac{-\lambda\beta}{\alpha_x + \lambda^2} \\ \dfrac{\alpha_x\beta b_\pi}{\alpha_x + \lambda^2} & \dfrac{\alpha_x\beta}{\alpha_x + \lambda^2} \end{pmatrix},$$

and its eigenvalues are 0 and $0 < \alpha_x\beta b_x/(\alpha_x + \lambda^2) < 1$.

References

Branch, W.A. and G.W. Evans. 2006. "A Simple Recursive Forecasting Model." *Economic Letters* 91(2): 158–66.

Bullard, J. 2006. "The Learnability Criterion and Monetary Policy." *Federal Reserve Bank of St. Louis Review* 88 (May): 203–17.

Bullard, J. and K. Mitra. 2002. "Learning about Monetary Policy Rules." *Journal of Monetary Economics* 49(6): 1105–29.

Clarida, R., J. Galí and M. Gertler. 1999. "The Science of Monetary Policy: A New Keynesian Perspective." *Journal of Economic Literature* 37(4): 1661–707.

Duffy, J. and W. Xiao. 2007. "The Value of Interest Rate Stabilization Policies When Agents Are Learning." *Journal of Money, Credit, and Banking* 39(8): 2041–56.

Evans, G.W. and R. Guesnerie. 1993. "Rationalizability, Strong Rationality, and Expectational Stability." *Games and Economic Behaviour* 5(4): 632–46.

Evans, G.W. and S. Honkapohja. 1998. "Economic Dynamics with Learning: New Stability Results." *Review of Economic Studies* 65(1): 23–44.

———. 2001. *Learning and Expectations in Macroeconomics*. Princeton University Press.

———. 2003a. "Adaptive Learning and Monetary Policy Design." *Journal of Money, Credit, and Banking* 35(6): 1045–72.

———. 2003b. "Expectations and the Stability Problem for Optimal Monetary Policies." *Review of Economic Studies* 70(4): 807–24.

———. 2006. "Monetary Policy, Expectations, and Commitment." *Scandinavian Journal of Economics* 108: 15–38.

Evans, G.W., S. Honkapohja and N. Williams. Forthcoming. "Generalized Stochastic Gradient Learning." *International Economic Review*.

McCallum, B.T. 1999. "Issues in the Design of Monetary Policy Rules." In *Handbook of Macroeconomics*, vol. 1, edited by J. Taylor and M. Woodford, chap. 23. Amsterdam: Elsevier.

McCallum, B.T. and E. Nelson. 2004. "Timeless Perspective vs. Discretionary Monetary Policy in Forward-Looking Models." *Federal Reserve Bank of St.Louis Review* 86: 43–56.

Milani, F. 2005. "Adaptive Learning and Inflation Persistence." Working paper 050607. University of California at Irvine, Department of Economics.

————. 2007a. “Expectations, Learning, and Macroeconomic Persistence.” *Journal of Monetary Economics* 54(7): 2065–82.

————. 2007b. “Learning and Time-Varying Macroeconomic Volatility.” Working paper 070802. University of California at Irvine, Department of Economics.

Orphanides, A. and J.C. Williams. 2005a. “Imperfect Knowledge, Inflation Expectations, and Monetary Policy.” In *The Inflation-Targeting Debate*, edited by B. Bernanke and M. Woodford, chap. 5. University of Chicago Press.

————. 2005b. “The Decline of Activist Stabilization Policy: Natural Rate Misperceptions, Learning, and Expectations.” *Journal of Economic Dynamics and Control* 29(11): 1927–50.

Slobodyan, S., A. Bogomolov and D. Kolyuzhnov. 2006. “Stochastic Gradient Learning versus Recursive Least Squares Learning.” Working paper 309. CERGE-EI.

Svensson, L.E.O. and M. Woodford. 2005. “Implementing Optimal Policy through Inflation-Forecast Targeting.” In *The Inflation-Targeting Debate*, edited by B. Bernanke and M. Woodford, pp. 19–83. University of Chicago Press.

Woodford, M. 1999. “Optimal Monetary Policy Inertia.” Working paper 7261. Cambridge, Mass.: National Bureau of Economic Research.

————. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Macroeconomic and Monetary Policies from the Eductive Viewpoint

Roger Guesnerie
Collège de France and Paris School of Economics

The quality of the coordination of expectations, a key issue for monetary policy, obtains from different, but interrelated, channels: both the credibility of the central bank intervention and the ability of decentralized agents to coordinate on a dynamical equilibrium matter. For both purposes, it is important to understand how agents learn. Indeed, many studies on monetary policy focus on learning processes involving evolutive, real-time learning rules (such as adaptive learning rules).

The eductive viewpoint, as illustrated in Guesnerie (2005) and other references cited herein, partly abstracts from the real-time dimension of learning, with the aim of more directly capturing the systems' coordination-friendly characteristics. The paper first presents the analytical philosophy of expectational coordination underlying the eductive viewpoint. Providing a synthesis of the eductive viewpoint is a prerequisite to comparing the methods that this viewpoint suggests with those actually adopted in most present studies of learning in the context of macroeconomic and monetary policy. Such a comparison rests on the review of existing learning results in the context of dynamic systems, which is currently the main field for applying the eductive method to macroeconomics.[1] Such applications, however, have not had a direct bearing on monetary policy issues. Following the review, the

I thank Carl E. Walsh for useful comments on an earlier draft and Xavier Ragot for discussions on these issues. I am especially grateful to Antoine d'Autume for pointing out an error in a previous version. Also, section 5 borrows significantly from the joint study of eductive learning in RBC-like models undertaken with George Evans and Bruce McGough (Evans, Guesnerie, and McGough, 2007).

1. See, in particular Evans and Guesnerie (2005); for a static macroeconomic example, see Guesnerie (2001).

Monetary Policy under Uncertainty and Learning, edited by Klaus Schmidt-Hebbel and Carl E. Walsh, Santiago, Chile.

paper explores the differences between the traditional viewpoint and this competing viewpoint as they relate to standard monetary policy analysis. This exploration is tentative, yet promising.

The paper proceeds as follows. The next section lays out the logic behind the eductive viewpoint and compares it with the evolutive approach. I then review results that support a comparison between the most standard expectational criteria and the eductive criterion, first in the framework of a simple one-dimensional dynamic system (section 2) and then in a multidimensional system (section 3). The comparison with standard approaches is completed in section 4. The analysis emphasizes the role of heterogeneity of expectations and may suggest that the alternative view completes and deepens—rather than contradicts—the conclusions of more standard approaches. However, section 5 undertakes an eductive analysis of a simple cashless economy in an infinite-horizon model with infinitely-lived agents, which stresses conditions for expectational coordination that are strikingly different from the classical ones. In particular, the eductive evaluation of the stabilizing performance of the Taylor rule suggests that its reaction coefficient to inflation has to be severely restricted.

1. Expectational Stability: The Eductive Viewpoint

The notion of an eductively stable or strongly rational equilibrium has game-theoretical underpinnings and draws on ideas like rationalizability, dominance solvability, common knowledge. These concepts serve to provide a high-tech justification of the proposed expectational stability criteria. The next subsection emphasizes this high-tech approach for proposing global stability concepts that have a clearly eductive flavor. The local view of the global approach allows a more intuitive, low-tech interpretation which is presented in the second subsection, and the section closes with comments on the connections between the eductive viewpoint and the standard evolutive learning viewpoint.

1.1 Global Eductive Stability

The model assumes rational economic agents (modeled as a continuum), who know the logic of the collective economic interactions (that is, the underlying model). Both the rationality of the agents and the model are common knowledge. The state of the system is denoted E and belongs to some subset ε of some vector space. The state E can be a number, (the value of an equilibrium price or a growth rate), a vector (of equilibrium prices, for example), a function (an equilibrium demand function), an infinite trajectory of states, or a probability distribution. For example, in the variant of the Muth model considered in Guesnerie (1992), E is a number—namely, the market clearing price tomorrow on the wheat market. The agents are farmers whose profits depend on the wheat price. They know the model in the sense that they understand how the market price depends on the total amount of wheat available tomorrow: the market clearing price, as a function of the total crop, is determined from the inverse of a known demand function. Agents know all this, (Bayesian) rationality and the model, and they know that it is known, and they know that it is known that it is known, and so on. With straightforward notation, (it is known)^p for any p (that is, it is common knowledge. In general equilibrium models (Guesnerie, 2001, 2002; Ghosal, 2006), E is a price or quantity vector. In models focusing on the transmission of information through prices (Desgranges, 2000; Desgranges and Heinemann, 2005; Desgranges, Geoffard, and Guesnerie, 2003), E is a function that relates the non-noisy part of excess demand to the asset price. In infinite horizon models, E is an infinite trajectory consisting, at each date t, of either a number or a vector, describing the state of the system at this date. Introducing uncertainty in these partial equilibrium, general equilibrium, and intertemporal models leads to substituting E with a probability distribution over the set of finite or infinite dimensional vectors previously considered.

In this paper, I focus on rational expectations or perfect foresight equilibria. Emphasizing the expectational aspects of the problem, I view an equilibrium of the system as a state, E^*, that prevails if everybody believes that it prevails. Note that in the described context, E^* is such that the assertion, "it is common knowledge that $E = E^*$" is meaningful.

I say that E^* is eductively stable or strongly rational if the following assertion A implies assertion B (given that Bayesian rationality and the model are common knowledge):

Assertion A: It is common knowledge that $E \in \varepsilon$;

Assertion B: It is common knowledge that $E = E^*$.

The mental process that leads from assertion A to assertion B is as follows. First, since everybody knows that $E \in \varepsilon$, everybody knows that everybody limits their responses to actions that are the best responses to some probability distributions over ε. It follows that everybody knows that the state of the system will be in $\varepsilon(1) \subset \varepsilon$. Second, if $\varepsilon(1)$ is a proper subset of ε, the mental process goes on as in the first step, but it is now based on $\varepsilon(1)$ instead of ε. Third, the process continues indefinitely, resulting in a (weakly) decreasing sequence $\varepsilon(n) \subset \varepsilon(n-1) \subset \ldots \subset \varepsilon(1) \subset \varepsilon$. When the sequence converges to E^*, the equilibrium is strongly rational or eductively stable. When convergence does not occur, the limit set is the set of rationalizable equilibria of the model (see Guesnerie and Jara-Moroni, 2007).

Global eductive stability is clearly very demanding, although it can be shown to hold under plausible economic conditions in a variety of models, including partial and general equilibrium standard market contexts (Guesnerie, 1992, 2001), financial models of the transmission of information through prices (Desgranges, Geoffard, and Guesnerie, 2003), and general settings involving strategic complementarities or substitutabilities (Guesnerie and Jara-Moroni, 2007).

1.2 Local Eductive Stability

Local eductive stability may be defined through the same high-tech or hyperrational view. However, the local criterion also has a very intuitive, low-tech, and in a sense boundedly rational interpretation.

1.2.1 Local eductive stability as a common knowledge statement

I say that E^* is locally eductively stable or locally strongly rational if there is some nontrivial neighborhood of E^*, $V(E^*)$, such that assertion A implies assertion B:

Assertion A: It is common knowledge that $E \in V(E^*)$;

Assertion B: it is common knowledge that $E = E^*$.

Hypothetically, the state of the system is assumed to be in some nontrivial neighborhood of E^*, and this hypothetical assumption of common knowledge implies common knowledge of E^*. In other words, the deletion of non-best responses starts under the assumption that the system is close to its equilibrium state. In that sense, this is the same hyperrational view referred to above. However, the statement can be read in a simpler way.

1.2.2 Local eductive stability as a common sense requirement

An intuitively plausible definition of local expectational stability is as follows: there is a nontrivial neighborhood of the equilibrium such that if everybody believes that the state of the system is in this neighborhood, it is necessarily the case that the state is, in fact, in this neighborhood, regardless of the specific form of everybody's belief. Intuitively, the absence of such a neighborhood signals some tendency to instability: there can be facts falsifying any universally shared conjecture on the set of possible states, unless this set reduces to the equilibrium itself. The failure of local expectational stability in the precise sense defined above is (roughly) equivalent to a failure of the local intuitive requirement.

1.3 Eductive versus Evolutive Learning Stability

Milgrom and Roberts (1990) suggest an informal argument according to which, in a system that repeats itself, non-best responses to existing observations will be deleted after a while, initiating a real-time counterpart of the notional-time deletion of non-best responses that underlies eductive reasoning. I focus here on the connections between local eductive stability and the local convergence of standard evolutive learning rules. Local eductive stability, as just defined, implies that once the (possibly stochastic) beliefs of the agents are, for whatever reason, trapped in $V(E^*)$, they will remain in $V(E^*)$ whenever updating satisfies natural requirements that are met in particular by Bayesian updating rules. Although this does not guarantee that any evolutive learning rule will converge, local eductive stability does mean that every reasonable evolutive real-time learning rule will converge asymptotically in many settings (see Guesnerie, 2002; Gauthier and Guesnerie, 2005). Furthermore, the failure to find a set $V(E^*)$ for which the equilibrium is locally strongly rational signals a tendency to trigger away in some cases reasonable states of beliefs that are close to the equilibrium (and are thus likely to be reachable with some reasonable evolutive updating process) a fact that threatens the convergence of the corresponding learning rule.[2]

2. It also forbids a strong form of monotonic convergence.

The very abstract and hyperrational criterion thus provides a shortcut for understanding the difficulties of expectational coordination, without entering into the business of specifying the real-time bounded rationality considerations. Naturally, the eductive criterion is generally more demanding than most fully specified evolutive learning rules (as strongly suggested by the argument sketched above and precisely shown in the previously cited works).

The connection, however, is less clear-cut than just suggested in models with extrinsic uncertainty. In such cases, the equilibrium, as well as a state of the system in the sense of the word used here, is a probability distribution. However, an observation is not an observation on the state in this sense, but information on the state in the standard sense of the word. Evolutive and eductive learning may then differ significantly.

2. Expectational Coordination: Infinite Horizon and One-Dimensional State

Models used for monetary policy generally adopt an infinite horizon approach. This section and the following review existing results on expectational coordination in general and eductive stability in particular, in infinite horizon models. They are based on Gauthier (2003), Evans and Guesnerie (2003, 2005), and Gauthier and Guesnerie (2005). The review will support an expansion of the comparison of the game-theoretical viewpoint stressed in this paper with the standard macroeconomic approach to the problem as reported in Evans and Honkapohja (2001). I start with one-dimensional one-step-forward models with one-period memory.

2.1 The Model

Consider a model in which the one-dimensional state of the system today is determined from its value yesterday and its expected value tomorrow, according to the following linear (for the sake of simplicity) equation:

$$\gamma E\,[x(t+1)\mid I_t] + x(t) + \delta\, x(t-1)=0,$$

where x is a one-dimensional variable and γ and δ are real parameters (γ, $\delta \neq 0$).[3]

A perfect foresight trajectory is a sequence ($x(t)$, $t \geq -1$) such that

$$\gamma x(t + 1) + x(t) + \delta x(t - 1)=0 \tag{1}$$

in any period $t \geq 0$, given the initial condition $x(-1)$.

Assume that the equation $g_1 = -\gamma g_1^2 - \delta$ has only two real solutions, λ_1 and λ_2 (which arise if and only if $1 - \delta\gamma \geq 0$), with different moduli (with $|\lambda_1| < |\lambda_2|$ by definition). Given an initial condition $x(-1)$, there are many perfect foresight solutions, but only two perfect foresight solutions have the simple form

$$x(t) = \lambda_1 x(t - 1)$$

and

$$x(t) = \lambda_2 x(t - 1).$$

They are called constant growth rate solutions.

The steady-state sequence ($x(t) = 0$, $t \geq -1$) is a perfect foresight equilibrium if and only if the initial state $x(-1)$ equals 0. The steady state is a sink if $|\lambda_2| < 1$, a saddle if $|\lambda_1| < 1 < |\lambda_2|$, or a source if $|\lambda_1| > 1$. I focus here on the saddle case, for which the solution, $x(t) = \lambda_1 x(t - 1)$, is generally called the saddle path. Economists have long considered this the focal solution, on the basis of arguments that refer to expectational plausibility. The rest of this section reviews the standard expectational criteria that are used and confirms that the saddle-path solution fits them.

2.2 The Standard Expectational Criteria

The standard expectational criteria basically fall into four categories: determinacy, immunity to sunspots, evolutive learning, and iterative expectational stability. I briefly explore each of these in turn and then relate their solutions in an equivalence theorem.

3. Such dynamics obtain, in particular, from linearized versions of overlapping generations models with production, at least for particular technologies (Reichlin, 1986), or infinite horizon models with a cash-in-advance constraint (Woodford, 1994).

2.2.1 Determinacy

The first criterion is determinacy. Determinacy means that the equilibrium under consideration is locally isolated. In an infinite horizon setting, determinacy has to be viewed as a property of trajectories: a trajectory $(x(t), t \geq -1)$ is determinate if there is no other equilibrium trajectory $(x'(t), t \geq -1)$ that is close to it. This calls for a reflection about the notion of proximity of trajectories, that is, on the choice of a topology. While the choice of the suitable topology is open, the most natural candidate is the C_0 topology, according to which two different trajectories, $(x(t), t \geq -1)$ and $(x'(t), t \geq -1)$, are said to be close whenever $|x(t) - x'(t)| < \varepsilon$, for any arbitrarily small $\varepsilon > 0$ and any date $t \geq -1$. In fact, with such a concept of determinacy, the saddle-path solution, along which $x(t) = \lambda_1 x(t-1)$ when $|\lambda_1| < 1 < |\lambda_2|$, is the only solution to be locally isolated—that is, determinate—in the C_0 topology.

In the present context of models with memory, a saddle-path solution is characterized by a constant growth rate of the state variable. This suggests that determinacy should be applied in terms of growth rates, in which case the closeness of two trajectories, $(x(t), t \geq -1)$ and $(x'(t), t \geq -1)$, would require that the ratio $x(t) / x(t-1)$ be close to $x'(t) / x'(t-1)$ in each period $t \geq 0$. This is an ingredient of a kind of C_1 topology, as advocated by Evans and Guesnerie (2003). In this topology, two trajectories, $(x(t), t \geq -1)$ and $(x'(t), t \geq -1)$, are said to be close whenever both the levels $x(t)$ and $x'(t)$ are close, and the ratios $x(t) / x(t-1)$ and $x'(t) / x'(t-1)$ are close in any period.

As emphasized by Gauthier (2002), the examination of proximity in terms of growth rates leads to the analysis of the dynamics with perfect foresight in terms of growth rates. Define $g(t) = x(t) / x(t-1)$ for any $x(t-1)$ and any $t \geq 0$. The perfect foresight dynamics then imply either

$$x(t) = -[\gamma g(t+1)\, g(t) + \delta]\, x(t-1)$$

or

$$g(t) = -[\gamma g(t+1)\, g(t) + \delta]. \tag{2}$$

The perfect foresight dynamics of growth rates then follows from the initial perfect foresight dynamics defined in equation (1). The growth factor $g(t)$ is determined at date t from the correct forecast of the next growth factor $g(t + 1)$. This new dynamics of equation (2) are nonlinear, and they have a one-step-forward-looking structure, without predetermined variables.

The problem has thus been reassessed in terms of one-dimensional one-step-forward-looking models that are more familiar.

2.2.2 Immunity to sunspots on growth rates

Maintaining the focus on growth rates, I now define a concept of sunspot equilibrium, in the neighborhood of a constant growth rate solution. Suppose that agents a priori believe that the growth factor is perfectly correlated with sunspots. Namely, if the sunspot event is $s = 1, 2$, at date t, they a priori believe that $g(t) = g(s)$, that is,

$$x(t) = g(s)\, x(t-1).$$

Thus, their common expected growth forecast is

$$E\,[x(t + 1) \mid I_t\,] = \pi(s, 1)\, g(1)\, x(t) + \pi(s, 2)\, g(2)\, x(t),$$

where $\pi(s, 1)$ and $\pi(s, 2)$ are the sunspot transition probabilities.

As shown by Desgranges and Gauthier (2003), this consistency condition is written

$$g(s) = -\{\gamma\, [\pi(s, 1)\, g(1) + \pi(s, 2)\, g(2)]\, g(s) + \delta\}. \tag{3}$$

When $g(1) \neq g(2)$, the formula defines a sunspot equilibrium on the growth rate, as soon as the stochastic dynamics of growth rates are extended:[4]

$$g(t) = -\gamma\, E\,[g(t+1) \mid I_t\,]\, g(t) - \delta.$$

2.2.3 Evolutive learning on growth rates

It makes sense to learn growth rates from past observations. Agents then update their forecast of the next period growth rates from the observation of past or present actual rates. Reasonable learning

4. This equivalence relies on special assumptions about linearity and certainty equivalence.

rules in the sense of Guesnerie (2002) and Gauthier and Guesnerie (2005) consist of adaptive learning rules that are able to detect cycles of order two.

2.2.4 Iterative expectational stability

This subsection applies the iterative expectational (IE) stability criterion (see, for example, Evans, 1985; DeCanio, 1979; Lucas, 1978)[5] to conjectures on growth rates. Let agents believe a priori that the law of motion of the economy is given by

$$x(t) = g(\tau)\, x(t-1),$$

where $g(\tau)$ denotes the conjectured growth rate at step τ in some mental reasoning process. They expect the next state variable to be $g(\tau)x(t)$, so that the actual value is $x(t) = -\delta\, x(t-1) / [\gamma g(\tau) + 1]$. Assume that all agents understand that the actual growth factor is $-\delta / [\gamma g(\tau) + 1]$. When their initial guess is $g(\tau)$, they should revise their guess as

$$g(\tau+1) = -\frac{\delta}{\gamma g(\tau)+1}. \tag{4}$$

By definition, IE stability obtains whenever the sequence $(g(\tau), \tau \geq 0)$ converges toward one of its fixed points, a fact that is interpreted as reflecting the success of some mental process of learning (leading to the constant growth rate associated with the considered fixed point). These dynamics are the time mirror of the perfect foresight growth rate dynamics: then, a fixed point λ_1 or λ_2 is locally IE stable if and only if it is locally unstable in the previous growth rate dynamics, that is, in these dynamics, it is locally determinate.

This simple model provides a somewhat careful reminder of the four possible (and more or less standard) viewpoints on expectational stability. I later compare these viewpoints with the so called eductive viewpoint emphasized here. This comparison is facilitated by the fact that these a priori different approaches to the problem select the same solutions, as described in the proposition below.

5. This concept differs from the more usual concept of differential expectational stability (see Evans and Honkapohja, 2001).

2.2.5 An equivalence theorem for standard expectational criteria

Proposition 1. For a one-step-forward, one-dimensional linear model (with one lagged predetermined variable, where γ, $\delta \neq 0$), the following four statements are equivalent:

1. A constant growth rate solution is locally determinate in the perfect foresight growth rate dynamics and equivalently here is determinate in the C_1 topology of trajectories.

2. A constant growth rate solution is locally immune to (stationary) sunspots on growth rates.

3. For any a priori given reasonable learning rules bearing on growth rates, a constant growth rate solution is locally asymptotically stable.

4. A constant growth rate solution is locally IE stable.

In particular, a saddle-path solution that clearly meets the first requirement meets all the others. The argument presented in Guesnerie (2002) incorporates earlier findings. For example, the fact that reasonable learning processes converge relies on a definition of reasonableness integrating the suggestions of Grandmont and Laroque (1991) and the results of Guesnerie and Woodford (1991).

Section 4 will compare the standard criteria with the eductive viewpoint on learning, but some game theory flesh will have to be introduced into the model. Before doing that, I focus on a multidimensional version of the model.

3. Standard Expectational Criteria in Infinite Horizon Models: The Multidimensional Case

While keeping with one-step-forward-looking linear models with one-period memory, I now turn to the case of a multidimensional state variable.

3.1 The Framework

The dynamics of the multidimensional linear one-step-forward-looking economy with one predetermined variable are now governed by

$$\mathbf{G}E\,[\mathbf{x}(t+1) \mid I_t] + \mathbf{x}(t) + \mathbf{D}\mathbf{x}(t-1) = \mathbf{0},$$

where $\mathbf{x}$ is an $n \times 1$ dimensional vector, $\mathbf{G}$ and $\mathbf{D}$ are two $n \times n$ matrices, and $\mathbf{0}$ is the $n \times 1$ zero vector. A perfect foresight equilibrium is a sequence $(\mathbf{x}(t),\, t \geq 0)$ associated with the initial condition $\mathbf{x}(-1)$, such that

$$\mathbf{G}\mathbf{x}(t+1) + \mathbf{x}(t) + \mathbf{D}\mathbf{x}(t-1) = \mathbf{0}. \tag{5}$$

The dynamics with perfect foresight are governed by the $2n$ eigenvalues λ_i ($i = 1,\ldots, 2n$) of the following matrix (the companion matrix associated with the recursive equation):

$$\mathbf{A} = \begin{pmatrix} -\mathbf{G}^{-1} & -\mathbf{G}^{-1}\mathbf{D} \\ \mathbf{I}_n & (\mathbf{0}) \end{pmatrix},$$

where $(\mathbf{0})$ is the n-dimensional zero matrix.

The discussion centers on the perfect foresight dynamics restricted to a n-dimensional eigensubspace, especially the one spanned by the eigenvectors associated with the n roots of lowest modulus. I assume that the eigenvalues are distinct and define $|\lambda_i| < |\lambda_j|$ whenever $i < j$ ($i, j = 1,\ldots, 2n$). I then focus on the generalized saddle-path case, where $|\lambda_n| < 1 < |\lambda_{n+1}|$.

Let $\mathbf{u}_i$ denote the eigenvector associated with λ_i ($i = 1,\ldots, 2n$). Since all the eigenvalues are distinct, the n eigenvectors form a basis of the subspace associated with $\lambda_1,\ldots, \lambda_n$. Let

$$\mathbf{u}_i = \begin{pmatrix} \tilde{\mathbf{v}}_i \\ \mathbf{v}_i \end{pmatrix},$$

where $\mathbf{v}_i$ and $\tilde{\mathbf{v}}_i$ are of dimension n. If $\mathbf{u}_i$ is an eigenvector, then $\tilde{\mathbf{v}}_i = \lambda_i \mathbf{v}_i$.

Hence, on picking up some $\mathbf{x}(0)$, and if the n-dimensional subspace generated by $(\mathbf{u}_1,\ldots, \mathbf{u}_n)$ is in "general position," there is a single $\mathbf{x}(1)$ such that $(\mathbf{x}(0), \mathbf{x}(1)) = \Sigma a_i \mathbf{u}_i$ is in the subspace. This generates a sequence $(\mathbf{x}(t),\, t \geq 0)$, $(\mathbf{x}(1), \mathbf{x}(2)) = \Sigma a_i \lambda_i \mathbf{u}_i$ following the dynamics defined in equation (5). This generates a solution that is converging in the saddle-path case.

The methodology proposed in the previous section for constructing a constant growth rate solution can be replicated to obtain what is called a minimum-order solution. Assume that

$$\mathbf{x}(t) = \mathbf{B}\,\mathbf{x}(t-1) \tag{6}$$

in every period t and for any n-dimensional vector $\mathbf{x}(t-1)$($\mathbf{B}$ is an $n \times n$ matrix). Also, $\mathbf{x}(t+1) = \mathbf{B}\mathbf{x}(t)$. It must therefore be the case that

$$\mathbf{B} = -\,(\mathbf{GB} + \mathbf{I}_n)^{-1}\,\mathbf{D},$$

or, equivalently

$$(\mathbf{GB} + \mathbf{I}_n)\,\mathbf{B} + \mathbf{D} = \mathbf{0}. \tag{7}$$

A matrix $\bar{\mathbf{B}}$ satisfying this equation is a minimum-order solution in the sense of McCallum (1983).[6] Gauthier (2002) calls it a stationary extended growth rate. In view of the previous section's analysis of constant growth rate solutions, I use this latter terminology and focus on the expectational stability of extended growth rates.

3.2 The Expectational Plausibility of Extended Growth Rate Solutions According to Standard Criteria

This section concentrates on three of the above criteria: determinacy, immunity to sunspots, and IE stability. Determinacy is viewed through the dynamics of perfect foresight of extended growth rates, which extends the growth rate dynamics previously introduced. For every t, $\mathbf{B}(t)$ is an n-dimensional matrix whose ij^{th} entry is equal to $b_{ij}(t)$ and $\mathbf{x}(t) = \mathbf{B}(t)\mathbf{x}(t-1)$. This matrix is called an extended growth rate (EGR), in line with the terminology of stationary extended growth rates.

Assume that such a relationship holds in all t, so that $\mathbf{x}(t+1) = \mathbf{B}(t+1)\mathbf{x}(t)$; the dynamics with perfect foresight of the endogenous state variable $\mathbf{x}(t)$ imply

$$\mathbf{GB}(t+1)\,\mathbf{x}(t) + \mathbf{x}(t) + \mathbf{D}\mathbf{x}(t-1) = \mathbf{0},$$

that is,

$$\mathbf{x}(t) = -\,[\mathbf{GB}(t+1) + \mathbf{I}_n]^{-1}\,\mathbf{D}\mathbf{x}(t-1), \tag{8}$$

provided that $\mathbf{GB}(t+1) + \mathbf{I}_n$ is a n-dimensional regular matrix.

6. Evans and Guesnerie (2005) show that $\bar{\mathbf{B}} = \mathbf{V\Lambda V}^{-1}$, where $\mathbf{\Lambda}$ is an $n \times n$ diagonal matrix whose ii^{th} entry is λ_i ($i = 1,\ldots, n$) and $\mathbf{V}$ is the associated matrix of eigenvectors. In what follows, I focus on the saddle-path case, where $|\lambda_n| < 1 < |\lambda_{n+1}|$.

Then, a perfect foresight dynamics of such matrices $\mathbf{B}(t)$ may be associated with a sequence of matrices $(\mathbf{B}(t),\ t \geq 0)$ such that:

$$\mathbf{B}(t) = -[\mathbf{GB}(t+1) + \mathbf{I}_n]^{-1}\,\mathbf{D} \Leftrightarrow [\mathbf{GB}(t+1) + \mathbf{I}_n]\,\mathbf{B}(t) + \mathbf{D} = \mathbf{0}. \qquad (9)$$

This defines the perfect foresight EGR dynamics. Its fixed points are the stationary matrices $\bar{\mathbf{B}}$ such that $\mathbf{B}(t) = \bar{\mathbf{B}}$, in all t. They are solutions of equation (7).

The determinacy of the stationary extended growth rate associated with the matrix $\bar{\mathbf{B}}$, is standardly defined as the fact that $\bar{\mathbf{B}}$ (the infinite trajectory with constant extended growth rate) is locally isolated, that is, that there does not exist a sequence $\mathbf{B}(t)$ of perfect foresight extended growth rates converging to $\bar{\mathbf{B}}$.

A sunspot equilibrium on extended growth rates, in the spirit of the previous section, is a situation in which the whole matrix $\mathbf{B}(t)$ that links $\mathbf{x}(t)$ to $\mathbf{x}(t-1)$ is perfectly correlated with sunspots. If a sunspot event is s ($s = 1, 2$) at date t, then

$$E\,[\mathbf{x}(t+1) \,|\, s] = [\pi(s, 1)\,\mathbf{B}(1) + \pi(s, 2)\,\mathbf{B}(2)]\,\mathbf{B}(s)\,\mathbf{x}(t-1)$$

and

$$\mathbf{x}(t) = -\,\{\mathbf{G}[\pi(s, 1)\,\mathbf{B}(1) + \pi(s, 2)\,\mathbf{B}(2)]\,\mathbf{B}(s) + \mathbf{D}\}\,\mathbf{x}(t-1).$$

In a sunspot equilibrium, the a priori belief that $\mathbf{B}(t) = \mathbf{B}(s)$ is self-fulfilling in all $\mathbf{x}(t-1)$, so that

$$\mathbf{B}(s) = -\,\{\mathbf{G}[\pi(s, 1)\,\mathbf{B}(1) + \pi(s, 2)\,\mathbf{B}(2)]\,\mathbf{B}(s) + \mathbf{D}\}.$$

Finally, the (virtual-time) learning dynamics associated with the IE-stability criterion are as follows. At virtual time τ of the learning process, agents believe that, in all t,

$$\mathbf{x}(t) = \mathbf{B}(\tau)\,\mathbf{x}(t-1),$$

where $\mathbf{B}(\tau)$ is the τ^{th} estimate of the n-dimensional matrix $\mathbf{B}$. Their forecasts are accordingly

$$E\,[\mathbf{x}_{t+1} \,|\, I_t] = \mathbf{B}(\tau)\,\mathbf{x}_t.$$

The actual dynamics are obtained by reintroducing forecasts into the temporary equilibrium map:

$$\mathbf{GB}(\tau)\mathbf{x}_t + \mathbf{x}_t + \mathbf{Dx}_{t-1} = \mathbf{0} \Leftrightarrow \mathbf{x}_\tau = -\ [\mathbf{GB}(\tau) + \mathbf{I}_n]^{-1}\, \mathbf{Dx}_{\tau\text{-}1}.$$

As a result, the dynamics with learning are written

$$\mathbf{B}(\tau + 1) = -\ [\mathbf{GB}(\tau) + \mathbf{I}_n]^{-1}\mathbf{D}. \tag{10}$$

A stationary EGR $\bar{\mathbf{B}}$ is a fixed point of the above dynamics. It is locally IE stable if and only if the dynamics are converging when $\mathbf{B}(0)$ is close enough to $\bar{\mathbf{B}}$.

3.3 The Dynamic Equivalence Principle

The following proposition describes the equivalence principle in one-step-forward, multidimensional linear systems with one-period memory.

Proposition 2. For a stationary EGR, the following three statements are equivalent:

1. The EGR solution is determinate in the perfect foresight extended growth rate dynamics.
2. The EGR solution is immune to sunspots, that is, there are no neighboring local sunspot equilibria on extended growth rates with finite support, as defined above.
3. The EGR solution is locally IE stable.

In particular, the saddle-path solution—which exists when the n smallest eigenvalues of $\mathbf{A}$ have modulus less than 1, with the $(n + 1)^{\text{th}}$ having modulus greater than 1—meets all these conditions.

The proposition, which is proved in Gauthier and Guesnerie (2005), has a flavor similar to that of the one-dimensional case.[7] The connection between evolutive learning and eductive learning is now more intricate, however. It is not as easy to assess the performance of adaptive learning processes in the multidimensional extended growth rates context as in the one-dimensional situation of the previous section: part 3 of proposition 1 has no counterpart here.

7. The equivalence of propositions 1 and 3 follows easily from the above definitions and sketch of analysis. The equivalence with proposition 2 is clearly plausible.

4. Eductive Learning in Dynamic Models

The discussion of eductive learning requires fleshing out the dynamic models under scrutiny with elements from game theory. In other words, the dynamic model needs to be imbedded in a dynamic game. For the sake of completeness, I present the construct proposed in Evans and Guesnerie (2003), which is based on an overlapping generations (OLG) model.

At each period t, there exists a continuum of agents, some of whom react to expectations while others use strategies that are not reactive to expectations (in an OLG context, the latter are in the last period of their lives).[8] The former are denoted ω_t and belong to a convex segment of R, endowed with Lebesgue measure $d\omega_t$. More precisely, agents ω_t have a (possibly indirect) utility function that depends on three factors: their own strategy $s(\omega_t)$; sufficient statistics on the strategies played by others, that is, $\mathbf{y}_t = \mathbf{F}(\Pi_{\omega_t}\{s(\omega_t)\},*)$, where $\mathbf{F}$, in turn, depends first on the strategies of all agents who react to expectations at time t and second on $(*)$, which here represents sufficient statistics on the strategies played by agents who do not react to expectations and includes (but is not necessarily identified with) $\mathbf{y}_{t-1}$; and the sufficient statistics for time $t+1$ as perceived at time t, —that is, $\mathbf{y}_{t+1}(\omega_t)$, which may be random—and also, now directly, the $t-1$ sufficient statistics $\mathbf{y}_{t-1}$.

I assume that the strategies played at time t can be made conditional on the equilibrium value of the t sufficient statistics $\mathbf{y}_t$. Now, let $(\bullet)$ denote both (the product of) $\mathbf{y}_{t-1}$ and the probability distribution of the random variable $\tilde{\mathbf{y}}_{t+1}(\omega_t)$ (the random subjective forecasts held by ω_t of $\mathbf{y}_{t+1}$). Let $\mathbf{G}(\omega_t, \mathbf{y}_t, \bullet)$ be the best response function of agent ω_t. Under these assumptions, the sufficient statistics for the strategies of agents who do not react to expectations is $(*) = (\mathbf{y}_{t-1}, \mathbf{y}_t)$.

The equilibrium equations at time t are written as follows:

$$\mathbf{y}_t = \mathbf{F}\left\langle \Pi_{\omega_t}\{\mathbf{G}[\omega_t, \mathbf{y}_t, \mathbf{y}_{t-1}, \tilde{\mathbf{y}}_{t+1}(\omega_t)]\}, \mathbf{y}_{t-1}, \mathbf{y}_t \right\rangle. \tag{11}$$

8. An agent in period t is different from any other agent in period t', $t' \neq t$. This means either that each agent is physically different or that the agents have strategies that are independent from period to period. In an OLG interpretation of the model, each agent lives for two periods, but only reacts to expectations in the first period of his life.

When all agents have the same point expectations, denoted $\mathbf{y}^e_{t+1}$, the equilibrium equations determine what is called the temporary equilibrium mapping:

$$\mathbf{Q}(\mathbf{y}_{t-1}, \mathbf{y}_t, \mathbf{y}^e_{t+1}) = \mathbf{y}_t - \mathbf{F}\{\Pi_{\omega_t}[\mathbf{G}(\omega_t, \mathbf{y}_t, \mathbf{y}_{t-1}, \mathbf{y}^e_{t+1})], \mathbf{y}_{t-1}, \mathbf{y}_t\}.$$

Also assuming that all $\tilde{\mathbf{y}}_{t+1}$ have a very small common support around some given $\mathbf{y}^e_{t+1}$, decision theory suggests that $\mathbf{G}$, to the first order, depends on the expectation of the random variable $\tilde{\mathbf{y}}_{t+1}(\omega_t)$, which is denoted $\mathbf{y}^e_{t+1}(\omega_t)$ (and is close to $\mathbf{y}^e_{t+1}$). Equation (11), can be linearized around any initially given situation, denoted (0), as follows:

$$\mathbf{y}_t = \mathbf{U}(0)\mathbf{y}_t + \mathbf{V}(0)\mathbf{y}_{t-1} + \int \mathbf{W}(0, \omega_t)\mathbf{y}^e_{t+1}(\omega_t)d\omega_t,$$

where $\mathbf{y}_t$, $\mathbf{y}_{t-1}$, $\mathbf{y}^e_{t+1}(\omega_t)$ now denote small deviations from the initial values of $\mathbf{y}_t$, $\mathbf{y}_{t-1}$, $\mathbf{y}_{t+1}$, and $\mathbf{U}(0)$, $\mathbf{V}(0)$, $\mathbf{W}(0, \omega_t)$ are $n \times n$ square matrices.

If such a linearization is considered only around a steady state of the system, then $\mathbf{y}_t$, $\mathbf{y}_{t-1}$, and so on will denote deviations from the steady state and $\mathbf{U}(0)$, $\mathbf{V}(0)$, $\mathbf{W}(0, \omega_t)$ are simply $\mathbf{U}$, $\mathbf{V}$, $\mathbf{W}(\omega_t)$.

Adding an invertibility assumption yields two reduced forms. First, the standard temporary equilibrium reduced form, associated with homogenous expectations $\mathbf{y}^e_{t+1}(\omega_t) = \mathbf{y}^e_{t+1}$ is

$$\mathbf{y}_t = \mathbf{B}\mathbf{y}^e_{t+1} + \mathbf{D}\mathbf{y}_{t-1}, \tag{12}$$

Second, the stochastic beliefs reduced form is

$$\mathbf{y}_t = \mathbf{D}\mathbf{y}_{t-1} + \mathbf{B}\int \mathbf{Z}(\omega_t)\mathbf{y}^e_{t+1}(\omega_t)d\omega_t, \tag{13}$$

where $\int \mathbf{Z}(\omega_t)d\omega_t = \mathbf{I}$. I use the reduced form in equation (13) to analyze eductive stability.

4.1 Eductive Stability in a One-Dimensional Setting

Based on the above analysis, it seems natural to index beliefs to growth rates. As highlighted in Evans and Guesnerie (2003), beliefs on the proximity of trajectories in the C_0 sense do not have enough grip on the agents' actions. Hence, the hypothetical common knowledge assumption to be taken into account concerns growth rates (the C_1 topology).

(Hypothetical) common knowledge assumption: The growth rate of the system is between $\lambda_1 - \varepsilon$ and $\lambda_1 + \varepsilon$.

Such an assumption about growth rates triggers a mental process that, in successful cases, progressively reinforces the initial restriction and converges toward the solution. The mental process takes into account the variety of beliefs associated with the initial restriction. Common beliefs with point expectations are then a particular case, and it is intuitively easy to guess that convergence of the general mental process under consideration implies convergence of the special process under examination when studying IE stability. This is stressed as such: IE stability is a necessary condition of eductive stability (Evans and Guesnerie, 2003). Proposition 3 then follows from the earlier equivalence theorem (proposition 1):

Proposition 3: If a constant growth rate solution is locally eductively stable or locally strongly rational then it is determinate in growth rates, is locally IE stable, is locally immune to sunspots, and attracts all reasonable evolutive learning rules.

Eductive stability is thus more demanding in general than all the previous equivalent criteria. The fact that it is strictly more demanding is shown by Evans and Guesnerie (2003), although it becomes equally demanding when some behavioral homogeneity condition is introduced.

4.2 Eductive Stability in a Multidimensional Setting

The hypothetical common knowledge assumption to be taken into account naturally has to bear on extended growth rates.

(Hypothetical) common knowledge assumption: The extended growth rate of the system $\mathbf{B}$ belongs to $\mathbf{V}(\bar{\mathbf{B}})$, where $\mathbf{V}(\bar{\mathbf{B}})$ is a neighborhood in the space of matrices (which has to be defined with respect to some distance, normally evaluated from some matrix norm).

As mentioned earlier, if common knowledge of $\mathbf{B} \in \mathbf{V}(\bar{\mathbf{B}}) \Rightarrow \mathbf{B} = \bar{\mathbf{B}}$, then the solution is locally eductively stable or locally strongly rational. As in the one-dimensional case, proposition 4 now follows from proposition 2.

Proposition 4: If a stationary extended growth rate solution is locally eductively stable or locally strongly rational, then it is determinate, locally IE stable, and locally immune to sunspots.

Again, eductive stability is more demanding, in general, than all the standard and equivalent criteria. The reason is that it takes into account the stochastic nature of beliefs and the heterogeneity of beliefs. Both dimensions are neglected explicitly in the iterative expectational stability construct and implicitly in the other equivalent constructs. In fact, as soon as local eductive stability is concerned, point expectations and stochastic expectations may not make much difference (see Guesnerie and Jara-Moroni, 2007). At least locally, the key differences between strong rationality and standard expectational stability criteria stem from the heterogeneity of expectations.

4.3 Standard Expectational Coordination Approaches and the Eductive Viewpoint: A Tentative Conclusion

My comparison of the eductive viewpoint with the standard expectational coordination criteria (determinacy, absence of neighbor sunspot equilibria, and IE-stability) has been limited to the above class of models. An exhaustive attempt would have to extend the class of models under scrutiny in different directions. First, uncertainty (intrinsic uncertainty) would have to be introduced into the models. The analysis should extend, with some technical difficulties, the appropriate objects under scrutiny being respectively probability distributions on growth rates and extended growth rates. The equivalence proposition 2 would most likely have a close counterpart in the new setting. Second, the models would need to incorporate longer memory lags or more forward-looking perceptions (or both). The theory seems applicable, although the concept of extended growth rate becomes more intricate (Gauthier, 2004).

The next set of remarks brings me back to the models used in monetary theory (starting, for example, with Sargent and Wallace, 1975). A number of these models have a structure analogous to the ones examined here, although they often involve intrinsic uncertainty. This suggests two provisional conclusions that will be put under scrutiny in the next section. First, the standard criterion used in monetary theory for assessing expectational coordination, local determinacy, is less demanding than the eductive criterion. This can be seen, within the present perspective, as the reflection of a neglect of a dimension of heterogeneity of expectations that is present in the problem.

Second, the connections between the evolutive and eductive viewpoints are less clear-cut than in the prototype model. The differences have two sources: the theoretical connection between the two types of learning is less well established in the multidimensional case, which often obtains in monetary models of the New-Keynesian type, than in the one-dimensional one (that is, proposition 1-3 has no counterpart in proposition 2); and in a noisy system, agents do not observe, at each step, a state of the system, as defined in the construct (that is, a probability distribution), but a random realization drawn from this probability distribution. Rules on learning, aimed at being efficient, have to react slowly to new information. Intuitively, IE stability and thus eductive stability will be more demanding local criteria than the success of necessarily slow evolutive learning.

However, the above analysis and its provisional conclusions implicitly refer to a true overlapping-generations framework. The equations from which the expectational coordination aspects of monetary policy are most often examined are indeed overlapping, but they come from non-OLG infinite horizon models. Their interpretation within the framework of an eductive analysis should therefore be different.

5. Eductive Stability in a Cashless Economy

The objective here is to introduce very simple versions or models that are used for the discussion of monetary policy and central bank policy. The discussion centers on a simple model of a cashless economy, in the sense of Woodford (2003).

5.1 The Model and the Standard Viewpoint

Consider an economy populated by a continuum of identical agents, who live forever. Each agent α receives $\bar{y}$ units of a perishable good in every period.[9] There is money, and the good has a money price P_t in each period,

The agents have an identical utility function:

$$U = \sum \beta^t u(C_t),$$

where $u(C_t)$ is iso-elastic

9. Although the continuum interpretation continues to hold, the reasoning formally refers to a representative consumer, leaving aside the notation α.

$$u(C_t) = \frac{1}{1-\sigma}(C_t)^{(1-\sigma)}.$$

The first-order conditions are

$$(1+i_t) = \left(\frac{1}{\beta}\right)\left[\frac{u'(C_{t+1})}{u'(C_t)}\right]\left(\frac{P_t}{P_{t+1}}\right)^{-1} = \left(\frac{1}{\beta}\right)\left(\frac{P_{t+1}}{P_t}\right)\left(\frac{C_t}{C_{t+1}}\right)^{\sigma},$$

where i_t is the nominal interest rate.

The central bank decides on a nominal interest rate according to a Wicksellian rule. The rule takes the following form:

$$i_t^m = \phi\left(\frac{P_t}{P_{t-1}}\right),$$

where ϕ is increasing.
The targeted inflation rate is $\Pi^* > \beta$, so that

$$1+\phi(\Pi^*) = \frac{\Pi^*}{\beta}.$$

The money price at time 0 is denoted P_0^*. The targeted price path is

$$P_t^* = P_0^*(\Pi^*)^t.$$

The economy is considered to start at time 1.

The path $P_t = P_t^*$, $C_t(\alpha) = \bar{y}$, $t = 1, 2,\ldots + \infty$, defines a rational expectations (here a perfect foresight) equilibrium, associated with a nominal interest rate $\phi(\Pi^*) = (\Pi^*/\beta) - 1$.

Is this equilibrium determinate? Since all agents are similar and face the same conditions in any equilibrium, any equilibrium has to meet $C_t(\alpha) = \bar{y}$. It follows that any other (perfect foresight) equilibrium $\{P_t'\}$ has to meet

$$\left[1+\phi\left(\frac{P_t'}{P_{t-1}'}\right)\right]\beta = \left(\frac{P_{t+1}'}{P_t'}\right),$$

which can be rewritten, using π_t as the inflation rate:

$$[1 + \phi(\pi_t)] = \pi_{t+1}.$$

Any equilibrium close to the stationary equilibrium Π^* would satisfy (with straightforward notation)

$$\phi'(*)\beta(\delta\pi_t) = (\delta\pi_{t+1}),$$

an equation incompatible with the proximity of the new equilibrium trajectory to the steady-state trajectory, as soon as $\phi'(*)\beta > 1$. In other words, if $\phi'(*) > (1/\beta)$, then the equilibrium is locally determinate, and this is the condition associated with the Taylor rule (see, for example, Taylor, 1999).

The argument sketched above does not demonstrate that there are no other perfect foresight equilibria outside the neighborhood under consideration, although the one under scrutiny is the only stationary one. Moreover, if the equations are viewed as coming from an OLG framework, I would argue that the equilibrium is locally IE stable, or even here locally eductively stable. Indeed, assume that (a) it is initially common knowledge that inflation will remain forever in the neighborhood of Π^*, and (b) it is common knowledge that a (general) departure of inflation expectations of $\delta\pi_{t+1}$ involves a departure of period t inflation of $\delta\pi_t = (1/\beta\phi')\delta\pi_{t+1}$. The two assertions together imply that the steady-state inflation $*$ is common knowledge. In other words, the equilibrium $*$ is locally eductively stable.[10]

However, assertion (b), which is a core element of the OLG framework, makes no full sense here, where what happens today depends not only on expectations for tomorrow, but necessarily on the whole trajectory of agents' beliefs. To put it in another way, the fact that tomorrow's (period $t + 1$) inflation expectation is π_{t+1} has no final bite on what the equilibrium price may be today in period t. Indeed, an agent's demand in period t as seen from period 1 is

$$C_t(\alpha) = C_1(\alpha)\left\{\beta^{(t-1)/\sigma}\Pi_1^{t-1}\left[(1+i_s)\left(\frac{P_s}{P_{s+1}}\right)\right]^{1/\sigma}\right\}.$$

10. Strictly speaking, the sketched argument only shows that the equilibrium $*$ is locally IE stable. The fact that agents are identical here is more than needed to ensure that heterogeneity of beliefs does not matter, so that IE stability implies eductive stability.

In period *t,* agent α may be viewed as determining its demand as follows. First, take $C_t(\alpha)$ as a starting parameter and compute the infinite sequence,

$$C_{t+\tau}(\alpha) = C_t(\alpha)\left\{\beta^{(t+\tau-1)/\sigma}\Pi_t^{t+\tau-1}\left[(1+i_s)\left(\frac{P_s}{P_{s+1}}\right)\right]^{1/\sigma}\right\}.$$

Then choose $C_t(\alpha)$ so that it meets the consumer's discounted intertemporal budget constraint.

Clearly, such a computation has to be fed by the whole agents' beliefs over the period and not only by their beliefs over the next period! In other words, the connection between t and $t + 1$ emphasized above for the analysis of eductive stability only captures one intermediate step of the choice procedure and not the whole story, as it would in a true OLG framework.

The right question is then the following: if hypothetically it is common knowledge that π_s is close to $\Pi_s^* = \Pi^*$, then is the equilibrium common knowledge? The next section addresses this question.

5.2 Eductive Stability in the Infinite Horizon Cashless Economy: Preliminaries

Consider the world at time 1 and assume that, at the margin of the stationary equilibrium, where the real interest rate is r^*, all agents expect a small departure dr_s, $s = 1,\ldots$. At this stage, it does not matter whether such a departure comes from an expected change in nominal interest rate or an expected change in inflation. Given these changes in beliefs, what is the new first-period equilibrium?

Consumption will not change in period 1. The only adjustment variable is the first period interest rate, which will become $r^* + dr_1$. What will be the equilibrium dr_1? The answer is given by lemma 1.

Lemma 1: The new equilibrium real interest rate is, to the first-order approximation, $r^* + dr_1$, with

$$dr_1 = -\left(\frac{\beta}{1-\beta}\right)(dr_2).$$

Proof: Consider the first-order conditions:

$$C_t(\alpha) = C_1(\alpha)\left[\beta^{(t-1)/\sigma}\Pi_1^{t-1}(1+r_s)^{1/\sigma}\right].$$

Take the log,

$$\log C_t = \log C_1 + \left(\frac{t-1}{\sigma}\right)\log\beta + \left(\frac{1}{\sigma}\right)\sum_{1}^{t-1}\log(1+r_s),$$

so that, approximately, in the neighborhood of the stationary equilibrium with consumption C^* and interest rate r^* and with $\beta(1 + r^*) = 1$,

$$\left(\frac{dC_t}{C*}\right) = \left(\frac{dC_1}{C*}\right) + \left(\frac{\beta}{\sigma}\right)\left(\sum_{s=1}^{t-1} dr_s\right).$$

Singling out the adjustment variable dr_1,

$$\left(\frac{dC_t}{C*}\right) = \left(\frac{dC_1}{C*}\right) + \left(\frac{\beta}{\sigma}\right)dr_1 + \left(\frac{\beta}{\sigma}\right)\left(\sum_{s=2}^{t-1} dr_s\right).$$

A key remark is that the expected price change only induces a second-order welfare change for the consumer. As is known from consumption theory, the welfare change obtains to the first-order approximation, as the inner product of the price change and of the market exchange vector (the difference between the consumption and the endowment vector).[11] Since this latter vector is zero, the result obtains. Now, the above finding implies that

$$\sum_{1}^{+\infty}\beta^{t-1}\left(\frac{dC_t}{C*}\right) = 0.$$

I next compute the above expression:

$$\sum_{1}^{+\infty}\beta^{t-1}\left(\frac{dC_t}{C*}\right) = \left(\frac{1}{1-\beta}\right)\left(\frac{dC_1}{C*}\right) + \left(\frac{1}{\sigma}\right)\left\{\sum_{2}^{+\infty}\beta^{t}\left[dr_1 + \left(\sum_{s=2}^{t-1} dr_s\right)\right]\right\}.$$

11. The fact that this is an infinite-commodity setting does not modify the part of the theory under consideration.

In the case of $dr_s = dr_2$, $\forall$s,

$$\sum_{1}^{+\infty}\beta^{t-1}\left(\frac{dC_t}{C*}\right)=\left(\frac{1}{1-\beta}\right)\left(\frac{dC_1}{C*}\right)+\left(\frac{1}{\sigma}\right)\left\{\sum_{2}^{+\infty}\beta^t\left[(dr_1)+\sum_{3}^{+\infty}(t-2)\beta^t(dr_2)\right]\right\}.$$

Because $\sum_2^{+\infty}\beta^t = \beta^2/(1-\beta)$, $\sum_3^{+\infty}(t-2)\beta^t = \beta^3/(1-\beta)^2$, this implies that

$$\left(\frac{dC_1}{C*}\right)=-\left(\frac{\beta^2}{\sigma}\right)(dr_1)-\left(\frac{\beta^3}{\sigma}\right)(1-\beta)(dr_2).$$

As in equilibrium $dC_1 = 0$, the result follows.

5.3 Eductive Stability: The Core Analysis

As explained above, I implicitly assume that both the model and rationality are common knowledge. Also the monetary rule of the central bank (ϕ) is credibly committed and hence believed. The initial common knowledge restriction has to be a hypothetical restriction on the state of the system. Here the state of the system is entirely defined, once the monetary rule is adopted, by the sequence of inflation rates. Since the equilibrium inflation rate is Π^*, a natural local restriction on beliefs is that the inflation rate is in the range of $[\Pi^* - \epsilon, \Pi^* + \epsilon]$.

Does this belief trigger a collective mental process leading to the general conclusion that $*$ will emerge? The process under discussion takes place in period 1. To illustrate this process, I explore what will happen if in period 1, all agents believe that future inflation will be for ever $\Pi^* + \epsilon$. First, the expected price path will then be $P_t' = P_1(\Pi^* + \epsilon)^{t-1}$, $t = 2,\dots + \infty$. Second, the expected real interest rate between t and $t + 1$, $t \geq 2$ will be

$$\frac{1 + \varphi(\Pi^* + \epsilon)}{\Pi^* + \epsilon};$$

that is, it will differ from r^* by approximately

$$\left[\frac{1}{(\Pi^*)^2}\right]\left[\varphi'\Pi^* - (1+\varphi)\right]\epsilon.$$

That is,

$$\left(\frac{1}{\Pi^*}\right)\left(\varphi' - \frac{1}{\beta}\right)\epsilon.$$

I assume that in period one, agents make plans contingent on the interest rate (that is, they submit a demand curve). Their conditional inference of the nominal interest rate is then $\varphi(P_1/P_0^*)$.

With regard to their inference of the next period price P_2, $P_2 = P_1(\Pi^* + \epsilon)$.[12] Hence, the expected real interest rate is

$$\frac{1 + \varphi(P_1 / P_0^*)}{\Pi^* + \epsilon};$$

that is, approximately, when writing the first-period inflation rate $(P_1/P_0^*) = (\Pi^* - \epsilon')$,

$$\left(\frac{\varphi'\epsilon'}{\Pi^*}\right) - \left[\frac{(1+\varphi)\epsilon}{(\Pi^*)^2}\right] = \left(\frac{1}{\Pi^*}\right)\left[\varphi'\epsilon' - \left(\frac{1}{\beta}\right)\epsilon\right]$$

Setting $v = \varphi'$ yields the next lemma.

Lemma 2: Under the state of beliefs just considered, the first-period inflation rate is $(\Pi^* - \epsilon')$, where

$$v\epsilon' = \left[\left(\frac{1}{\beta}\right) - \left(\frac{\beta}{1-\beta}\right)\left(v - \frac{1}{\beta}\right)\right]\epsilon$$

Proof: The above formula is applied:

$$dr_1 = -\left(\frac{\beta}{1-\beta}\right)(dr_2),$$

with

$$dr_2 = \left(\frac{1}{\Pi^*}\right)\left(\varphi' - \frac{1}{\beta}\right)\epsilon$$

12. A different assumption on beliefs would be to see the expected price path as $P_t' = P_0^*(\Pi^* + \epsilon)^t$, $t = 2, \ldots +\infty$, so that $P_2^e = P_2^*(\Pi^* + \epsilon)^2$ in period 1.

and

$$dr_1 = \left(\frac{1}{\Pi^*}\right)\left(\varphi'\epsilon' - \frac{1}{\beta}\epsilon\right)$$

If $\varphi' = v$, then

$$v\epsilon' = \left[\left(\frac{1}{\beta}\right) - \left(\frac{\beta}{1-\beta}\right)\left(v - \frac{1}{\beta}\right)\right]\epsilon.$$

This leads to my main result, as presented in the following proposition.

Proposition 5. A necessary condition for the strong rationality of the equilibrium is $(1/\beta) \leq v \leq (1/\beta)[1/(2\beta - 1)]$. Since $1 + r^* = 1/\beta$, the condition can also be written

$$(1 + r^*) \leq v \leq \frac{(1 + r^*)^2}{(1 - r^*)}.$$

Proof: For eductive stability to hold, the initial belief must not be self-defeating. For that, it must be the case that

$$-1 \leq \left(\frac{1}{\beta v}\right) - \left(\frac{\beta}{1-\beta}\right)\left(1 - \frac{1}{\beta v}\right) \leq 1.$$

Take the inequality ≤ 1. It follows that

$$\frac{(1/\beta v)(1 - \beta + \beta)}{(1 - \beta)} \leq 1 + \left(\frac{\beta}{1-\beta}\right).$$

or

$$\left(\frac{1}{\beta v}\right) \leq 1.$$

Take the inequality $-1 \leq [\ ]$. Then,

$$\frac{(1/\beta v)(1 - \beta + \beta)}{(1 - \beta)} \geq -1 + \left(\frac{\beta}{1-\beta}\right),$$

or

$$\left(\frac{1}{\beta v}\right) \geq (-1+2\beta),$$

or

$$v \leq \left(\frac{1}{\beta}\right)\left(\frac{1}{2\beta-1}\right).$$

Indeed one conjecture is that this necessary condition is sufficient, as soon as one specifies the initial set of beliefs as avoiding sweeping beliefs (that is, alternating expectations of high and low inflation). In the sense of the general discussion at the beginning of the paper, this is like choosing an appropriate topology for the neighborhood of the steady state (with sweeping beliefs being considered as non-close to the initial one).[13] The proof would consist in showing that the initial beliefs induce a smaller deviation from the targeted inflation, not only in the first period but in any period, and then iterating the argument using the common knowledge assumption.

The result is striking. The range of $v = \varphi'$ that insures eductive stability is rather small. With β close to 1, the condition looks roughly as follows:

$$\left(\frac{1}{\beta}\right) \leq v \leq \left(\frac{1}{\beta}\right)\left[1+2(1-\beta)\right].$$

For the sake of illustration, with a high $\beta = 0.95$, this is roughly

$$(1.05) \leq v \leq (1.05)(1.1) = (1.15).$$

More generally, for small r^*, the window for the reaction coefficient is, to the first-order approximation, $[1 + r^*, 1 + 2r^*]$.

The analysis thus suggests that standard Taylor rules are too reactive. Another striking, but not surprising, conclusion is that a plausible intuition within the determinacy viewpoint (that is, the equilibrium is more determinate, and in a sense more expectationally

13. This is reminiscent of the distinction between C_0 and C_1 topology discussed in section 2.

stable, whenever v increases) is plainly wrong here; there is a small window, above $1/\beta$ (and shrinking with β and vanishing when β tends to 1), for expectational stability.

6. Conclusion

Any conclusions are necessarily provisional, since an outsider's random walk in monetary models (albeit starting from a well-established base camp) has to be subjected to criticism. It must also be enriched to develop an intuition that is somewhat missing in the present state of my understanding of the specialized issues that have been addressed. This outsider's walk has, however, attempted to raise interesting questions for insiders and thus open new fronts of thinking.

References

DeCanio, S.J. 1979. "Rational Expectations and Learning from Experience." *Quarterly Journal of Economics* 93(1): 47–57.

Desgranges, G. 2000. "CK-Equilibria and Informational Efficiency in a Competitive Economy." Contributed paper 1296. Econometric Society World Congress 2000.

Desgranges, G. and S. Gauthier. 2003. "Uniqueness of Bubble-Free Solution in Linear Rational Expectations Models." *Macroeconomic Dynamics* 7: 171–91.

Desgranges, G., P.Y. Geoffard and R. Guesnerie. 2003. "Do Prices Transmit Rationally Expected Information?" *Journal of the European Economic Association* 1(1): 124–53.

Desgranges, G. and M. Heinemann. 2005. "Strongly Rational Expectations Equilibria with Endogenous Acquisition of Information." Working paper series in economics 9. University of Lüneburg, Institute of Economics.

Evans, G.W. 1985. "Expectational Stability and the Multiple Equilibria Problem in Linear Rational Expectations Models." *Quarterly Journal of Economics* 100(4): 1217–33.

Evans, G.W. and R. Guesnerie. 2003. "Coordination on Saddle Path Solutions: The Eductive Viewpoint—Linear Univariate Models." *Macroeconomic Dynamics* 7: 42–62.

————. 2005. "Coordination on Saddle-Path Solutions: The Eductive Viewpoint, Part 2—Linear Multivariate Models." *Journal of Economic Theory* 124(2): 202–29.

Evans, G.W., R. Guesnerie and B. McGough. 2007. "Eductive Stability in a RBC Setting." University of Oregon.

Evans, G.W. and S. Honkapohja. 2001. *Learning and Expectations in Macroeconomics*. Princeton University Press.

Gauthier, S. 2002. "Determinacy and Stability under Learning of Rational Expectations Equilibria." *Journal of Economic Theory* 102(2): 354–74.

————. 2003. "On the Dynamic Equivalence Principle." *Macroeconomic Dynamics* 7: 63–88.

————. 2004. "Determinacy in Linear Rational Expectations Models." *Journal of Mathematical Economics* 40(7): 815–30.

Gauthier, S. and R. Guesnerie. 2005. "Successes and Failures in Coordinating Expectations." In *Assessing Rational Expectations 2: Eductive Stability in Economics*, edited by R. Guesnerie, chap. 11. MIT Press.

Ghosal, S. 2006. "Intertemporal Coordination in Two-Period Markets." *Journal of Mathematical Economics* 43(1): 11–35.

Grandmont, J.M. and G. Laroque. 1991. "Economic Dynamics with Learning: Some Instability Examples." In *Equilibrium Theory and Applications*, edited by W.A. Barnett, B. Cornet, C. D'Aspremont, J. Gabszewicz and A. Mas-Colell, pp. 247–73. Cambridge University Press.

Guesnerie, R. 1992. "An Exploration of the Eductive Justifications of the Rational Expectations Hypothesis." *American Economic Review* 82(5): 1254–78.

———. 2001. "Short-Run Expectational Coordination: Fixed versus Flexible Wages." *Quarterly Journal of Economics* 126(3): 1115–47.

———. 2002. "Anchoring Economic Predictions in Common Knowledge." *Econometrica* 70(2): 439–80.

———. 2005. *Assessing Rational Expectations 2: Eductive Stability in Economics*. MIT Press.

Guesnerie, R. and P. Jara-Moroni. 2007. "Expectational Coordination in a Class of Economic Models: Strategic Substitutabilities versus Strategic Complementarities." Working paper 2007-28. PSE.

Guesnerie, R. and M. Woodford. 1991. "Stability of Cycles with Adaptive Learning Rules." In *Equilibrium Theory and Applications: Proceedings of the Sixth International Symposium in Economic Theory and Econometrics*, edited by W.A. Barnett, B. Cornet, C. D'Aspremont, J. Gabszewicz and A. Mas-Colell. Cambridge University Press.

Lucas, R.E., Jr. 1978. "Asset Prices in an Exchange Economy." *Econometrica* 46(6): 1429–45.

McCallum, B.T. 1983. "On Nonuniqueness in Linear Rational Expectations Models: An Attempt at Perspective." *Journal of Monetary Economics* 11(2): 139–68.

Milgrom, P. and J. Roberts. 1990. "Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities." *Econometrica* 58(6): 1255–77.

Reichlin, P. 1986. "Equilibrium Cycles in an Overlapping Generations Economy with Production." *Journal of Economic Theory* 40(1): 89–102.

Sargent, T. and N. Wallace. 1975. "Rational Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply." *Journal of Political Economy* 83(2): 251–54.

Taylor, J., ed. 1999. *Monetary Policy Rules*. University of Chicago Press.

Woodford, M. 1994. “Monetary Policy and Price-Level Determinacy in a Cash-in-Advance Economy.” *Economic Theory* 4(3): 345–80.

———. 2003. *Interest and Prices: Foundations of Theory of Monetary Policy*. Princeton University Press.

Determinacy, Learnability, and Plausibility in Monetary Policy Analysis: Additional Results

Bennett T. McCallum
Carnegie Mellon University

It is almost superfluous to begin by emphasizing that recent research in monetary policy analysis has featured a great deal of work concerning conditions for determinacy—that is, existence of a unique dynamically stable rational expectations equilibrium—under various specifications of policy behavior.[1] Indeed, there are a number of papers in which determinacy is the only criterion for a desirable monetary policy regime that is explicitly mentioned.[2]

By contrast, I have argued in recent publications (McCallum, 2003a, 2007) that least-squares (LS) learnability is a compelling necessary condition for a rational expectations (RE) equilibrium to be considered plausible, since individuals must somehow learn about the exact nature of an economy from data generated by that economy itself, while the LS learning process is biased toward a finding of learnability. A similar position has also been expressed by Bullard (2006). From such a position it follows that in conditions in which there is more than one dynamically stable RE solution—that is, indeterminacy—there may still be only one RE solution that is economically relevant, if the others are not LS learnable. In this sense, LS learnability is arguably a more important criterion than determinacy.

I am indebted to Riccardo DiCecio, George W. Evans, Christopher Kent, and Carl E. Walsh for helpful comments on earlier versions.

1. Prominent examples include Benhabib et al. (2001), Clarida, Galí, and Gertler (1999), Rotemberg and Woodford (1999), Sims (1994), and Woodford (2003). Discussion in a leading textbook is provided by Walsh (2003).
2. See, for example, Carlstrom and Fuerst (2005). These authors would almost surely include other criteria if explicitly asked.

Monetary Policy under Uncertainty and Learning, edited by Klaus Schmidt-Hebbel and Carl E. Walsh, Santiago, Chile.

It may be useful to expand briefly on the contention that LS learnability is a compelling necessary condition. The argument begins with the idea that, in actual economies, agents must ultimately obtain quantitative details concerning their economy, necessary for forming expectations, from data generated by that economy. Accordingly, the same should be true for the model economy used by a researcher. There are many conceivable learning processes, of course, so it would be rash to presume that any single one is relevant. Thus, it is not argued here that LS learnability is a sufficient condition for a RE equilibrium to be plausible. But the setup for LS learnability (see Evans and Honkapohja, 2003) is specified in a way that is, in a sense, biased towards a finding of learnability. Specifically, it assumes that agents know the correct structure qualitatively—that is, they know which variables are relevant. In addition, the process assumes that agents are collecting an ever-increasing number of observations on all relevant variables while the structure is remaining unchanged. Furthermore, the agents are estimating the relevant unknown parameters with an appropriate estimator.[3] Consequently, it seems, all in all, that if a proposed RE solution is not learnable by the LS process in question, it is implausible that it could prevail in practice.

Substantively, McCallum (2007) demonstrates that, in a very wide class of linear RE models, determinacy implies LS learnability (but not the converse) when individuals have knowledge of current conditions available for use in the learning process. This strong result does not pertain, however, if individuals have available, in the learning process, only information regarding previous values of endogenous variables.[4] One task of the present paper, accordingly, is to investigate the situation that is obtained when only lagged information is available. In addition, the paper will explore results that pertain when an alternative criterion of model plausibility, provisionally termed "well-formulated," characterizes the model's structure. In particular, it is shown that models that are well formulated, in the defined sense, often (but not invariably) possess the property of E-stability and hence LS learnability if current-period information is available in the learning process, even if determinacy does not prevail. Thus plausibility of a RE solution requires both that it be learnable and that the model at

3. A bit of additional discussion of the process is given below in section 2. Also see Evans and Honkapohja (2001, pp. 232-38).

4. Another limitation of the analysis of McCallum (2007) is that it considers only solutions of a form that excludes "resonant frequency sunspot" solutions. That limitation, which is maintained here, is discussed briefly in section 5.

hand be well formulated. A sufficient condition for both of these to hold, requiring that certain matrices have positive dominant diagonals, is introduced and considered below. Unfortunately, the situation in the case of lagged information is less favorable—that is, learnability can be assured only in special cases, for which no general characterization has been found.

1. Model and Determinacy

It will be useful to begin with a summary of the formulation and results developed in McCallum (2007). Throughout, we will work with a model of the form

$$\mathbf{y}_t = \mathbf{A}E_t\, \mathbf{y}_{t+1} + \mathbf{C}\mathbf{y}_{t-1} + \mathbf{D}\mathbf{u}_t, \tag{1}$$

where $\mathbf{y}_t$ is a $m\times 1$ vector of endogenous variables, $\mathbf{A}$ and $\mathbf{C}$ are $m\times m$ matrices of real numbers, $\mathbf{D}$ is $m\times n$, and $\mathbf{u}_t$ is a $n\times 1$ vector of exogenous variables generated by a dynamically stable process

$$\mathbf{u}_t = \mathbf{R}\mathbf{u}_{t-1} + \varepsilon_t, \tag{2}$$

with ε_t a white noise vector. It will not be assumed, even initially, that $\mathbf{A}$ is invertible. This specification is useful in part because it is the one utilized in Section 10.3 of Evans and Honkapohja (2001), for which E-stability conditions are reported on their p. 238.[5] Furthermore, the specification is very broad; in particular, any model satisfying the formulations of King and Watson (1998) or Klein (2000), can be written in this form—which will accommodate any number of lags, expectational leads, and lags of leads (see the appendix).

Following McCallum (1983, 1998), consider solutions to model (1)–(2) of the form

$$\mathbf{y}_t = \boldsymbol{\Omega}\mathbf{y}_{t-1} + \boldsymbol{\Gamma}\mathbf{u}_t, \tag{3}$$

in which $\boldsymbol{\Omega}$ is required to be real. Then, $E_t\mathbf{y}_{t+1} = \boldsymbol{\Omega}(\boldsymbol{\Omega}\mathbf{y}_{t-1} + \boldsymbol{\Gamma}\mathbf{u}_t) + \boldsymbol{\Gamma}\mathbf{R}\mathbf{u}_t$, and straightforward undetermined-coefficient reasoning shows that $\boldsymbol{\Omega}$ and $\boldsymbol{\Gamma}$ must satisfy

5. Constant terms can be included in the equations of (1) by including an exogenous variable in $\mathbf{u}_t$ that is a random walk whose innovation has variance zero. In this case there is a borderline departure from process stability.

$$\mathbf{A\Omega^2} - \mathbf{\Omega} + \mathbf{C} = 0 \tag{4}$$

$$\mathbf{\Gamma} = \mathbf{A\Omega\Gamma} + \mathbf{A\Gamma R} + \mathbf{D}. \tag{5}$$

For any given $\mathbf{\Omega}$, equation (5) yields a unique $\mathbf{\Gamma}$ generically,[6] but there are many $m \times m$ matrices that solve (4) for $\mathbf{\Omega}$. Accordingly, the following analysis centers on equation (4). Since $\mathbf{A}$ is not assumed to be invertible, we write

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{\Omega}^2 \\ \mathbf{\Omega} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{-C} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{\Omega} \\ \mathbf{I} \end{bmatrix}, \tag{6}$$

in which the first row reproduces the matrix quadratic (4). Let the $2m \times 2m$ matrices on the left- and right-hand sides of equation (6) be denoted $\bar{\mathbf{A}}$ and $\bar{\mathbf{C}}$, respectively. Then, instead of focusing on the eigenvalues of $\bar{\mathbf{A}}^{-1}\bar{\mathbf{C}}$, which does not exist when $\mathbf{A}$ is singular, we solve for the (generalized) eigenvalues of the matrix pencil $(\bar{\mathbf{C}} - \lambda\bar{\mathbf{A}})$, alternatively termed the (generalized) eigenvalues of $\bar{\mathbf{C}}$ with respect to $\bar{\mathbf{A}}$ (see, for example, Uhlig, 1999). Thus, instead of diagonalizing $\bar{\mathbf{A}}^{-1}\bar{\mathbf{C}}$, as in Blanchard and Khan (1980), we use the Schur generalized decomposition, which serves the same purpose. Specifically, the Schur generalized decomposition theorem establishes that there exist unitary matrices $\mathbf{Q}$ and $\mathbf{Z}$ such that $\mathbf{Q}\,\bar{\mathbf{C}}\,\mathbf{Z} = \mathbf{T}$ and $\mathbf{Q}\,\bar{\mathbf{A}}\,\mathbf{Z} = \mathbf{S}$ with $\mathbf{T}$ and $\mathbf{S}$ triangular.[7] Then, eigenvalues of the matrix pencil $(\bar{\mathbf{C}} - \lambda\bar{\mathbf{A}})$ are defined as t_{ii}/s_{ii}. Some of these eigenvalues may be "infinite," in the sense that some s_{ii} may equal zero. This will be the case, indeed, whenever $\mathbf{A}$ and therefore $\bar{\mathbf{A}}$ are of less than full rank, since then $\mathbf{S}$ is also singular. All of the foregoing is true for any ordering of the eigenvalues and associated columns of $\mathbf{Z}$ (and rows of $\mathbf{Q}$). For the present, let us focus on the arrangement that places the t_{ii}/s_{ii} in order of decreasing modulus.[8]

6. Generically, $\mathbf{I}\text{-}\mathbf{R}' \otimes [(\mathbf{I} - \mathbf{A\Omega})^{-1}\,\mathbf{A}]$ will be invertible, permitting solution of (5) for vec($\mathbf{\Gamma}$). Invertibility of $(\mathbf{I} - \mathbf{A\Omega})$ is discussed in section 3.

7. Provided only that there exists some λ for which $\det[\bar{\mathbf{C}} - \lambda\bar{\mathbf{A}}] \neq 0$. See Klein (2000) or Golub and Van Loan (1996, p. 377). Note that in McCallum (2007) the matrices $\bar{\mathbf{A}}$ and $\mathbf{A}$ are denoted $\mathbf{A}$ and $\mathbf{A}_{11}$, respectively.

8. The discussion proceeds as if none of the t_{ii}/s_{ii} equals 1.0 exactly. If one does, the model can be adjusted, by multiplying some relevant coefficient by (for example) 0.9999.

To begin the analysis, pre-multiply equation (6) by **Q**. Since $\mathbf{Q\bar{A}} = \mathbf{SH}$ and $\mathbf{Q\bar{C}} = \mathbf{TH}$, where $\mathbf{H} \equiv \mathbf{Z}^{-1}$, the resulting equation can be written as

$$\begin{bmatrix} \mathbf{S}_{11} & 0 \\ \mathbf{S}_{21} & \mathbf{S}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{\Omega}^2 \\ \mathbf{\Omega} \end{bmatrix} = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{0} \\ \mathbf{T}_{21} & \mathbf{T}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{H}_{11} & \mathbf{H}_{12} \\ \mathbf{H}_{21} & \mathbf{H}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{\Omega} \\ \mathbf{I} \end{bmatrix}. \tag{7}$$

The first row of equation (7) reduces to

$$\mathbf{S}_{11}(\mathbf{H}_{11}\mathbf{\Omega} + \mathbf{H}_{12})\, \mathbf{\Omega} = \mathbf{T}_{11}(\mathbf{H}_{11}\mathbf{\Omega} + \mathbf{H}_{12}). \tag{8}$$

Then, if $\mathbf{H}_{11}$ is invertible, the latter can be used to solve for $\mathbf{\Omega}$ as

$$\mathbf{\Omega} = -\mathbf{H}_{11}{}^{-1}\, \mathbf{H}_{12} = -\mathbf{H}_{11}{}^{-1}(-\mathbf{H}_{11}\, \mathbf{Z}_{12}\, \mathbf{Z}_{22}{}^{-1}) = \mathbf{Z}_{12}\, \mathbf{Z}_{22}{}^{-1}, \tag{9}$$

where the second equality comes from the upper-right-hand submatrix of the identity $\mathbf{HZ} = \mathbf{I}$, provided that $\mathbf{H}_{11}$ is invertible, which is assumed without significant loss of generality.[9],[10]

As mentioned above, there are many solutions $\mathbf{\Omega}$ to equation (4). These correspond to different arrangements of the eigenvalues, which result in different groupings of the columns of **Z** and therefore different compositions of the submatrices $\mathbf{Z}_{12}$ and $\mathbf{Z}_{22}$. Here, with the eigenvalues t_{ii}/s_{ii} arranged in order of decreasing modulus, the diagonal elements of $\mathbf{S}_{22}$ will all be non-zero, provided that **S** has at least m non-zero eigenvalues, which is assumed to be the case.[11] Clearly, for any solution under consideration to be dynamically stable, the eigenvalues of $\mathbf{\Omega}$ must be smaller than 1.0 in modulus. In McCallum (2007) it is shown that

$$\mathbf{\Omega} = \mathbf{Z}_{22}\, \mathbf{S}_{22}{}^{-1}\, \mathbf{T}_{22}\, \mathbf{Z}_{22}{}^{-1}, \tag{10}$$

9. This invertibility condition, also required by King and Watson (1998) and Klein (2000), obtains except in degenerate special cases of equation (1) that can be solved by simpler methods than considered here. Note that the invertibility of $\mathbf{H}_{11}$ implies the invertibility of $\mathbf{Z}_{22}$, given that **Z** and **H** are unitary.

10. Note that it is not being claimed that all solutions are of the form (9).

11. From its structure it is obvious that $\mathbf{\bar{A}}$ has at least m nonzero eigenvalues so, since **Q** and **Z** are nonsingular, **S** must have rank of at least m. This necessary condition is not sufficient for **S** to have at least m nonzero eigenvalues, however; hence the assumption.

so $\mathbf{\Omega}$ has the same eigenvalues as $\mathbf{S}_{22}{}^{-1}\ \mathbf{T}_{22}$. The latter is triangular, moreover, so the relevant eigenvalues are the m smallest of the $2m$ ratios t_{ii}/s_{ii} (given the decreasing-modulus ordering). For dynamic stability, the modulus of each of these ratios must then be less than 1. (In many cases, some of the m smallest moduli will equal zero.)

Let us henceforth refer to the solution under the decreasing-modulus ordering as the MOD solution. Now suppose that the MOD solution is stable. For it to be the only stable solution, there must be no other arrangement of the t_{ii}/s_{ii} that would result in a $\mathbf{\Omega}$ matrix with all eigenvalues smaller in modulus than 1.0. Thus, each of the t_{ii}/s_{ii} for $i = 1,\ldots, m$ must have modulus greater than 1.0, some perhaps infinite. Is there some $m \times m$ matrix whose eigenvalues relate cleanly to these ratios? Yes, it is the matrix $\mathbf{F} \equiv (\mathbf{I} - \mathbf{A\Omega})^{-1}\mathbf{A}$, which appears frequently in the analysis of Binder and Pesaran (1995, 1997).[12] Regarding this $\mathbf{F}$ matrix, it is shown that, for any ordering such that $\mathbf{H}_{11}$ is invertible, including the MOD ordering, we have the equality

$$\mathbf{H}_{11}\mathbf{F}\ \mathbf{H}_{11}{}^{-1} = \mathbf{T}_{11}{}^{-1}\mathbf{S}_{11}, \tag{11}$$

which implies that $\mathbf{F}$ has the same eigenvalues as $\mathbf{T}_{11}{}^{-1}\mathbf{S}_{11}$. In other words, it is the case that the eigenvalues of $\mathbf{F}$ are the same, for any given arrangement of the system's eigenvalues, as the *inverses* of the values of t_{ii}/s_{ii} for $i = 1, \ldots, m$. Under the MOD ordering, these are the inverses of the first (largest) m of the eigenvalues of the system's matrix pencil. Accordingly, for solution (9) to be the only stable solution, all the eigenvalues of the corresponding $\mathbf{F}$ must be smaller than 1.0 in modulus. This result, stated in different ways, is well known from Binder and Pesaran (1995), King and Watson (1998), and Klein (2000), and is an important generalization of one result of Blanchard and Khan (1980) for a model with nonsingular $\mathbf{A}$.

Thus we have established notation for models of form (1)–(2) and have reported results showing that the existence of a unique stable solution requires that all eigenvalues of the defined $\mathbf{\Omega}$ matrix and the corresponding $\mathbf{F}$ be less than 1.0 in modulus. It will be convenient to express that condition as follows: all $|\lambda_{\mathbf{\Omega}}| < 1$ and all $|\lambda_{\mathbf{F}}| < 1$.

12. There is no general proof of invertibility of $[\mathbf{I} - \mathbf{A\Omega}]$, but if $\mathbf{A\Omega}$ were by chance to have some eigenvalue exactly equal to 1.0, that condition could be eliminated by making some small adjustment to elements of $\mathbf{A}$ or $\mathbf{C}$. Also, see section 4 below.

2. E-Stability in Two Cases

Let us now turn to conditions for learnability under two different information assumptions. First we will review the main results from McCallum (2007), which assumes that agents have full information on current values of endogenous variables during the learning process, and then we will go on to the second assumption, namely, that only lagged values of endogenous variables are known during the learning process. The manner in which learning takes place in Evans and Honkapohja's analysis is as follows. Agents are assumed to know the structure of the economy as specified in equations (1) and (2), in the sense that they know what variables are included, but do not know the numerical values of the parameters. What they need to know, to form expectations, is values of the parameters of the solution equations (3). In each period *t*, they form forecasts on the basis of a least squares regression of the variables in $\mathbf{y}_{t-1}$ on previous values of $\mathbf{y}_{t-2}$ and any exogenous observables. Given those regression estimates, however, expectations of $\mathbf{y}_{t+1}$ may be calculated assuming knowledge of $\mathbf{y}_t$ or, alternatively, assuming that $\mathbf{y}_{t-1}$ is the most recent observation possessed by agents and is thus usable in the forecasting process. In the former case, the conditions for E-stability reported by Evans and Honkapohja (2001) are that the following three matrices must have all eigenvalues with real parts less than 1.0:

$$\mathbf{F} \equiv (\mathbf{I} - \mathbf{A\Omega})^{-1}\mathbf{A}, \tag{12a}$$

$$\left[(\mathbf{I}\text{-}\mathbf{A\Omega})^{-1}\,\mathbf{C}\right]' \otimes \mathbf{F}, \tag{12b}$$

$$\mathbf{R}' \otimes \mathbf{F}. \tag{12c}$$

In the second case, however, the analogous condition (Evans and Honkapohja, 2001) is that the following matrices must have all eigenvalues with real parts less than 1.0:

$$\mathbf{A}\,(\mathbf{I} + \mathbf{\Omega}), \tag{13a}$$

$$\mathbf{\Omega}' \otimes \mathbf{A} + \mathbf{I} \otimes \mathbf{A\Omega}, \tag{13b}$$

$$\mathbf{R}' \otimes \mathbf{A} + \mathbf{I} \otimes \mathbf{A\Omega}. \tag{13c}$$

Except in the case that $\mathbf{\Omega} = \mathbf{0}$, which will result when $\mathbf{C} = \mathbf{0}$, these conditions are not equivalent to those in equation (12).

It is important to note that use of the first information assumption is not inconsistent with a model specification in which supply and demand decisions in period t are based on expectations formed in the past, such as $E_{t-1}\mathbf{y}_{t+j}$ or $E_{t-2}\mathbf{y}_{t+j}$. It might also be mentioned parenthetically that conditions (12) and (13) literally pertain to the E-stability of the model (1)–(2) under the two information assumptions, not its learnability. Under quite broad conditions, however, E-stability is necessary and sufficient for LS learnability. This near-equivalence is referred to by Evans and Honkapohja as the "E-stability principle" (Evans and Honkapohja, 1999, 2001). Since E-stability is technically easier to verify, applied analysis typically focuses on it, rather than on direct exploration of learnability.

Given the foregoing discussion, it is a simple matter to verify that if a model of form (1)–(2) is determinate, then it satisfies conditions (12). First, determinacy requires that all eigenvalues of $\mathbf{F}$ have modulus less than 1.0, so their real parts must all be less than 1.0, thereby satisfying (12a). Second, from equation (4) it can be seen that $(\mathbf{I}–\mathbf{A\Omega})^{-1}\mathbf{C} = \mathbf{\Omega}$. Therefore, matrix (12b) can be written as $\mathbf{\Omega}' \otimes \mathbf{F}$. Furthermore, it is a standard result (Magnus and Neudecker, 1988) that the eigenvalues of a Kronecker product are the products of the eigenvalues of the relevant matrices (for example, the eigenvalues of $\mathbf{\Omega}' \otimes \mathbf{F}$ are the products $\lambda_{\mathbf{\Omega}}\, \lambda_{\mathbf{F}}$). Therefore, condition (12b) holds. Finally, since $|\lambda_{\mathbf{F}}| < 1$, condition (12c) holds provided that all $|\lambda_{\mathbf{R}}| \leq 1$, which has been assumed by specifying that equation (2) is dynamically stable.

Determinacy does not imply learnability, however, under the second information assumption. This point, which is developed by Evans and Honkapohja (2001), can be illustrated by means of a bivariate example.[13] Let the $\mathbf{y}_t$ vector in equation (1) include two variables, y_{1t} and y_{2t}, related by the dynamic model that follows:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} -0.01 & 0.01 \\ 0.99 & -0.01 \end{bmatrix} \begin{bmatrix} E_t y_{1t+1} \\ E_t y_{2t+1} \end{bmatrix} + \begin{bmatrix} 0.02 & 1.10 \\ 0.01 & 0.06 \end{bmatrix} \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \tag{14}$$

13. Its specification is close numerically to the qualitative version of the Evans and Honkapohja example that is used in McCallum (2007), pp. 1386–88.

Then, for the MOD solution we have

$$\mathbf{A\Omega} = \begin{bmatrix} -0.01 & 0.01 \\ 0.99 & -0.01 \end{bmatrix} \begin{bmatrix} 0.0218 & 1.1133 \\ -0.095 & -0.774 \end{bmatrix} = \begin{bmatrix} -0.0012 & -0.0189 \\ 0.0225 & 1.1099 \end{bmatrix}, \quad (15)$$

with eigenvalues of $\mathbf{\Omega}$ being –0.148 and –0.604, while

$$\mathrm{F} = \begin{bmatrix} 0.1604 & 0.00831 \\ -9.040 & 0.0893 \end{bmatrix},$$

which has (complex) eigenvalues $0.1249 \pm 0.2717\, i$. Inspection of these shows that this solution is determinate, and that conditions (12a) and (12b), relevant for E-stability in the case in which current information is available during learning, are satisfied. Let us assume $\mathbf{R} = \mathbf{0}$, that is, white noise disturbances, for simplicity. Then the determinate RE solution is E-stable and learnable under the first information assumption.

But for the case with only lagged information during learning, it is necessary to consider the eigenvalues of the matrices shown in expressions (13). For equation (13a), the matrix $\mathbf{A}(\mathbf{I} + \mathbf{\Omega})$ is

$$\begin{bmatrix} -0.0112 & -0.0089 \\ 1.0125 & 1.0999 \end{bmatrix}$$

whose eigenvalues are -0.0030 and 1.0918. The last of these violates the condition for equation (13a), however, so under the lagged-information assumption, the relevant E-stability condition is not satisfied and the determinate RE equilibrium is not LS learnable.

This result exemplifies the fact that determinacy is not generally sufficient for learnability of RE solutions, although it is sufficient under the first information assumption. Of equal importance, in my opinion, is the fact that determinacy is not necessary for learnability. In particular, the MOD solution can be learnable, and be the only learnable solution of form (3), in cases in which indeterminacy prevails. One such example is given in McCallum (2007).[14] In such cases, the

14. I take this opportunity to point out that McCallum (2007, p. 1386), errs in stating that when the eigenvalues are … “30.65, −0.532, −0.123, and 0.000 … both stable solutions are learnable.” Actually, only the MOD solution is learnable.

position that learnability is necessary for a solution to be plausible would suggest that there may be no problem implied by the absence of determinacy.[15]

3. Well-Formulated Models

McCallum (2003b) suggests that there is a distinct and neglected property that dynamic models should possess to be considered "well-formulated" and plausible for the purposes of economic analysis. To begin the discussion, consider first the single-variable case of specification (1),

$$y_t = \alpha E_t y_{t+1} + c y_{t-1} + u_t, \tag{16}$$

with $u_t = (1 - \rho)\eta + \rho u_{t-1} + w_{t,}$ with $|\rho| < 1$ and w_t white noise. Thus, u_t is an exogenous forcing variable with an unconditional mean of η (assumed nonzero) and units have been chosen so that there is no constant term. Applying the unconditional expectation operator to equation (16) yields

$$Ey_t = \alpha Ey_{t+1} + cEy_{t-1} + \eta. \tag{17}$$

In this case, y_t will be covariance stationary, and we have

$$Ey_t = \frac{\eta}{[1-(a+c)]}. \tag{18}$$

But from the latter, it is clear that as $a + c$ approaches 1.0 from above, the unconditional mean of y_t approaches $-\infty$ (assuming, without loss of generality, that $\eta > 0$), whereas if $a + c$ approaches 1.0 from below, the unconditional mean approaches $+\infty$. Thus, there is an infinite discontinuity at $a + c = 1.0$. This implies that a tiny change in $a + c$ could alter the average (that is, steady–state) value Ey_t from an arbitrarily large positive number to an arbitrarily large negative number. Such a property seems highly implausible and therefore unacceptable for a well-formulated model.[16] The

15. Disregarding, that is, "sunspot" solutions not of form (3).

16. The model could be formulated with the exogenous variable also written in terms of percent or fractional deviations from the reference level η, for example, $\hat{u}_t = u_t - \eta$. But that would not alter the relationship between Ey_t and η, which can be extremely sensitive to tiny changes in $a + c$.

substantive problem is not eliminated, obviously, by adoption of the zero-measure exclusion $a + c \neq 1$.

In light of the foregoing observation, it is my contention that, to be considered well formulated (WF), the model at hand needs to include a restriction on its admissible parameter values; a restriction that rules out $a + c = 1$ and yet admits a large interval of values that includes $(a,c) = (0,0)$. In the case at hand, the appropriate restriction is $a + c < 1$. Of course, $a + c > 1$ would serve just as well mathematically to avoid the infinite discontinuity, but it seems clear that $a + c < 1$ is vastly more appropriate from an economic perspective since it includes the values (0,0)[17] Since we want this condition to apply to $a + c$ sums between zero and that value that pertains to the model at hand, our requirement for WF is that a and c satisfy $1 - \varepsilon\,(a + c) > 0$ for all $0 \leq \varepsilon \leq 1$. [It should be clear, in addition, that the foregoing argument could be easily modified to apply to y_t processes that are trend stationary, rather than strictly (covariance) stationary.] It is shown in McCallum (2003b) that under this requirement, plus a second one to be discussed shortly, the univariate model (16) is invariably E-stable.[18]

Next, for the bivariate case of model (1), extension of the foregoing WF property requires that $\mathbf{A}$ and $\mathbf{C}$ be such that $\det[\mathbf{I} - \varepsilon(\mathbf{A} + \mathbf{C})]$ is positive for all $0 \leq \varepsilon \leq 1$; otherwise, the steady-state values of the variables may possess infinite discontinuities. But there are other requirements as well. Let ac_{ij} temporarily denote the ij^{th} element of $\mathbf{A} + \mathbf{C}$. Then the model with $y_1 = Ey_{1t}$, $y_2 = Ey_{2t}$, $\eta_1 = Eu_{1t}$ and $\eta_2 = Eu_{2t}$ implies

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} ac_{11} & ac_{12} \\ ac_{21} & ac_{22} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \tag{19}$$

so that $E\mathbf{y} = [\mathbf{I} - (\mathbf{A} + \mathbf{C})]^{-1}\,\boldsymbol{\eta}$ can be written as

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \frac{1}{\Delta} \begin{bmatrix} 1 - ac_{22} & ac_{12} \\ ac_{21} & 1 - ac_{11} \end{bmatrix} \begin{bmatrix} \eta_1 \\ \eta_2 \end{bmatrix} \tag{20}$$

17. In models of the linear form (16), one would expect coefficients a and c typically to represent elasticities and often to be numerically small relative to 1.

18. That paper's analysis of multivariate systems is, however, unsatisfactory.

where $\Delta = \det[\mathbf{I} - (\mathbf{A} + \mathbf{C})] = (1 - ac_{11})(1 - ac_{22}) - ac_{12}\, ac_{21}$. Then the counterpart of the univariate requirement that $1 - \varepsilon\,(a + c) > 0$ includes the condition $\Delta > 0$ [for all $0 \leq \varepsilon \leq 1$].[19] We must rule out, however, the case in which $\Delta > 0$ results from $1 - \varepsilon ac_{11}$ and $1 - \varepsilon ac_{22}$ both being negative.[20] The condition on Δ should be extended, therefore, to also require $1 - \varepsilon ac_{11} > 0$ and $1 - \varepsilon ac_{22} > 0$.

How are these WF requirements extended to pertain to cases with more than two variables? It appears that the appropriate requirement is that $[\mathbf{I} - \varepsilon(\mathbf{A} + \mathbf{C})]$ be a P-matrix, which by definition has all its principal minors positive and thereby imposes the conditions discussed for the cases above in which m equals 1 and 2. Other properties of any P-matrix are that its inverse exists and is itself a P-matrix, and that all its real eigenvalues are positive.[21]

An alternative possibility that is of interest would be to require $[\mathbf{I} - \varepsilon(\mathbf{A} + \mathbf{C})]$ to be a positive dominant-diagonal matrix.[22] This requirement would have implications for the E-stability status of the model, as will be discussed below, and positive dominant-diagonal (PDD) matrices have an important tradition in dynamic economics stemming from the literature on multimarket stability analysis. This condition is, however, somewhat stronger than is actually required by our objective of ruling out specifications in which leading implications of the model are hyper-sensitive to parameter values.

As a brief but relevant digression, one example of a matrix that is a P-matrix and yet is not positive dominant-diagonal is as follows:

$$\begin{bmatrix} 0.08 & -0.92 & 0.90 \\ 0.92 & 0.07 & -0.03 \\ -0.72 & 0.30 & 0.04 \end{bmatrix}. \tag{21}$$

Clearly, the entries in any row show immediately that this matrix is not positive dominant diagonal (PDD). But its determinant is 0.3087 and the three second-order minors are 0.0118, 0.651, and 0.852. Since the diagonal elements are also all positive, the matrix

19. Henceforth the bracketed condition is to be understood wherever relevant.
20. This is clear for the case in which $\mathbf{A} + \mathbf{C}$ is a diagonal matrix.
21. On the topic of P-matrices, see Horn and Johnson (1991) and Gale and Nikaido (1965).
22. Again, see Horn and Johnson (1991) and Gale and Nikaido (1965).

is a P-matrix. For future reference, note that its eigenvalues are –0.0067 + 1.2319i, –0.0067 – 1.2319i, and 0.2034. Thus the example illustrates the fact that, although a P-matrix cannot have a negative real eigenvalue, it can have a complex eigenvalue pair with negative real parts.[23]

Returning now to the main line of argument, there is a second type of discontinuity that should also be eliminated for a model to be viewed as WF, namely, infinite discontinuities in its impulse response functions. In model (1)–(2) with solution (3), the impulse response to the shock vector $\mathbf{u}_t$ involves the matrix $\mathbf{\Gamma}$, which is given by

$$\mathbf{\Gamma} = \mathbf{A\Omega\Gamma} + \mathbf{A\Gamma R} + \mathbf{D}. \tag{22}$$

Thus, $(\mathbf{I} - \mathbf{A\Omega})\,\mathbf{\Gamma} = \mathbf{A\Gamma R} + \mathbf{D}$ so using $\mathbf{F} = (\mathbf{I} - \mathbf{A\Omega})^{-1}\mathbf{A}$, equation (22) can be written as

$$\mathbf{\Gamma} = \mathbf{F\Gamma R} + (\mathbf{I} - \mathbf{A\Omega})^{-1}\mathbf{D}. \tag{23}$$

Then, using the well-known identity that, for any conformable matrix product $\mathbf{ABC}$ it is true that vec $\mathbf{ABC} = (\mathbf{C}' \otimes \mathbf{A})$ vec $\mathbf{B}$,[24] it follows that

$$\text{vec}\mathbf{\Gamma} = (\mathbf{R}' \otimes \mathbf{F})\text{vec}\mathbf{\Gamma} + \text{vec}\left[(\mathbf{I} - \mathbf{A\Omega})^{-1}\mathbf{D}\right] \tag{24}$$

implying

$$\text{vec}\mathbf{\Gamma} = \left[\mathbf{I} - (\mathbf{R}' \otimes \mathbf{F})\right]^{-1}\text{vec}\left[(\mathbf{I} - \mathbf{A\Omega})^{-1}\mathbf{D}\right]. \tag{25}$$

Accordingly, our second WF requirement is for $[\mathbf{I} - (\mathbf{R}' \otimes \mathbf{F})]$ and $(\mathbf{I} - \mathbf{A\Omega})$ to be well behaved in the same manner as $\mathbf{I} - (\mathbf{A} + \mathbf{C})$, that is, that each is a P-matrix. Again it is of interest to consider the possibility of requiring that each of these be a PDD matrix.

23. See Horn and Johnson (1991, p. 123).

24. See, for example, Evans and Honkapohja (2001, p. 117) or Magnus and Neudecker (1988, p. 28).

4. E-Stability in WF Models?

In this section, the concern is with the relationship between models that are WF and those in which the MOD solution is learnable. That there may be some significant relationship is suggested by the following identity:

$$(\mathbf{I} - \mathbf{A\Omega})(\mathbf{I} - \mathbf{F})(\mathbf{I} - \mathbf{\Omega}) = \mathbf{I} - (\mathbf{A} + \mathbf{C}), \tag{26}$$

which is mentioned by Binder and Pesaran (1995).[25] From this equation, it is clear that that non-singularity of $\mathbf{I} - (\mathbf{A} + \mathbf{C})$ implies that the three matrices $(\mathbf{I} - \mathbf{A\Omega})$, $(\mathbf{I} - \mathbf{F})$ and $(\mathbf{I} - \mathbf{\Omega})$ are all nonsingular. In addition, we can see that the WF requirement that $\det[\mathbf{I} - \varepsilon(\mathbf{A} + \mathbf{C})]$ is positive for all $0 \leq \varepsilon \leq 1$ also implies that the real eigenvalues of $\mathbf{\Omega}$, $\mathbf{A\Omega}$, and $\mathbf{F}$ must all be less than 1.0 in value.[26] To make that argument, consider the situation when $\mathbf{A}$ and $\mathbf{C}$ are multiplied by ε, $0 \leq \varepsilon \leq 1$. For very small values of ε, the matrices $\mathbf{\Omega}$, $\mathbf{A\Omega}$, and $\mathbf{F}$ will all be small so the eigenvalues of all four matrices in equation (26) will be close to 1.0 and their determinants will be positive. Now let ε increase and approach 1.0. If $\mathbf{I} - \varepsilon(\mathbf{A} + \mathbf{C})$ remains nonsingular throughout this process, so too will each of the three matrices on the left-hand side of equation (26). Since a real eigenvalue of zero would imply singularity for any of the matrices in question, and since eigenvalues are continuous functions of the matrix elements, the stated result is valid.

Accordingly, the WF requirement that $\det[\mathbf{I} - \varepsilon(\mathbf{A} + \mathbf{C})]$ is positive for all $0 \leq \varepsilon \leq 1$ also implies that the real eigenvalues of $\mathbf{\Omega}$, $\mathbf{A\Omega}$, and $\mathbf{F}$ are all less than 1.0 in value. In addition, the requirement that the matrix $[\mathbf{I} - (\mathbf{R}' \otimes \mathbf{F})]$ be a P-matrix implies that all the real eigenvalues of $(\mathbf{R}' \otimes \mathbf{F})$ will be smaller than 1.0. Therefore, condition (12c), as well as (12a), is satisfied. What about the remaining condition, for the current-information case, (12b)? Here we recognize that, by rearrangement of equation (4), $(\mathbf{I} - \mathbf{A\Omega})^{-1}\mathbf{C} = \mathbf{\Omega}$. Accordingly, condition (12b) becomes $\mathbf{\Omega}' \otimes \mathbf{F}$. But then note that with the MOD ordering it is the case that all $|\lambda_{\mathbf{\Omega}}| < 1/|\lambda_{\mathbf{F}}|$ so all $|\lambda_{\mathbf{\Omega}}||\lambda_{\mathbf{F}}| < 1$. But $|\lambda_{\mathbf{\Omega}}||\lambda_{\mathbf{F}}| = |\lambda_{\mathbf{\Omega}}\lambda_{\mathbf{F}}| \geq \mathrm{Re}(\lambda_{\mathbf{\Omega}}\lambda_{\mathbf{F}})$ so it follows that this condition is invariably satisfied. Accordingly,

25. The identity can be verified by writing out $\mathbf{F}$ in the left side of equation (26), multiplying, cancelling, and inserting $\mathbf{C}$ for $\mathbf{\Omega} - \mathbf{A\Omega}^2$.

26. Here, and often in what follows, I use the fact that the eigenvalues of a matrix of form $(\mathbf{I} - \mathbf{B})$ satisfy $\lambda_{\mathbf{I-B}} = 1 - \lambda_{\mathbf{B}}$.

with current information available during the learning process, the MOD solution would be learnable, when the model is WF, if all eigenvalues were real.

Unfortunately, there is no guarantee that the real part of all complex eigenvalues will be smaller than 1.0. The situation is described by Horn and Johnson (1991) as follows: "if **A** is a n–by–n P-matrix ... then every eigenvalue of **A** lies in the open angular wedge $\mathbf{W}_n \equiv \{z = re^{i\theta}: |\theta| < \pi - (\pi/n), r > 0\}$. Moreover, every point in $\mathbf{W}_n$ is an eigenvalue of some n-by-n P-matrix." But for $n > 2$, $\mathbf{W}_n$ includes points in the in the two left-hand quadrants in the complex plane. Therefore, it cannot be argued that, in general, the WF condition implies LS learnability for the MOD solution.

In this regard, note that, since **A** and **C** are matrices of real numbers, $\mathbf{I} - (\mathbf{A} + \mathbf{C})$ will have only real eigenvalues if $\mathbf{A} + \mathbf{C}$ is symmetric. And since eigenvalues are continuous functions of the elements of the matrix in question, these eigenvalues will be real if $\mathbf{A} + \mathbf{C}$ does not depart too far from symmetry. Diagonal matrices are of course symmetric, so it is not surprising that dominant-diagonal matrices have strong properties pertaining to their eigenvalues. In particular, if a real matrix is positive diagonal dominant (PDD), that is, is diagonal-dominant with all diagonal elements positive, then all its eigenvalues will have positive real parts—see Horn and Johnson (1985). Accordingly, if we were to require (as mentioned above) that $\mathbf{I} - (\mathbf{A} + \mathbf{C})$, $(\mathbf{I} - \mathbf{A}\boldsymbol{\Omega})$, and $[\mathbf{I} - (\mathbf{R}' \otimes \mathbf{F})]$ were PDD, rather than just P-matrices, then learnability would be implied. That possibility is not, however, justified by the line of argument used to motivate the WF condition, that is, by the desirability of ruling out infinite discontinuities in impulse response functions (and the model's steady-state values).

The argument, then, is that being WF is an additional, distinct, plausibility condition to be required along with learnability. Only if a RE solution is both learnable, and results from a model that is WF, would it be considered as a plausible candidate for a RE solution that might prevail in reality. This may seem like a rather demanding requirement. But most realistic models utilized in monetary policy analysis easily meet both of these conditions; difficulties arise primarily in the case of zero-lower-bound situations, very strong policy responses to expected future conditions, and other extreme conditions.

In any case, the potential attractiveness of the WF requirement, in addition to that of LS learnability, is exemplified by an example considered for other purposes in McCallum (2004). The example in table 2 of that paper combines two univariate models of form

(1)–(2), one of which has two explosive solutions and the other of which has two stable solutions.[27] Small off-diagonal elements of the **A** and **C** matrices are added to make the combined model a bivariate example that is not reducible (while barely changing the system eigenvalues). In this bivariate model it is found that there is a unique stable solution.[28] Under the current-information assumption, then, this equilibrium is learnable as well as determinate. It hardly seems plausible, however, to believe that the combination of an explosive sector plus an indeterminate sector, with only minimal interaction between them, would result in overall behavior reflecting a well-behaved, unique equilibrium. Thus the finding that the determinate and learnable solution pertains to a model that is not well-formulated, is highly relevant and leads to a conclusion that seems entirely sensible.[29] The appropriate conclusion is that this solution is not plausible. The other solution (of form (3)) is the MSV solution. It is learnable but not dynamically stable.[30] Thus the conclusion of an analysis based on the requirement that a plausible RE equilibrium must be stable, learnable, and WF is that the system under discussion has no such equilibrium. That seems eminently sensible, for a model that is the combination of one explosive sector and one indeterminate sector with very little interaction.

Next we consider learnability for WF models under the second information assumption, for which the relevant conditions are that all eigenvalues of the matrices in conditions (13a)–(13c) have real parts less than 1.0. Let us assume that $\mathbf{I} - (\mathbf{A} + \mathbf{C})$, $(\mathbf{I} - \mathbf{A}\boldsymbol{\Omega})$ and $[\mathbf{I} - (\mathbf{R}' \otimes \mathbf{F})]$ are all PDD matrices, which makes the MOD solution both learnable and WF. First consider condition (13a), which implies that $\mathbf{I} - \mathbf{A}(\mathbf{I} + \boldsymbol{\Omega})$ must have all eigenvalues with real parts that are positive. Using the definition of **F**, we can write

$$(\mathbf{I} - \mathbf{A}\boldsymbol{\Omega})(\mathbf{I} - \mathbf{F}) = (\mathbf{I} - \mathbf{A}\boldsymbol{\Omega})\,[\mathbf{I} - (\mathbf{I} - \mathbf{A}\boldsymbol{\Omega})^{-1}\mathbf{A}] = (\mathbf{I} - \mathbf{A}\boldsymbol{\Omega}) - \mathbf{A} = \mathbf{I} - \mathbf{A}(\mathbf{I} + \boldsymbol{\Omega}). \quad (27)$$

27. Incidentally, in that paper's equation (29), the lower-left element of **C** is 0.3, not 0.5.

28. Which differs from the minimum-state-variable (MSV) solution in the sense of McCallum (2003b).

29. The non-WF conclusion is based on violations of both steady-state and impulse response requirements. For the other solution of form (3), the steady-state WF conditions are violated.

30. For learning of explosive solutions, a modified condition pertaining to shock variances is required. See Evans and Honkapohja (2001, pp. 219–20).

Now, our discussion above indicates that $\mathbf{I} - \mathbf{A\Omega}$ and $\mathbf{I} - \mathbf{F}$ will both have eigenvalues with all real parts positive under the WF assumption, so equation (27) indicates that this property would carry over to $\mathbf{I} - \mathbf{A}(\mathbf{I} + \mathbf{\Omega})$. This would not be the case, however, if the only specification is that $\mathbf{I} - (\mathbf{A} + \mathbf{C})$, $(\mathbf{I} - \mathbf{A\Omega})$ and $[\mathbf{I} - (\mathbf{R}' \otimes \mathbf{F})]$ are P-matrices.

Even in the more favorable case, with PDD matrices, no general results pertaining to conditions (13b) and (13c) have been found. The problem is that sums of Kronecker products do not in general yield matrices for which eigenvalues are cleanly related to those of the individual matrices. Nevertheless, there are two special cases that can be treated readily. First, consider the case in which $\mathbf{C} = \mathbf{0}$, so there are no predetermined variables in the solution, which implies that $\mathbf{\Omega} = \mathbf{0}$. Then, $\mathbf{F} = (\mathbf{I} - \mathbf{A\Omega})^{-1}\mathbf{A} = \mathbf{A}$, and thus condition (13a) becomes the same as (12a). Furthermore, (13b) is irrelevant with $\mathbf{\Omega} = \mathbf{0}$ and (13c) becomes $(\mathbf{R}' \otimes \mathbf{A})$, which is the same as in (12c). So in this case, the two information assumptions yield the same E-stability conditions. Second, suppose that $\mathbf{C} \neq \mathbf{0}$, but that the exogenous variables are white noise, that is, $\mathbf{R} = \mathbf{0}$. Then condition (13c) becomes $(\mathbf{I} \otimes \mathbf{A\Omega})$ and the result based on $(\mathbf{I} - \mathbf{A\Omega})^{-1}$ shows that this condition will be satisfied if the latter matrix is PDD. But conditions pertaining to (13a) and (13b) are not necessarily satisfied. Of course, it is a simple matter to examine specific cases numerically.

5. General Issues

A number of possible objections to the foregoing argument need to be addressed. Probably the most prominent among researchers in the area would be the fact that our analysis has been concerned only with solutions of form (3), which excludes sunspot solutions of the "resonant frequency" type. It is my position, however, that the learning process pertaining to solutions of this type is much less plausible than for solutions of form (3). In particular, the solutions are not of the standard vector-autoregression (VAR) form. Therefore, an agent who experimented with many different specifications of VAR models, using the economy's generated time series data, would still not be led to such a solution. Indeed, it seems to me that arguments suggesting that that type of learning could exist in actual economies are utterly implausible. Of course, literally speaking, RE itself is implausible—as early critics emphasized. Nevertheless, RE is rightly regarded by mainstream researchers as the appropriate assumption for economic analysis, especially policy analysis. That is the case because RE is fundamentally the assumption that agents optimize with respect to their expectational behavior, just as they do (according to neoclassical economic analysis) with respect to other basic economic activities such as selection of consumption bundles, selection of quantities produced and inputs utilized, etc.—for a necessary condition for optimization is that individuals eliminate any systematically erroneous component of their expectational behavior. Also, RE is doubly attractive (to researchers) from a policy perspective, for it assures that a researcher does not propose policy rules that rely upon policy behavior that is designed to exploit patterns of suboptimal expectational behavior by individuals.

Another issue is the possible use of learning behavior not as a device for assessing the plausibility of rational expectations equilibria, but as a replacement for the latter. This type of approach is discussed by Evans and Honkapohja (2001) and has been prominent in the work of Orphanides and Williams (2005), among others. Use of constant-gain learning (Evans and Honkapohja, 2001) provides a sensible alternative to the decreasing-gain learning implicit in the LS learning/E-stability literature. This approach, however, does not seem to solve the "startup" problem, that is, the issue of how the economy will behave in the first several periods following the adoption of a new policy rule or the occurrence of some other structural change. It is highly unlikely that economies will move promptly to new RE equilibria following such a change, and I doubt that they would move promptly to a modeled learning path. In both cases, I share the opinion voiced by Lucas (1980), to the effect that, after a structural change (including policy regime changes), reliable analysis should pertain to the economy's behavior after it has had time to settle into a new dynamic stochastic equilibrium.

6. Conclusion

Let us now conclude with a very brief review of the points developed above. First, the paper reviews a previous result to the effect that, under the "first" information assumption that agents possess knowledge of current endogenous variables in the learning process, determinacy of a RE equilibrium is a sufficient but not necessary condition for least-squares learnability of that equilibrium. Thus, since learnability is an attractive necessary condition for plausibility of any equilibrium, there may exist a single plausible RE solution even in cases of indeterminacy. In addition, the paper proposes and outlines a distinct criterion that plausible models should possess, termed "well formulated" (WF), that rules out infinite discontinuities in the model's implied steady–state values of endogenous variables and in its impulse response functions. The paper then explores the relationship between this WF property and learnability, under the first information assumption, and finds that (although they often agree) neither implies the other. Extending the P-matrix requirement, implied for specified matrices by the WF property, to one that demands positive dominant-diagonal matrices would guarantee both WF and learnability, but a suitable rationale for such a requirement has not been found. Finally, under the second information assumption, which gives the agents only lagged information on endogenous variables during the learning process, the situation is less favorable in the sense that learnability can be guaranteed only under special assumptions.

APPENDIX

To demonstrate that a very wide variety of linear RE models can be written in form (1)–(2), consider the formulation of King and Watson (1998) or Klein (2000), as exposited by McCallum (1998), as follows:

$$\begin{bmatrix} \mathbf{A}_{11}^{*} & 0 \\ 0 & \mathbf{I} \end{bmatrix} \begin{bmatrix} E_t \mathbf{x}_{t+1} \\ \mathbf{k}_{t+1} \end{bmatrix} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{x}_t \\ \mathbf{k}_t \end{bmatrix} + \begin{bmatrix} \mathbf{G}_1 \\ \mathbf{G}_2 \end{bmatrix} [\mathbf{v}_t]. \tag{A1}$$

Here $\mathbf{v}_t$ is an AR(1) vector of exogenous variables (including shocks) with stable AR matrix $\mathbf{R}$, while $\mathbf{x}_t$ and $\mathbf{k}_t$ are $m_1 \times 1$ and $m_2 \times 1$ vectors of non-predetermined and predetermined endogenous variables, respectively. It is assumed, without significant loss of generality, that $\mathbf{B}_{11}$ is invertible[31] and that $\mathbf{G}_2 = 0$.[32] Define $\mathbf{y}_t = [\mathbf{x}_t{}' \quad \mathbf{k}_t{}' \quad \mathbf{x}_{t-1}{}' \quad \mathbf{k}_{t-1}{}']'$ and write the system in form (1) with $\mathbf{u}_t = \mathbf{v}_t$ and the matrices given, as follows:

$$\mathbf{A}_{11} = \begin{bmatrix} \mathbf{B}_{11}^{-1}\mathbf{A}_{11}^{*} & 0 & 0 & -\mathbf{B}_{11}^{-1}\mathbf{B}_{12} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \ \mathbf{C} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ \mathbf{B}_{21}\mathbf{B}_{22} & 0 & 0 & 0 \\ \mathbf{I} & 0 & 0 & 0 \\ 0 & \mathbf{I} & 0 & 0 \end{bmatrix} \text{ and } \mathbf{D} = \begin{bmatrix} -\mathbf{B}_{11}^{-1}\mathbf{G} \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{A2}$$

This representation is important because it is well known that the system (A1) permits, via use of auxiliary variables, any finite number of lags, expectational leads, and lags of expectational leads for the basic endogenous variables. Also, any higher-order AR process for the exogenous variables can be written in AR(1) form.[33] Thus it has been shown that the Evans and Honkapohja (2001) formulation is in fact rather general, although it does not pertain to asymmetric information models.

31. For the system (A1) to be cogent, each of the m_1 non-predetermined variables must appear in at least one of the m_1 equations of the first matrix row. Then the diagonal elements of $\mathbf{B}_{11}$ will all be non-zero and to avoid inconsistencies the rows of $\mathbf{B}_{11}$ must be linearly independent. This implies invertibility.

32. If it is desired to include a direct effect of $\mathbf{v}_t$ on $\mathbf{k}_{t+1}$, this can be accomplished by defining an auxiliary variable (equal to $\mathbf{v}_{t-1}$) in $\mathbf{x}_t$ (in which case $\mathbf{v}_t$ remains in the information set for period *t*). Also, auxiliary variables can be used to include expectations of future values of exogenous variables.

33. Binder and Pesaran (1995) show that virtually any linear model can be put in form (1), but in doing so admit a more general specification than (2) for the process generating the exogenous variables.

References

Benhabib, J., S. Schmitt-Grohé and M. Uribe. 2001. "The Perils of Taylor Rules." *Journal of Economic Theory* 96: 40–69.

Binder, M. and M.H. Pesaran. 1995. "Multivariate Rational Expectations Models and Macroeconometric Modeling: A Review and Some New Results." In *Handbook of Applied Econometrics*, edited by M.H. Pesaran and M. Wickens. Basil Blackwell Publishers.

————. 1997. "Multivariate Linear Rational Expectations Models: Characterization of the Nature of the Solutions and Their Fully Recursive Computation." *Econometric Theory* 13: 877–88.

Blanchard, O. and C.M. Kahn. 1980. "The Solution of Linear Difference Models under Rational Expectations." *Econometrica* 48: 1305–11.

Bullard, J.B. 2006. "The Learnability Criterion and Monetary Policy." *Federal Reserve Bank of St. Louis Review* 88(3): 203–17.

Carlstrom, C.T. and T.S. Fuerst. 2005. "Investment and Interest Rate Policy." *Journal of Economic Theory* 123: 4–20.

Clarida, R., J. Galí and M. Gertler. 1999. "The Science of Monetary Policy: A New Keynesian Perspective." *Journal of Economic Literature* 37: 1661–707.

Evans, G.W. and S. Honkapohja. 1999. "Learning Dynamics." In *Handbook of Macroeconomics*, edited by J.B. Taylor and M. Woodford. North-Holland.

————. 2001. *Learning and Expectations in Macroeconomics*. Princeton University Press.

————. 2003. "Expectations and the Stability Problem for Optimal Monetary Policies." *Review of Economic Studies* 70: 807–24.

Gale, D. and H. Nikaido. 1965. "The Jacobian Matrix and Global Univalence of Mappings." *Mathematische Annalen* 159: 81–93.

Golub, G.H and C.F. Van Loan. 1996. *Matrix Computations*, 3rd ed. Johns Hopkins University Press.

Horn, R.A. and C.R. Johnson. 1985. *Matrix Analysis*. Cambridge University Press.

————. 1991. *Topics in Matrix Analysis*. Cambridge University Press.

King, R.G. and M.W. Watson. 1998. "The Solution of Singular Linear Difference Systems under Rational Expectations." *International Economic Review* 39: 1015–26.

Klein, P. 2000. "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model." *Journal of Economic Dynamics and Control* 24: 1405–23.

Lucas, R.E. Jr., 1980. "Rules, Discretion, and the Role of the Economic Advisor." In *Rational Expectations and Economic Policy*, edited by S. Fischer. University of Chicago Press for National Bureau of Economic Research.

Magnus, J.R. and H. Neudecker. 1988. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley and Sons.

McCallum, B.T. 1983. "On Non-Uniqueness in Rational Expectations Models: An Attempt at Perspective." *Journal of Monetary Economics* 11: 139–68.

————. 1998. "Solutions to Linear Rational Expectations Models: A Compact Exposition." *Economics Letters* 61: 143–47.

————. 2003a. "Multiple-Solution Indeterminacies in Monetary Policy Analysis." *Journal of Monetary Economics* 50: 1153–75.

————. 2003b. "The Unique Minimum State Variable RE Solution Is E-Stable in All Well-Formulated Linear Models." Working paper 9960. Cambridge, Mass.: National Bureau of Economic Research.

————. 2004. "On the Relationship between Determinate and MSV Solutions in Linear RE Models." *Economics Letters* 84: 55–60.

————. 2007. "E-stability vis-à-vis Determinacy Results for a Broad Class of Linear Rational Expectations Models." *Journal of Economic Dynamics and Control* 31: 1376–91.

Orphanides, A. and J.C. Williams. 2005. "Imperfect Knowledge, Inflation Expectations, and Monetary Policy." In *The Inflation-Targeting Debate*, edited by B.S. Bernanke and M. Woodford. University of Chicago Press for National Bureau of Economic Research.

Rotemberg, J.J. and M. Woodford. 1999. "Interest Rate Rules in an Estimated Sticky Price Model." In *Monetary Policy Rules*, edited by J.B. Taylor. University of Chicago Press for National Bureau of Economic Research.

Sims, C.A. 1994. "A Simple Model for the Study of the Determination of the Price Level and the Interaction of Monetary and Fiscal Policy." *Economic Theory* 4: 381–99.

Uhlig, H. 1999. "A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily." In *Computational Methods for the Study of*

Dynamic Economies, edited by R. Marimon and A. Scott. Oxford University Press.

Walsh, C. 2003. *Monetary Theory and Policy*, 2d ed. MIT Press.

Woodford, M. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

A Sticky-Information General Equilibrium Model for Policy Analysis

Ricardo Reis
Columbia University

Following on Keynes's desire that economists be as useful as dentists, Lucas (1980) argues that this would amount to the following: "Our task, as I see it, is to write a FORTRAN program that will accept specific economic policy rules as 'input' and will generate as 'output' statistics describing the operating characteristics of time series we care about, which are predicted to result from these policies." Starting with Kydland and Prescott (1982), and with Rotemberg and Woodford (1997) in the context of monetary policy, the computer program that Lucas asked for has taken the form of dynamic stochastic general equilibrium (DSGE) models.[1] This paper follows the seminal work of Taylor (1979) in using one of these models to ask a series of hypothetical monetary policy questions.

However, the initial versions of monetary DSGE models suffer from one problem: they imply a rapid adjustment of many macroeconomic variables to shocks, while in the data, these responses tend to be gradual and delayed. The predictions of the standard classical model regarding investment, consumption, real wages, or inflation lack stickiness, to use the term coined by Sims (1998) and Mankiw and Reis (2006). The most popular approach for addressing this disconnect between theory and data follows the

At the time of the conference, Ricardo Reis was affiliated with Princeton University.

I am grateful to Olivier Coibion, Anton Nakov, Klaus Schmidt-Hebbel, Claudio Soto, Carl E. Walsh, Volker Wieland, and John Williams for useful comments and to Tiago Berriel for research assistance.

1. These are quickly growing in richness and being used in central banks, including the European Central Bank (Smets and Wouters, 2003) and the Board of Governors of the U.S. Federal Reserve (Erceg, Guerrieri, and Gust, 2006), as well as the International Monetary Fund (Bayoumi, 2004).

Monetary Policy under Uncertainty and Learning, edited by Klaus Schmidt-Hebbel and Carl E. Walsh, Santiago, Chile.

influential work of Christiano, Eichenbaum, and Evans (2005) by adding many rigidities that stand in the way of adjustment: sticky but indexed prices in goods markets, adjustment costs in investment markets, habits in consumption markets, and sticky but indexed wages in labor markets.

This paper contributes to the literature by providing an alternative DSGE model of business cycles and monetary policy. The only source of rigidity is inattention in all markets by agents who choose to only update their information sporadically in order to save on the fixed costs of acquiring, absorbing, and processing information (Reis, 2006a, 2006b). Information is sticky because different agents update their information at different dates, so they only gradually learn of news. I call it the sticky information in general equilibrium, or SIGE, model. Mankiw and Reis (2006, 2007) provided a first glimpse of SIGE, and this paper presents the model and its solution in full. I then proceed to estimate it for the United States after 1986 and the euro area after 1993 and to conduct a few policy experiments.

The paper is organized as follows. Section 1 presents the model and discusses its current limitations. Section 2 log-linearizes the model to arrive at a set of reduced-form relations that characterize the equilibrium. Section 3 describes an algorithm to compute a solution and derives formulas to calculate the key inputs into estimation (the likelihood function) and policy analysis (a social welfare function). Section 4 reviews the literature on estimating models with sticky information and describes the approach taken in this paper. Section 5 presents the estimation results for the United States and the euro area, while section 6 examines the sensitivity of the estimates. Section 7 answers a few policy questions, and section 8 concludes.

1. The SIGE Model

The SIGE model belongs to the wide class of general-equilibrium models with monopolistic competition that have become the workhorse for the study of monetary policy (surveyed in Woodford, 2003b). There are three sets of markets where agents meet every period: markets for different varieties of goods, where monopolistic firms sell varieties of goods to households; a market for savings, where households trade bonds and interest rates change to balance borrowing and lending; and markets for labor, where monopolistic households sell varieties of labor to firms. I present each of these markets in turn, before describing the assumptions on information and attention.

1.1 The Goods Market

On the buying side, there is a continuum of shoppers indexed by j that consume a continuum of varieties of goods in the unit interval indexed by i, denoted by $C_{t,j}(i)$. A bundle of these varieties of goods yields utility according to a Dixit-Stiglitz function with a time-varying and random elasticity of substitution $\tilde{\nu}_t$. Each good trades at price $P_{t,i}$ and the problem of a shopper with $Z_{t,j}$ to spend that observes current prices is

$$\max_{\{C_{t,j}(i)\}_{i\in[0,1]}} C_{t,j} = \left[\int_0^1 C_{t,j}(i)^{\frac{\tilde{\nu}_t}{\tilde{\nu}_t - 1}} di\right]^{\frac{\tilde{\nu}_t - 1}{\tilde{\nu}_t}}, \tag{1}$$

$$\text{subject to } \int_0^1 P_{t,i} C_{t,j}(i) < Z_{t,j}. \tag{2}$$

The solution to this problem is $C_{t,j}(i) = C_{t,j}\left(P_{t,i} / P_t\right)^{-\tilde{\nu}_t}$, where the price index is defined as

$$P_t = \left(\int_0^1 P_{t,i}{}^{1-\tilde{\nu}_t} di\right)^{1/(1-\tilde{\nu}_t)}$$

and implies that, conditional on the optimal choices of the shopper, $Z_{t,j} = P_t C_{t,j}$. Integrating over the continuum of shoppers gives the total demand for variety i:

$$\int_0^1 C_{t,j}(i) dj = \left(P_{t,i} / P_t\right)^{-\tilde{\nu}_t} \int_0^1 C_{t,j} dj. \tag{3}$$

On the selling side of the market, there is a monopolistic firm for each variety of the good. Each of these firms, indexed by i, operates a technology that uses labor $N_{t,i}$ at cost W_t to produce good i under diminishing returns to scale with $\beta \in (0, 1)$ and a common technology shock A_t. The firm's sales department is in charge of setting the price $P_{t,i}$ and selling the output $Y_{t,i}$ to maximize real after-tax profits subject to the technology and the demand for the good:

$$\max_{P_{t,i}} E_t^{(i)}\left[\frac{\left(1-\tau_p\right) P_{t,i} Y_{t,i}}{P_t} - \frac{W_t N_{t,i}}{P_t}\right], \tag{4}$$

subject to $Y_{t,i} = A_t N\beta_{t,i}$, (5)

$$Y_{t,i} = G_t \int_0^1 C_{t,j}(i)dj. \tag{6}$$

The $E_t^{(i)}(.)$ expectations operator of the sales department of firm i depends on its information, which I will discuss later. The government intervenes in two ways in the actions of the firm: collecting a fixed sales tax, τ_p, and buying a time-varying and random share, $1 - 1/G_t$, of the goods in the market. These governmental purchases are wasted, and I refer to them broadly as aggregate demand shocks. Aggregate output is $Y_t = \int_0^1 Y_{t,i} di$.[2]

After some rearranging, the first-order condition from this problem becomes

$$P_{t,i} = \frac{E_t^{(i)}\left[(1-\tau_p)\tilde{\nu}_t W_t N_{t,i} / P_t\right]}{E_t^{(i)}\left[(\tilde{\nu}_t - 1)\beta Y_{t,i} / P_t\right]}. \tag{7}$$

If the firm observed all the variables on the right-hand side, this condition would state that the nominal price charged, $P_{t,i}$, is equal to a markup, $(1-\tau_p)\,\tilde{\nu}_t/(\tilde{\nu}_t - 1)$, stemming from taxes and the ability to exploit an elastic demand curve, over nominal marginal costs, which equal the cost of an extra unit of labor, W_t, divided by its marginal product, $\beta Y_{t,i} / N_{t,i}$.

1.2 The Bond Market

In this market, saver-planners meet each other to trade one-period bonds. Their aim is to maximize the expected discounted utility from consumption:

$$E_t^{(j)} \sum_{t=0}^{\infty} \xi^t \left(\frac{C_{t,j}^{1-1/\theta}}{1 - 1/\theta} \right), \tag{8}$$

2. Defining aggregate output as $Y_t = \left(\int_0^1 Y_{t,i}^{(1-1/\tilde{\nu}_t)} di \right)^{\tilde{\nu}_t/(\tilde{\nu}_t - 1)}$ leads to the same results, up to a first-order log-linear approximation.

where ξ is the discount factor and θ is the intertemporal elasticity of substitution. They have an intertemporal budget constraint:

$$M_{t+1,j} = \Pi_{t+1}\left[M_{t,j} - C_{t,j} + \frac{(1-\tau_w)W_{t,j}L_{t,j}}{P_t} + T_{t,j}\right]. \tag{9}$$

The saver-planner j enters the period with real wealth $M_{t,j}$, uses some of it to consume, earns labor income at the wage rate $W_{t,j}$ after paying a fixed labor income tax τ_w, and receives a lump-sum transfer $T_{t,j}$. The transfer $T_{t,j}$ includes lump-sum taxes, profits and losses from firms, and payments from an insurance contract that all households signed at date 0 that ensures that every period they are all left with the same wealth. Savings accumulate at the real interest rate Π_{t+1}, although, in equilibrium, bonds are in zero net supply, so savings integrate to zero over all consumers.

The dynamic program that characterizes the saver-planner's problem is messy, so it is covered in the appendix. If $j = 0$ denotes the saver-planner that forms expectations rationally based on up-to-date information, so $E_t^{(0)} = E_t$, then the optimality conditions are

$$C_{t,0}^{-1/\theta} = \xi E_t\left(\Pi_{t+1}C_{t+1,0}^{-1/\theta}\right), \tag{10}$$

$$C_{t,j}^{-1/\theta} = E^{(j)}\left(C_{t,0}^{-1/\theta}\right). \tag{11}$$

The first equation is the standard Euler equation for a well-informed agent. It states that the marginal utility of consuming today is equal to the expected discounted marginal utility of consuming tomorrow times the return on savings. The second equation notes that agents who are not so well informed set their marginal utility of consumption to what they expect it would be with full information.

The monetary policymaker intervenes in this market by supplying reserves at an interest rate. Because these reserves are substitutable with the bonds that consumers trade among themselves, the central bank can target a value for the nominal interest rate, $i_t \equiv \log[E_t(\Pi_{t+1}P_{t+1}/P_t)]$, standing ready to issue as many reserves as necessary to ensure it. Alternatively, one could introduce money directly as an additive term in the agents' utility function and then have the central bank control the money supply to target an interest rate (see Woodford, 1998, for an elaboration of this point). The nominal interest rate follows some policy rule subject to exogenous monetary shocks ε_t. To fix ideas, and because it will be the policy rule used in the estimation, consider a Taylor rule:

$$i_t = \phi_y \log\left(\frac{Y_t}{Y_t^c}\right) + \phi_p \log\left(\frac{P_t}{P_{t-1}}\right) - \varepsilon_t, \tag{12}$$

where Y_t^c is the level of output in the classical or attentive equilibrium (sometimes called the natural output level).

1.3 The Labor Market

This market features workers on the selling side and firms on the buying side. The firms, indexed by i, have a purchasing department hiring a continuum of varieties of labor indexed by k in the amount $N_{t,i}(k)$ at price $W_{t,k}$ and combining them into the labor input $N_{t,i}$ according to a Dixit-Stiglitz function with a random and time-varying elasticity of substitution $\tilde{\gamma}_t$. The purchasing department's problem is to solve the following problem, given current wages and a total desired amount of inputs $N_{t,i}$:

$$\min_{\{N_{t,i}(k)\}_{k\in[0,1]}} \int_0^1 W_{t,k} N_{t,i}(k) dk, \tag{13}$$

$$\text{subject to } \left[\int_0^1 N_{t,i}(k)^{\frac{\tilde{\gamma}_t}{\tilde{\gamma}_t - 1}} dk\right]^{\frac{\tilde{\gamma}_t - 1}{\tilde{\gamma}_t}} = N_{t,i},$$

The solution to this problem is $N_{t,i}(k) = N_{t,i}\left(W_{t,k} / W_t\right)$, where $W_t N_{t,i} = \int_0^1 W_{t,k} N_{t,i}(k) dk$ for a static wage index $W_t = \left(\int_0^1 W_{t,k}{}^{1-\tilde{\gamma}_t} dk\right)^{1/(1-\tilde{\gamma}_t)}$. Aggregating over all firms gives the total demand for labor variety k:

$$\int_0^1 N_{t,i}(k) di = \left(\frac{W_{t,k}}{W_t}\right)^{-\tilde{\gamma}_t} \int_0^1 N_{t,i} di. \tag{14}$$

Each worker is a monopolistic supplier of a variety of labor. The workers' aim is to minimize their expected discounted disutility of labor:

$$E_t^{(k)} \sum_{t=0}^{\infty} \xi^t \left(\frac{\kappa L_{t,k}^{1+1/\psi}}{1+1/\psi} \right), \tag{15}$$

where ξ is the discount factor and ψ is the Frisch elasticity of labor supply. They face the same intertemporal budget constraint as the consumers in equation (9), and they also take into account the demand for their good $L_{t,k} = \int_0^1 N_{t,i}(k)di$ and equation (14). Aggregate labor employed is $L_t = \int L_{t,k} dk$.[3] The optimality conditions are

$$\frac{\tilde{\gamma}_t}{\tilde{\gamma}_t - 1} \times \frac{L_{t,0}^{1/\psi} P_t}{W_{t,0}} = \xi E_t \left(\Pi_{t+1} \times \frac{\tilde{\gamma}_{t+1}}{\tilde{\gamma}_{t+1} - 1} \times \frac{L_{t+1,0}^{1/\psi} P_{t+1}}{W_{t+1,0}} \right); \tag{16}$$

$$W_{t,k} = \frac{E_t^{(k)} \left[(1-\tau_w)\kappa \tilde{\gamma}_t L_{t,k}^{1/\psi} \right]}{E_t^{(k)} \left[\tilde{\gamma}_t L_{t,k} L_{t,0}^{1/\psi - 1} / W_{t,0} \right]}. \tag{17}$$

The first condition is the standard intertemporal labor supply Euler equation for a well-informed worker. If $\tilde{\gamma}_t$ is fixed, the equation states that the marginal disutility of supplying labor today ($L_{t,0}^{1/\psi}$) divided by the real wage ($W_{t,0}/P_t$) equates the discounted marginal disutility tomorrow ($L_{t+1,0}^{1/\psi}$) divided by the real wage tomorrow ($W_{t+1,0}/P_{t+1}$) times the real interest rate. With time-varying $\tilde{\gamma}_t$, the Euler equation takes into account the change in the markup that the monopolistic worker wants to charge. The second condition is the counterpart to condition (11) in the consumer problem—for the fully-informed case $E_t^{(k)} = E_t$, it simply states that $W_{t,k} = W_{t,0}$.

1.4 Information, Agents, and Attention

Uncertainty in this economy arises because every period there is a different realization of the random variables characterizing

3. As with output, defining aggregate labor as $L_t = \left(\int_0^1 L_{t,k}^{(1-1/\tilde{\gamma}_t)} dk \right)^{\tilde{\gamma}_t/(\tilde{\gamma}_t - 1)}$ leads to the same results up to a log-linear approximation.

productivity (A_t), aggregate demand (G_t), price and wage markups ($\tilde{\nu}_t$ and $\tilde{\gamma}_t$), and monetary policy (ε_t).

If all agents are fully informed, then the model described above is a standard classical model. While the discussion presented consumers (shoppers and saver-planners) and workers separately, they are all members of one household with period preferences

$$U\left(C_{t,j}, L_{t,k}\right) = \frac{C_{t,j}^{1-1/\theta}}{1-1/\theta} - \frac{\kappa L_{t,k}^{1+1/\psi}}{1+1/\psi}, \tag{18}$$

and with $j = k$ since there is common information. The decisions on the consumption of each variety, total consumption, and the wage to charge, are all made with rational expectations using all available information. Likewise, if the two departments of the firm share their information, they can be thought of as a single decisionmaker.

The SIGE model introduces only one new assumption relative to this classical benchmark: while the expectations of each agent are formed rationally, they do not necessarily use all available information. More concretely, it assumes that there are fixed costs of acquiring, absorbing, and processing information, so that agents optimally choose to only update their information sporadically (Reis, 2006a, 2006b). This inattentiveness is present in all of the markets—by the planner-savers in the savings market, by the sales departments of firms in the goods markets, and by the workers in the labor markets. Separating consumers from workers allows them to potentially update their information at different frequencies. In this case, while they share a household, in the sense of a common objective (equation 18) and a common budget constraint (equation 9), they do not necessarily need to share information. When workers update their information, they also learn about what the consumers have been doing, and vice versa for consumers when they update.

While inattentiveness occurs in all markets, not all agents in this economy are inattentive. In the goods market, the model assumes that the consumer is separated into two units: the saver-planner who updates information infrequently and the shopper who knows about the expenditure plan of the saver and observes the relative prices of the different goods. This assumption is not implausible: while the choice of how much to spend in total and how much to save requires solving an intertemporal optimization problem and making forecasts into the infinite future, choosing the relative proportion of each good to buy requires only seeing goods' prices. The main reason to make this assumption, though, is a current limitation in our knowledge. If the monopolistic firms in the goods' market faced inattentive shoppers, they would want to exploit them to raise profits, but the shoppers would then take this into account in choosing how often to be inattentive. The equilibrium of this game has not yet been fully studied, and assuming that shoppers are attentive avoids it entirely. The same argument leads to separating the firm into an inattentive sales-production team and an attentive purchasing department.

Within the inattentiveness model, the SIGE model adds an extra restriction: that the stochastic process for the expected costs of planning is such that the distribution of inattentiveness for consumers, workers, and firms is exponential. Reis (2006b) establishes the strict conditions under which this will hold for the firms' problem. Under these conditions, for a linearized homoskedastic economy, the optimal rate of arrival of information is fixed so that it can be treated as a parameter (bearing in mind that it maps into the monetary cost of updating information). Therefore, every period, a fraction of planner-savers δ updates its information, so there are δ agents who have current information, $\delta(1-\delta)$ that have one-period-old information, $\delta(1-\delta)^2$ with two-period-old information, and so on. Because agents only differ on the date at which they last updated, we can group them and let j denote how long ago the planner last updated. Likewise, a share λ of firms and ω of workers update their information every period, so they can be grouped into groups i of size $\lambda(1-\lambda)^i$ and groups k of size $\omega(1-\omega)^k$, according to how long it has been since they last updated.

The inattentive equilibrium is defined as follows: the set of aggregate variables $\{Y_t, L_t\}$, the output of each variety $\{Y_{t,i}\}$, the labor of each variety $\{L_{t,j}\}$, the prices of each good $\{P_{t,i}\}$, wages $\{W_{t,i}\}$, and interest rates $\{i_t\}$, such that consumers, workers, and firms behave optimally (as described above), all markets clear, and monetary policy follows a rule like equation (12), with $P_{-1} = 0$ for all dates t from 0 to infinity as a function of the exogenous paths for technology $\{A_t\}$, monetary policy shocks $\{\varepsilon_t\}$, aggregate demand $\{G_t\}$, goods' substitutability $\{\tilde{\nu}_t\}$, and labor substitutability $\{\tilde{\gamma}_t\}$. The classical equilibrium is the equilibrium when $\delta = \lambda = \omega = 1$, so that all agents are attentive.

1.5 Missing Work on the Micro-Foundations of the Model

In the tradition of Kydland and Prescott (1982) and Rotemberg and Woodford (1997), the SIGE model presented above makes a few simplifying assumptions, some of which are more common and others perhaps more unusual. Each of these presents an opportunity for future work to improve the model. I now discuss a few that seem particularly promising.

First, the model lacks investment and capital accumulation. Whether this absence significantly affects the dynamics of the other variables in this class of models is open to debate (Woodford, 2005; Sveen and Weinke, 2005), but modelling investment has the benefit of extending the model to explain one more macroeconomic variable. The SIGE model omits investment because the behavior of inattentive investors accumulating capital has not yet been studied, whereas there is previous work on the micro-foundations and implications of inattentiveness on the part of consumers (Reis, 2006a), price-seting firms (Mankiw and Reis, 2002; Reis, 2006b), and workers (Mankiw and Reis 2003). Gabaix and Laibson (2002) and Abel, Eberly, and Panageas (2007) study financial investment decisions with inattentiveness, but the step from this work to studying physical investment and capital accumulation remains to be taken.

Second, the model lacks international trade and exchange rates. The reason for this omission is the same as for investment: the models of inattentive behavior in international markets are still missing. Progress in this area will likely come soon, as Bachetta and van Wincoop (2006) have already filled some of the gap. Once this is completed, one can build an open economy SIGE to use for economies other than the United States or the euro area.

Third, the model lacks wealth heterogeneity since it assumes a complete insurance contract with which households fully diversify their risks. Most business cycle models make this assumption because it makes them more tractable by collapsing the wealth distribution to a single point. Relaxing this assumption and numerically computing the equilibria should not be difficult, but it has not yet been undertaken.

With regard to the micro-foundations of inattentiveness, the model assumes that when agents pay the cost to obtain new information, they can observe everything. While there is an explicit fixed cost of information, the variable cost is zero. This assumption is useful because it allows the model to emphasize the decision of when and how often to pay attention, which can then be studied in detail. It can easily be relaxed to allow people to observe only some things but not everything when they update (see, for example, Carroll and Slacalek, 2006). A harder extension would be to also consider the decision of how much to pay attention, by letting people pick which pieces of news to look at when they update. Mackowiack and Wiederholt (2007) have made promising progress in this area, following Sims (1998), but the models are still not at the point where they can be put in general equilibrium and taken to the data.

One implication of removing the assumption that updating agents learn everything, is that there is no longer common knowledge in the economy. This leads to a new source of strategic interactions between agents who have different information and know that no one knows everything. Woodford (2003a), Hellwig (2002), Amato and Shin (2006), Morris and Shin (2006), and Adam (2007) all study some of the implications of this behavior, and recent work by Lorenzoni (2008) moves toward turning these insights into a business cycle model that could be taken to the data. Hellwig and Veldkamp (2008) study another source of strategic interaction, namely, whether agents coordinate their attention times. These extra ingredients promise to enrich future models of inattentiveness.

The SIGE model ignores another source of strategic interaction. The model assumes that consumers have inattentive planners and attentive shoppers, while firms have inattentive sales departments and attentive purchasing departments. Consequently, monopolists only face attentive agents in every market. This is important because if a monopolist sells its product to some buyers that are inattentive, then it will want to exploit their inattentiveness to raise its profits (Gabaix and Laibson, 2006). These inattentive buyers would take into account this extra cost of being inattentive and alter their choices of when to update their information and how to act when uninformed. The equilibrium of this game has not, to my knowledge, been fully studied.

Overall, the SIGE model ignores many features that could lead to new and interesting insights. They were omitted mainly because they are not sufficiently understood to put them into the full DSGE setup used in this paper.

2. The Reduced-Form Log-Linear Equilibrium

The appendix describes how to log-linearize the equilibrium conditions around the Pareto-optimal steady state, where all the random variables are equal to their mean and the tax rates ensure that markups are zero. This gives a set of reduced-form relations characterizing the equilibrium of the log-linearized values of key aggregate variables (denoted with small letters and a t subscript), as a function of parameters and steady-state values (in small letters but no subscript).

First, summing the production function for the individual firms gives an aggregate relation between output (y_t), productivity (a_t), and labor (l_t) with decreasing returns to scale at rate β:

$$y_t = \alpha_t + \beta l_t. \tag{19}$$

Second, the equilibrium in the goods market leads to a Phillips curve (or aggregate supply) linking the price level (p_t) to marginal costs and desired markups. Real marginal costs rise with real wages ($w_t - p_t$), since these are the cost of inputs; they rise with output (y_t), as a result of decreasing returns to scale; and they fall with productivity (a_t). Desired markups are lower the higher the elasticity of substitution across goods' varieties (ν_t), where ν is the steady-state elasticity of substitution for goods:

$$p_t = \lambda \sum_{i=0}^{\infty} (1-\lambda)^i E_{t-i} \left\{ \begin{array}{l} p_t + \dfrac{\beta(w_t - p_t) + (1-\beta) y_t - a_t}{\beta + \nu(1-\beta)} \\ -\dfrac{\beta \nu_t}{(\nu - 1)[\beta + \nu(1-\beta)]} \end{array} \right\}. \tag{20}$$

Since only a fraction λ of firms update their information and set their plans, current shocks only have an immediate impact of λ on prices.

Third, the equilibrium in the bond market leads to an IS curve (or aggregate demand) relating output to three variables: a measure of wealth, namely, $y_\infty^c = \lim_{i \to \infty} E_t(y_{t+i})$, since higher expected future output stimulates current spending; the long real interest rate, defined as $R_t = E_t \sum_{j=0}^{\infty} (i_{t+j} - \Delta p_{t+1+j})$, since higher expected interest rates encourage postponing consumption; and shocks to government spending (g_t), since these subtract from consumption:

$$y_t = \delta \sum_{j=0}^{\infty} (1-\delta)^j E_{t-j} \left(y_\infty^c - \theta R_t \right) + g_t, \tag{21}$$

Every period, only a randomly drawn share δ of consumers update their plan, so the larger the value of δ, the more consumption responds to shocks as they occur.

Fourth, equilibrium in the labor market leads to a wage curve (or labor supply) according to which current wages (w_t) are higher with higher prices, since workers care about real wages; with higher expected real wages, since these push up the demand for a worker's variety of labor; with higher employment, since the marginal disutility of working rises; with higher wealth, since leisure is a normal good; with lower interest rates, since the return on savings is lower and the incentive to work to save is thus also lower; and with a lower elasticity of substitution across labor varieties, since desired markups are then higher:

$$w_t = \omega \sum_{k=0}^{\infty} (1-\omega)^k E_{t-k} \begin{bmatrix} p_t + \dfrac{\gamma(w_t - p_t)}{\gamma+\psi} + \dfrac{l_t}{\gamma+\psi} \\ + \dfrac{\psi\left(y_\infty^c - \theta R_t\right)}{\theta(\gamma+\psi)} - \dfrac{\psi\gamma_t}{(\gamma+\psi)(\gamma-1)} \end{bmatrix}. \tag{22}$$

The fraction of up-to-date workers is ω, with the remaining workers setting their wage to what they expected would be optimal when they last updated.

Finally, the policy rule gives the last reduced-form equilibrium relation. In the case of the Taylor rule, this relation is

$$i_t = \phi_p \Delta p_t + \phi_y (y_t - y_t^c) - \varepsilon_t. \tag{23}$$

These five equations give the equilibrium values for inflation, nominal interest rates, output growth, employment, and real wage growth, $x_t = \{\Delta p_t, i_t, \Delta y_t, l_t, \Delta(w_t - p_t)\}$, as a function of the five exogenous shocks to aggregate productivity growth, aggregate demand, goods markups, labor markups, and monetary policy, $s_t = \{\Delta a_t, g_t, \nu_t, \gamma_t, \varepsilon_t\}$. I assume that each of these shocks follows an independent stationary stochastic process with (potentially infinite) moving-average representation. This assumption allows for a very general representation of the shocks hitting the economy. One implication is that there is a stochastic trend in the economy driven by productivity, which seems consistent with the data.

3. Solving for the Equilibrium

I first solve for the equilibrium when all are attentive and then solve for the inattentive equilibrium under different policy rules.

Finally, I derive expressions for the likelihood and social welfare functions.

3.1 The Classical Equilibrium

In the classical equilibrium, all the agents are attentive, and simple algebra shows that output:

$$y_t^c = a_t + \Xi\left(g_t + \frac{\gamma_t}{\gamma - 1} + \frac{\nu_t}{\nu - 1}\right), \tag{24}$$

where $\Xi = \beta\psi/(1 + \psi)$, under the assumption that $\theta = 1$. Assuming that the elasticity of intertemporal substitution equals one implies that output moves one-to-one with the nonstationary productivity shocks, while hours worked, $l_t^c = (y_t^c - a_t)/\beta$, are stationary, as seems to be the case in the data.

In the classical equilibrium, output rises with each of the four real shocks, but it is independent of monetary policy shocks and the monetary policy rule. There are no nominal rigidities in this classical economy, so the classical dichotomy holds, with real variables being independent of monetary shocks.

Finally, it is important to note that this classical equilibrium is not necessarily optimal. The definition of a Pareto optimum is not obvious when there are changes in preferences. However, if the shocks to the preferences lead to an inefficiency relative to their steady-state values, then the optimal output is $y_t^o = a_t + \Xi g_t$, so shocks to the markups leads to inefficient fluctuations even if all agents are attentive.

3.2 The Inattentiveness Equilibrium

The solution of the inattentiveness equilibrium is a little more involved. One useful piece of notation is to write each variable in terms of its moving-average representation. For instance, for the generic shock $s \in S = \{\Delta a, g, \nu, \gamma, \varepsilon\}$, Wold's theorem implies that there is a representation $s_t = \sum_{n=0}^{\infty} \hat{s}_n e_{t-n}^s$, where the e_t^s are independent zero-mean random variables. For the endogenous variables that depend on all five shocks, $y_t^c = \sum_s \sum_n \Xi(s) s_n$, where the new coefficients $\Xi(s)$ follow easily from equation (24) and the definitions of Ξ and $\hat{s}_n$. Another useful piece of notation is to denote the share of people that have updated after n periods by $\Lambda_n = \lambda \sum_{i=0}^{n} (1-\lambda)^i$, $\Delta_n = \delta \sum_{i=0}^{n} (1-\delta)^i$, and $\Omega_n = \omega \sum_{i=0}^{n} (1-\omega)^i$.

The first result gives the first key step in the algorithm to solve the model:

Proposition 1. Writing the solution for the price level as $p_t = \sum_{s \in S} \sum_{n=0}^{\infty} \hat{p}_n(s) e_{t-n}^s$, where $\hat{p}_n(s)$ is a scalar measuring the impact of shock s at lag n, and likewise for output with $\hat{y}_n(s)$, then, regardless of the policy rule,

$$\hat{y}_n(s) = \Psi_n \hat{p}_n(s) + \Upsilon_n(s) \hat{s}_n, \tag{25}$$

where

$$\Psi_n^{den} = (1-\beta)(\gamma+\psi)\theta\Delta_n + \Omega_n \left\{\theta\Delta_n \left[1-\gamma(1-\beta)\right] + \psi\beta\right\}, \tag{26}$$

$$\Psi_n = \frac{\theta\Delta_n \left\{\left[\psi + \gamma(1-\Omega_n)\right]\left[\dfrac{\beta+\nu(1-\beta)}{\Lambda_n} - \nu(1-\beta)\right] - \beta\psi\Omega_n\right\}}{\Psi_n^{den}}, \tag{27}$$

and

$$\Upsilon_n(s) = \begin{cases} \theta\Delta_n\left(\gamma+\psi+\Omega_n-\gamma\Omega_n\right)a_n / \Psi_n^{den} & \text{for } s = a \\ \beta\psi\Omega_n g_n / \Psi_n^{den} & \text{for } s = g \\ \beta\theta\psi\Omega_n\Delta_n\gamma_n / \Psi_n^{den}(\gamma-1) & \text{for } s = \gamma \\ \beta\theta\Delta_n\left(\psi+\gamma-\gamma\Omega_n\right)\nu_n / \Psi_n^{den}(\nu-1) & \text{for } s = \nu \\ 0 & \text{for } s = \varepsilon. \end{cases} \tag{28}$$

The proof of this (and all other results) is in the appendix. It implies that given a solution for prices, one can easily compute the solution for output. A closely associated result is the following:

Proposition 2. The moving-average coefficients for the short-term real interest, wages, and hours worked as a function of those for prices and output are,

$$\hat{r}_n(s) = \frac{\hat{y}_{n+1}(s)}{\theta\Delta_{n+1}} - \frac{\hat{y}_n}{\theta\Delta_n} + \begin{cases} \dfrac{\hat{s}_n}{\theta\Delta_n} - \dfrac{\hat{s}_{n+1}}{\theta\Delta_{n+1}} & \text{for } s = g \\ 0 & \text{for } s = a, \gamma, \nu, \varepsilon \end{cases} \tag{29}$$

$$(\hat{w}_n - \hat{p}_n)(s) = \left[1 + \nu\left(\frac{1}{\beta} - 1\right)\right]\left(\frac{1}{\Lambda_n} - 1\right) \hat{p}_n(s)$$
$$+\left(1 - \frac{1}{\beta}\right)\hat{y}_n(s) + \begin{cases} \dfrac{\hat{s}_n}{\beta} & \text{for } s = a \\ \dfrac{\hat{s}_n}{\nu - 1} & \text{for } s = \nu \\ 0 & \text{for } s = g, \gamma, \varepsilon \end{cases} \tag{30}$$

$$\hat{l}_n(s) = \frac{\hat{y}_n(s)}{\beta} - \begin{cases} \dfrac{\hat{s}_n}{\beta} & \text{for } s = a \\ 0 & \text{for } s = g, \gamma, \nu, \varepsilon \end{cases} \tag{31}$$

With these two propositions and a solution for prices, we have the equilibrium values of all the real variables independently of the monetary policy rule. We can therefore focus on solving for prices alone.

If the policy rule is the one proposed by Taylor, then using the Fisher equation, $i_t = r_t + E_t(\Delta p_{t+1})$, and the results in the previous two propositions leads to the solution for the price level: [4]

Proposition 3. If the policy rule is a Taylor rule, $i_t = \phi_p \Delta p_t + \phi_y(y_t - y_t^n) - \varepsilon_t$, the undetermined coefficients for the price level satisfy the second-order difference equation:

$$A_{n+1}\hat{p}_{n+1}(s) - B_n \hat{p}_n(s) + C_{n-1}\hat{p}_{n-1}(s) = D_n(s)\hat{s}_n \text{ for } n = 0,1,2,... \tag{32}$$

where

$$A_n = 1 + \Psi_n / \theta\Delta_n,\ B_n = A_n + \phi_p + \phi_y \Psi_n, \text{ and } C_n = \phi_p, \tag{33}$$

and where

$$D_n(s) = \begin{cases} \dfrac{\Upsilon_n(s)}{\theta\Delta_n} - \dfrac{\Upsilon_{n+1}(s)\hat{s}_{n+1}}{\theta\Delta_{n+1}\hat{s}_n} + \phi_y\left[\Upsilon_n(s) - \Xi(s)\right] & \text{for } s = a, \gamma, \nu \\ \dfrac{\left[\Upsilon_n(s) - 1\right]}{\theta\Delta_n} - \dfrac{\left[\Upsilon_{n+1}(s) - 1\right]\hat{s}_{n+1}}{\theta\Delta_{n+1}\hat{s}_n} + \phi_y\left[\Upsilon_n(s) - \Xi(s)\right] & \text{for } s = g \\ -1 & \text{for } s = \varepsilon \end{cases} \tag{34}$$

4. Mankiw and Reis (2007) present an initial version of this result, limited to AR(1) shocks.

Solving the difference equation requires two boundary conditions. As the time from the shock goes to infinity, all agents become aware of it, so the effect of the shock on the inattentive equilibrium is the same as that in the attentive equilibrium. Since the price level converges to a constant (nonzero for the technology shocks and zero for the other shocks) regardless of the shock, one boundary condition is $\lim_{n\to\infty}(\hat{p}_n - \hat{p}_{n+1}) = 0$. The other boundary condition is $\hat{p}_{-1} = 0$.

I solve the difference equations by writing, separately for each shock, a system of $N + 1$ equations for the $N + 1$ undetermined coefficients from $\hat{p}_0(s)$ to $\hat{p}_N(s)$:

$$\begin{pmatrix} -B_0 & A_1 & \dots & 0 & 0 & 0 \\ C_0 & -B_1 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & -B_{N-2} & A_{N-1} & 0 \\ 0 & 0 & \dots & C_{N-2} & -B_{N-1} & A_N \\ 0 & 0 & \dots & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} \hat{p}_0(s) \\ \hat{p}_1(s) \\ \dots \\ \hat{p}_{N-2}(s) \\ \hat{p}_{N-1}(s) \\ \hat{p}_N(s) \end{pmatrix} = \begin{pmatrix} D_0(s) \\ D_1(s) \\ \dots \\ D_{N-2}(s) \\ D_{N-1}(s) \\ 0 \end{pmatrix}. \tag{35}$$

Because the system has a special tri-diagonal structure, it is numerically easy to solve. I have set N at either 100, 500, or 1,000. In almost all cases, both the ignored terms of order above N and the change in the first 100 coefficients as N changed were negligible.

Because the goal of this paper is to provide a model that can be used to study monetary policy, it is important to consider alternative policy rules to the Taylor rule. The main alternative to interest-rate rules are targeting rules (Svensson, 2003). Ball, Mankiw, and Reis (2005) show that if only firms are inattentive, an elastic price standard is optimal:

Proposition 4. If policy follows an elastic price-level standard, $p_t = K_t - \phi(y_t - y_t^o)$, the undetermined coefficients for the price level are as follows:

$$p_n(s) = \frac{\phi\left[\tilde{\Xi}(s) - \Upsilon_n(s)\right]s_n}{1 + \phi\Psi_n} \text{ for } n = 0,1,2,\dots \tag{36}$$

where $\tilde{\Xi}(s) = \Xi(s)$ for $s = a, g$; and $\tilde{\Xi}(s) = 0$ for $s = \gamma, \nu$.

The literature contains many alternative policy rules, and the appendix presents a few more and their corresponding solution. Together with the results in this section, this should provide sufficient evidence that despite the infinite number of expectations going backward and the lack of a recursive representation for the endogenous variables, the SIGE model is still easy to solve.[5]

3.3 The Likelihood and Welfare Functions

The key input in likelihood-based estimation is the likelihood function. Letting $\mathbf{x}_t$ denote the 5×1 column vector with the endogenous variables of the model and $\mathbf{e}_t$ denote the column vector with the 5 exogenous shocks, the solution in propositions 1 to 4 can be expressed as a set of 5×5 matrices Φ_n, such that $\mathbf{x}_t = \sum_{n=1}^{N} \Phi_n \mathbf{e}_{t-n}$. The data consists of time-series on $\mathbf{x}_t$ from $t = 1$ to $t = T$ for the endogenous variables, which can be stacked in a $5T\times 1$ vector $\mathbf{X}$, and the unknown parameters can be collected in the vector θ. The likelihood function is then denoted by $L(\mathbf{X}|\theta)$.

I assume that the five zero-mean shocks e_t^s are normally distributed with variances σ_s^2. The vector $\mathbf{e}_t$ therefore follows a multivariate normal distribution with diagonal covariance matrix $\mathbf{\Sigma}$. The notation $\mathbf{I}_N$ denotes an identity matrix of size N and $\otimes$ for the Kronecker product of two matrices. Since the model is linear, $\mathbf{X}$ follows a multivariate normal distribution. This leads to the next proposition, taken from Mankiw and Reis (2007):

Proposition 5. Let $\mathbf{\Omega}$ be the $5T\times 5N$ matrix,

$$\begin{pmatrix} \Phi_0 & \Phi_1 & \Phi_2 & \dots & \dots & \dots & \Phi_{N-3} & \Phi_{N-2} & \Phi_{N-1} \\ 0 & \Phi_0 & \Phi_1 & \dots & \dots & \dots & \Phi_{N-4} & \Phi_{N-3} & \Phi_{N-2} \\ 0 & 0 & \Phi_0 & \dots & \dots & \dots & \Phi_{N-5} & \Phi_{N-4} & \Phi_{N-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & \Phi_0 & \Phi_1 & \dots & \Phi_{N-T-1} & \Phi_{N-T} \end{pmatrix}; \tag{37}$$

the likelihood function is then

$$L(\mathbf{X}|\theta) = -2.5T\ln(2\pi) - 0.5\ln\left|\mathbf{\Omega}(\mathbf{I}_N \otimes \mathbf{\Sigma})\mathbf{\Omega}'\right| - 0.5\mathbf{X}'\left(\mathbf{\Omega}\left(\mathbf{I}_N \otimes \mathbf{\Sigma}\right)\mathbf{\Omega}'\right)^{-1}\mathbf{X}.$$

Mankiw and Reis (2007) note that the large $5T\times 5T$ matrix $\mathbf{\Omega}(\mathbf{I}_N\otimes\mathbf{\Sigma})\mathbf{\Omega}'$ can be inverted either with a Choleski decomposition

5. Building on some of these results, Meyer-Gohde (2007) combines this approach with others in the literature to provide a unified user-friendly algorithm that can solve most DSGE models with forward and lagged expectations without requiring almost any algebra on the part of the user (unlike the propositions above). His set of programs holds the promise of further advancing this literature.

or by choosing $N = T$ to re-express the problem in terms of a system of linear equations. Either way, one can evaluate the log-likelihood function quickly and reliably.

A natural way to compare the performance of different policy rules is to compute the utility of the agents in the model. I focus on the unconditional expectation of a utilitarian measure of social welfare:

$$E\left[(1-\xi)\sum_{t=0}^{\infty}\xi^t \int\int U\left(C_{t,j}, L_{t,k}\right)djdk\right]. \tag{38}$$

Because the model assumes that all households are ex ante identical and there are complete insurance markets, it is natural to assume that all households get the same weight in the integral. Moreover, because one wants a rule that performs well across circumstances, it makes sense to take the ex ante perspective provided by the unconditional expectation that integrates over all possible initial conditions. The appendix proves the following result:

Proposition 6. An approximate formula for the welfare benefits in percentage units of steady-state consumption of a policy $\boldsymbol{\theta}^{(1)}$ starting from a policy $\boldsymbol{\theta}^{(0)}$ are

$$\exp\left\{0.5\beta\left(1+\frac{1}{\psi}\right)\left[W\left(\boldsymbol{\theta}^{(1)}\right)-W\left(\boldsymbol{\theta}^{(0)}\right)\right]\right\}, \tag{39}$$

where

$$W(\boldsymbol{\theta}) = -\sum_{s\in S}\sum_{n=0}^{\infty}\left[(1-\Omega_n)\varsigma_n(s)^2 + \Omega_n\zeta_n(s)^2\right]\sigma_s^2, \tag{40}$$

$$\varsigma_n\left(s\right) = \hat{l}_n\left(s\right) + \gamma\hat{w}_n\left(s\right), \text{ for all } s, \tag{41}$$

and

$$\frac{(\gamma+\psi)\zeta_n(s)}{\gamma\psi} = \frac{\hat{l}_n(s)}{\gamma} + \left(\hat{w}-\hat{p}\right)_n(s) + \frac{\hat{y}_n(s)}{\Delta_n} + \begin{cases} 0 & \text{for } s = \varepsilon, a, \nu \\ \dfrac{\hat{s}_n}{\gamma-1} & \text{for } s = \gamma \\ -\dfrac{\hat{s}_n}{\Delta_n} & \text{for } s = g \end{cases} \tag{42}$$

Combining this result with those in propositions 1 to 4, it is easy to evaluate this expression and compare the performance of different policy rules.

4. Estimating Sticky Information

Taking sticky information models to the data has been an active field of research. One approach is to look for direct evidence of inattentiveness using microeconomic data. Carroll (2003) uses surveys of inflation expectations to show that the public's forecasts lag the forecasts made by professionals.[6] Mankiw, Reis, and Wolfers (2004) show that the disagreement in the inflation expectations in the survey data have properties consistent with sticky information.[7] Reis (2006a) and Carroll and Slacalek (2006) interpret some of the literature on the sensitivity and smoothness of microeconomic consumption data in the light of sticky information, and Klenow and Willis (2007) and Knotek (2006) find slow dissemination of information in the microeconomic data on prices. For the most part, this literature supports the assumption of sticky information, and the associated estimates of the information-updating rates are consistent.

A second approach is to estimate Phillips curves assuming sticky information on the part of price setters only.[8] These limited-information approaches typically use data on inflation, output, marginal costs, and expectations to estimate simpler versions of equation (20), and the results are typically good or mixed. One interesting finding that comes out of many of these studies is that the main source of discrepancy between the model and the data is not the inattentiveness or the slow dissemination of information, but the assumption that, conditional on their information sets, agents form expectations rationally.

This paper takes a third approach, of estimating the model using full-information techniques that exploit the restrictions imposed by general equilibrium. The few papers that attempt this exercise typically find either mixed or poor fits between the model and the data.[9] Mankiw and Reis (2006) explain the contrast between the

6. See also Dopke and others (2008) and Nunes (2006).

7. Also focussing on disagreement, see Gorodnichenko (2006), Branch (2007), and Rich and Tracy (2006).

8. See Khan and Zhu (2006), Dopke and others (2006), Korenok (2005), Pickering (2004), Coibion (2007), and Molinari (2007).

9. See Trabandt (2007), Andrés, Nelson, and López-Salido (2005), Kiley (2007), Laforte (2007), Korenok and Swanson (2005, 2007), and Paustian and Pytlarczyk (2006).

negative results in some of these papers and the mostly positive results found by the other two approaches. They note that the papers in this literature assume inattentiveness only in price setting, while assuming that the other agents in the model are fully attentive. To fit the data, however, stickiness should be pervasive, and for the internal coherence of the model, inattentiveness should apply to all decisions. By assuming attentive consumer and workers, the general-equilibrium restrictions imposed in these papers are misspecified.

Allowing for pervasive stickiness, I take a Bayesian approach to deal with the uncertainty, starting with a prior joint probability density $p(\boldsymbol{\theta})$ and using the likelihood function $L(\mathbf{X}|\boldsymbol{\theta})$ to obtain the posterior density of the parameters $p(\theta|\mathbf{X})$. This is done numerically, using Markov chain Monte Carlo simulations.[10]

The prior density $p(\boldsymbol{\theta})$ follows the convention in the DSGE literature (for example, An and Schorfheide, 2007), including assuming that the shocks s_t follow first order autoregressive, or AR(1), processes with coefficients ρ_s and innovation standard deviations σ_s. There are twenty parameters in the model: $\boldsymbol{\theta} = \{\theta, \psi, \nu, \gamma, \beta, \rho_{\Delta a}, \sigma_{\Delta a}, \rho_{\varepsilon}, \sigma_{\varepsilon}, \rho_{g}, \sigma_{g}, \rho_{\nu}, \sigma_{\nu}, \rho_{\gamma}, \sigma_{\gamma}, \phi_p, \phi_y, \delta, \omega, \lambda\}$. Table 1 shows the moments of the prior densities.

Four of the parameters have a tight prior with zero variance: θ, which is set to one to ensure stationary hours worked; β, which equals two-thirds to match the labor share in the data; and $\rho_{\Delta a}$ and $\sigma_{\Delta a}$, since a series for productivity growth follows from the data on output and employment in equation (19), so we can recover these parameters by a simple least-squares regression.[11]

Each of the remaining sixteen parameters is treated independently and is assigned a particular distribution (gamma, beta, or uniform) with a relatively large variance. The mean elasticity of labor supply, ψ, is 2 and the elasticities of substitution across goods and labor varieties, ν and γ, are set at 11, in line with the typical assumptions in the literature. The mean ρ_s for the four shocks other than productivity are set to 0.9, so that the half-life of the shocks is approximately six quarters and the σ_s are set to 0.5, which lies between the two values estimated for $\sigma_{\Delta a}$.[12] The

10. The exact algorithm is described in the appendix.

11. The values for $\rho_{\Delta a}$ and $\sigma_{\Delta a}$ are 0.03 and 0.51, respectively, for the United States and 0.66 and 0.28 for the euro area.

12. For the markups, the value for the standard deviation is multiplied by ten and the elasticities of substitution are multiplied by minus one, to counteract the multiplier that is visible in equations (20) and (22).

Table 1. Prior Distribution

Parameter	*Density*[a]	*Mean*	*Standard deviation*	*Percentile*		
				2.5	*50.0*	*97.5*
Preferences						
ν	$1 + G$	11	3.16	5.80	10.67	18.08
γ	$1 + G$	11	3.16	5.80	10.67	18.08
ψ	G	2	2.00	0.05	1.39	7.38
Nonpolicy shocks						
ρ_g	B	0.90	0.20	0.23	0.99	1.00
σ_g	$G^{-1/2}$	0.50	0.24	0.21	0.39	1.02
ρ_ν	B	0.90	0.20	0.23	0.99	1.00
σ_ν	$G^{-1/2}$	0.50	0.24	0.21	0.39	1.02
ρ_γ	B	0.90	0.20	0.23	0.99	1.00
σ_γ	$G^{-1/2}$	0.50	0.24	0.21	0.39	1.02
Monetary policy						
ϕ_p	$1 + G$	1.24	0.25	1.00	1.16	1.92
ϕ_y	G	0.33	0.25	0.03	0.27	0.97
ρ_ε	B	0.90	0.22	0.23	0.99	1.00
σ_ε	$G^{-1/2}$	0.50	0.11	0.21	0.39	1.02
Inattentiveness						
δ	U	0.50	0.29	0.03	0.50	0.98
ω	U	0.50	0.29	0.03	0.50	0.98
λ	U	0.50	0.29	0.03	0.50	0.98

Source: Author's calculations.
a. The densities are the gamma (G), beta (B) and uniform (U).

monetary policy parameters are set at $\phi_p = 1.24$ and $\phi_y = 0.33$, which are the values estimated by Rudebusch (2002) on U.S. data. Finally, the inattentiveness parameters δ, ω, and λ have a flat prior in the unit interval.

As for the data, I use quarterly observations for two large economies: the United States from 1986:3 to 2006:1 and the euro area from 1993:4 to 2005:4. I chose these countries because they are closer to the closed-economy approximation in the model. The starting dates coincide with the start of Alan Greenspan's term as chairman of the Federal Open Market Committee (FOMC) in the United States and with the signing of the Maastricht treaty that created the European Union and started the coordination of monetary policy towards the euro, so they are consistent with assuming a stable monetary policy rule. They come after the "great moderation" in economic activity, consistent with assuming constant variances of the shocks.

The data for the United States are seasonally adjusted, refer to the nonfarm business sector, and comprise observations on growth in real output per capita, growth in total real compensation per hour, hours worked per capita, and inflation. All series are demeaned; they use the implicit nonfarm business price deflator for the price level and for deflating nominal values; and growth rates refer to the change in the natural logarithm. The nominal interest rate is the effective Federal funds rate. The data for the euro area are the area-wide quarterly dataset that combines data from each country's national accounts to build consistent pseudo-aggregates for the whole region. Inflation is the change in the log of the GDP deflator, output growth is the change in log real GDP, and wages are measured using total compensation. To obtain variables per capita, I use an interpolated euro area population series. The hours data are detrended using a linear trend.

5. Estimates of the Model

I discuss the estimates for the two regions separately.

5.1 The United States

Table 2 displays summary statistics of the posterior distribution of the parameters. The posterior moments for the elasticities of

Table 2. Posterior Distribution for the United States[a]

Parameter	*Mean*	*Standard deviation*	*Percentile*		
			2.5	*50.0*	*97.5*
Preferences					
ν	10.09	2.67	5.83	9.75	15.93
γ	9.09	2.64	4.74	8.83	14.63
ψ	5.15	2.52	1.18	4.94	10.95
Nonpolicy shocks					
ρ_g	0.99	0.01	0.98	1.00	1.00
σ_g	0.83	0.16	0.59	0.81	1.23
ρ_ν	0.28	0.10	0.08	0.28	0.48
σ_ν	0.11	0.06	0.03	0.09	0.26
ρ_γ	0.86	0.08	0.71	0.85	1.00
σ_γ	0.12	0.06	0.05	0.11	0.27
Monetary policy					
ϕ_p	1.17	0.16	1.01	1.12	1.60
ϕ_y	0.06	0.03	0.01	0.06	0.14
ρ_ε	0.29	0.12	0.07	0.30	0.52
σ_ε	0.44	0.09	0.30	0.43	0.65
Inattentiveness					
δ	0.08	0.03	0.03	0.08	0.16
ω	0.74	0.17	0.34	0.78	0.98
λ	0.52	0.17	0.28	0.48	0.94

Source: Author's calculations.
a. All numbers are based on 450,000 draws from the posterior.

substitution across varieties are close to the prior assumptions from the literature. The elasticity of labor supply is quite large, but still in line with typical assumptions in the business cycle literature. As for the shocks, the aggregate demand disturbances are very persistent and quite volatile, so one can already guess that they are playing an important part in the volatility of the economy.

The more interesting estimates are those of the inattentiveness parameters, on which the prior had less information. Firms are estimated to be inattentive for six months, on average, which is slightly more attentive than what was found in the studies described in the previous section. Consumers are very inattentive, updating their information once every three years, on average. This is not too shocking considering that fixed costs of planning of less than $100 per household can easily generate this length of inattentiveness. Moreover, between 20 percent and 50 percent of the U.S. population lives hand-to-mouth, which is equivalent to being inattentive forever (Reis, 2006a).

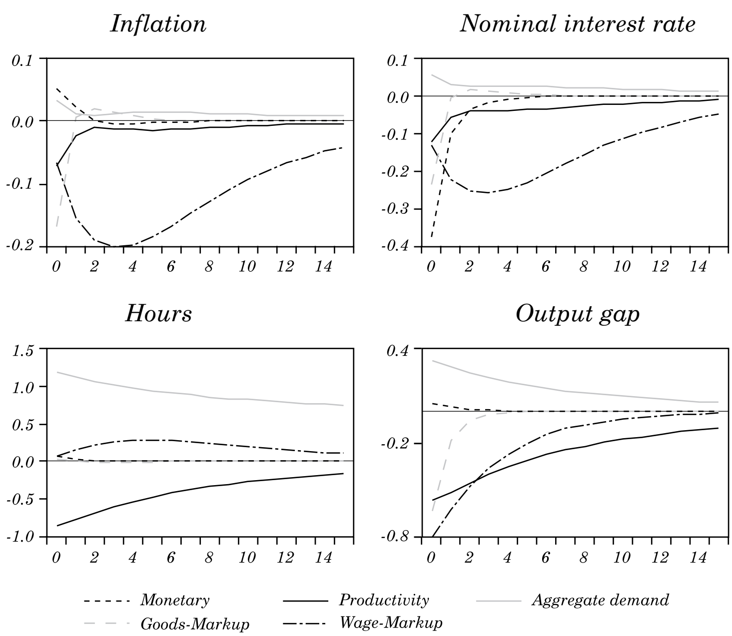
The more surprising estimate in the table is the inattention of workers, who update their information very often, on average once every four months. One possible explanation for this result is that the data series used for wages measured total compensation, a large fraction of which is accounted for by nonwage payments. It is conceivable that the many dimensions of an employee's compensation may actually be updated to include new information quite often, even if the wage component of this compensation is not. Preliminary calculations using a wage series find more inattentive workers, and workers are also more inattentive in the euro area, where nonwage compensation is less important.

Figure 1 shows the impulse responses of four variables (namely, inflation, nominal interest rates, hours worked, and the output gap) to one-standard-deviation impulses to the five shocks. The most surprising finding is perhaps the quick response of inflation to monetary policy shocks. The conventional wisdom from studies using postwar U.S. data is that this response should be delayed and hump shaped. As recent studies have shown, however, inflation responds much faster to monetary policy after 1980, which some researchers attribute to changes in monetary policy.[13] From the perspective of the SIGE model, inflation responds quickly to monetary policy because monetary policy shocks are quite short-lived. When policy changes, the SIGE model predicts a

13. See Boivin and Giannoni (2006) and the references therein.

change in the dynamics of the model that matches the data, surviving the Lucas critique in a way that pricing models that always produce a hump shape do not.

Figure 1. Impulse Response Functions to the Five Shocks: United States



Source: Author's calculations.

Table 3 presents the predicted variance decompositions at different horizons. Monetary policy shocks play a small role in the variance of most macroeconomic variables in the United States after 1986, with the exception of the nominal interest rate and wages. Productivity shocks are important for real wages at all horizons and for hours worked at short horizons, while aggregate demand shocks explain much of the variability of output growth and hours worked.[14] Finally, inflation is significantly driven by the markup shocks.

14. Of all the model's shocks, these aggregate demand shocks are closest to the shocks to the marginal rate of substitution between consumption and leisure that Hall (1997) argues account for most of the U.S. business cycle.

Table 3. Variance Decompositions for the United States[a]
In percentage points

	Shock				
Variable	*Monetary*	*Aggregate productivity*	*Aggregate demand*	*Goods markup*	*Labor markup*
A. Contribution to unconditional variance					
Inflation	1 (0, 35)	7 (1, 35)	3 (0, 58)	11 (1, 58)	68 (7, 92)
Interest rate	24 (1, 69)	6 (1, 30)	3 (0, 44)	8 (2, 28)	49 (8, 84)
Output growth	0 (0, 3)	4 (1, 11)	94 (83, 98)	0 (0, 1)	1 (0, 5)
Hours worked	0 (0, 0)	4 (1, 21)	94 (26, 98)	0 (0, 0)	1 (0, 70)
Wage growth	12 (1, 30)	45 (20, 77)	6 (1, 18)	8 (3, 22)	24 (0, 59)
B. Contribution to one-quarter-ahead, one-year-ahead, and four-year-ahead variance					
Inflation	7, 3, 1	12, 7, 7	3, 2, 2	64, 24, 12	8, 57, 71
Interest rate	63, 40, 26	6, 6, 6	1, 2, 2	21, 12, 9	5, 31, 49
Output growth	0, 0, 0	1, 3, 4	98, 95, 94	0, 0, 0	0, 1, 1
Hours worked	0, 0, 0	35, 32, 21	64, 67, 75	0, 0, 0	0, 1, 3
Wage growth	11, 12, 12	47, 46, 45	6, 6, 6	6, 9, 8	26, 23, 24

Source: Author's calculations.
a. Panel A reports the median of 10,000 parameter draws from the posterior distribution; the numbers in parentheses are the 2.5 and 97.5 percentiles. Panel B reports medians from the same number of draws. Rows will not add up to 100, since the medians are cell by cell.

5.2 The Euro Area

Table 4 shows moments from the posterior distribution for the euro area. Relative to the U.S. estimates, there are two differences. First, the estimated average markups are larger for the euro area than for the United States. Second, the elasticity of labor supply is somewhat smaller, although it is still large compared with typical estimates based on microeconomic data. The inattentiveness of European firms is similar to that of American firms, while consumers are more attentive and workers less attentive. This brings the two members of the household in line, with both updating every nine to fifteen months, on average.

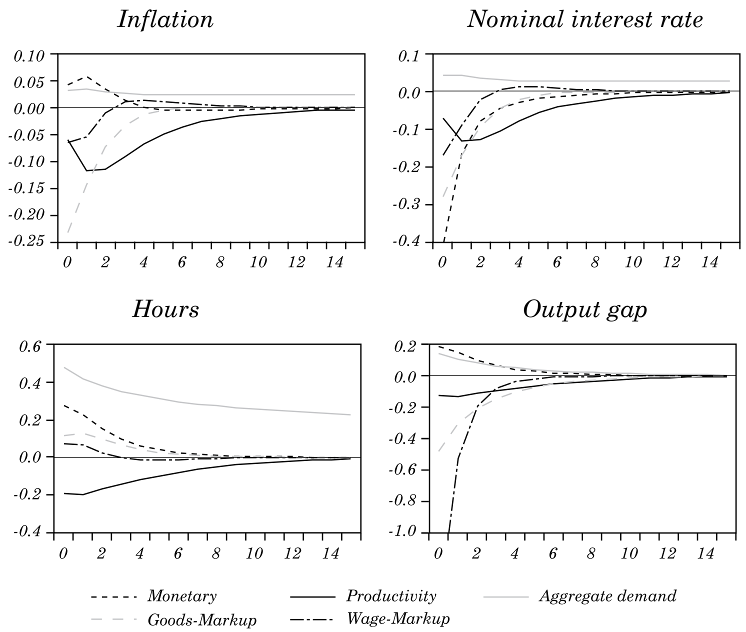
Table 4. Posterior Distribution for the Euro Area[a]

			Percentile		
Parameter	*Mean*	*Standard deviation*	*2.5*	*50.0*	*97.5*
Preferences					
ν	8.16	1.31	5.94	7.98	10.80
γ	7.11	0.75	5.49	7.26	8.34
ψ	2.70	0.43	1.92	2.74	3.46
Nonpolicy shocks					
ρ_g	0.99	0.01	0.95	0.99	1.00
σ_g	0.37	0.10	0.22	0.35	0.62
ρ_ν	0.70	0.21	0.31	0.67	0.98
σ_ν	0.08	0.05	0.03	0.07	0.20
ρ_γ	0.37	0.15	0.09	0.39	0.62
σ_γ	0.19	0.09	0.08	0.17	0.41
Monetary policy					
ϕ_p	1.06	0.10	1.00	1.01	1.35
ϕ_y	0.07	0.02	0.01	0.05	0.24
ρ_ε	0.51	0.11	0.27	0.54	0.66
σ_ε	0.46	0.12	0.30	0.44	0.75
Inattentiveness					
δ	0.21	0.11	0.10	0.17	0.52
ω	0.31	0.18	0.15	0.26	0.93
λ	0.58	0.15	0.26	0.62	0.79

Source: Author's calculations.
a. All numbers are based on 450,000 draws from the posterior.

Figure 2 shows the impulse responses to shocks in the euro area. The response of inflation to a monetary shock is now slightly hump shaped, but it peaks just two quarters after the shock. Moreover, the response of all variables to a monetary shock is more delayed than in the United States.

Figure 2. Impulse Response Functions to the Five Shocks: Euro Area



Source: Author's calculations.

As was the case for the United States, a positive productivity shocks raises total output but lowers hours worked and the output gap on impact, consistent with the evidence in Galí (2004). Because many firms initially do not know about the shock, they do not raise their output as much as they would with full information. Likewise, an increase in the elasticities of substitution (that is, a positive markup shock) raises hours worked and output, but leads to a negative output gap, because the expansion is smaller than would be the case with full information. Aggregate demand shocks boost inflation and the output gap and thus raise nominal interest rates, via the Taylor rule.

Table 5 has the variance decompositions for the euro area. Monetary policy shocks play a significantly larger role in explaining the variability of output growth and hours worked than they did in the United States, while productivity shocks are also more important drivers of output and inflation. Aggregate demand shocks are still important in explaining output and hours worked, as are markup shocks for inflation.

6. Robustness of the Estimates

This section summarizes the impact of several changes to the specification choices on the posterior estimates. Starting with the priors, I attempted a few variations from the baseline in table 1. Because fully characterizing the posterior distributions is computationally time consuming, I focused only on their modes. The three experiments were as follows: raising the prior mean for the elasticity of labor supply from 2 to 4; lowering the prior mean correlation of the shocks from 0.9 to 0.5; and setting the prior standard deviation of the shocks equal to $\sigma_{\Delta a}$ in each region, rather than to the 0.5 in-between value. Each of these changes had a negligible difference in the mode of the posterior distribution.

With regard to the policy rule, an alternative to the Taylor rule in equation (23) with serially correlated shocks is an inertial rule:

$$i_t = \phi_p \Delta p_t + \phi_y \left(y_t - y_t^c \right) + \rho_i i_{t-1} - \varepsilon_t, \tag{43}$$

where the ε_t are serially uncorrelated. I estimated this alternative model and obtained a mean posterior estimate for ρ_i of 0.25 for the United States and 0.16 for the euro area. In terms of overall fit to the data, the results are mixed. For the United States, the marginal density for the inertial rule is higher, whereas for the euro area, the Taylor rule with correlated shocks dominates.

In terms of the data, the main issue to address is a clear upward trend in hours worked in the euro area, associated with the slow decline in European unemployment. In the main results, I dealt with it by removing a linear trend from the data. Using a Hodrick-Prescott (HP) filter led to the same results. There is no trend in the U.S. data, so detrending it with the HP filter or even not detrending it at all led to almost indistinguishable data series.

Finally, for the sample periods, Mankiw and Reis (2007) estimate a subset of the parameters using postwar U. S. data.

Table 5. Variance Decompositions for the Euro Area[a]
In percentage points

	Shock				
Variable	*Monetary*	*Aggregate productivity*	*Aggregate demand*	*Goods markup*	*Labor markup*
A. Contribution to unconditional variance					
Inflation	2	15	12	40	2
	(0, 18)	(1, 52)	(0, 85)	(5, 91)	(0, 46)
Interest rate	30	10	7	22	4
	(8, 78)	(1, 26)	(0, 71)	(4, 78)	(0, 37)
Output growth	16	25	50	3	1
	(4, 42)	(11, 42)	(20, 79)	(1, 16)	(0, 12)
Hours worked	2	5	84	1	0
	(0, 17)	(0, 27)	(28, 99)	(0, 64)	(0, 2)
Wage growth	2	31	1	36	20
	(0, 22)	(10, 53)	(0, 5)	(4, 71)	(3, 80)
B. Contribution to one-quarter-ahead, one-year-ahead, and four-year-ahead variance					
Inflation	3, 4, 3	8, 31, 26	2, 3, 4	78, 53, 48	5, 4, 3
Interest rate	60, 46, 40	3, 15, 14	1, 1, 2	25, 25, 25	7, 6, 5
Output growth	18, 16, 16	12, 24, 26	63, 51, 49	3, 3, 3	1, 1, 1
Hours worked	17, 13, 6	14, 18, 13	60, 59, 69	3, 2, 1	1, 0, 0
Wage growth	2, 2, 2	13, 31, 31	1, 1, 1	47, 34, 36	32, 21, 20

Source: Author's calculations.
a. Panel A reports the median of 10,000 parameter draws from the posterior distribution; the numbers in parentheses are the 2.5 and 97.5 percentiles. Panel B reports medians from the same number of draws. Rows will not add up to 100, since the medians are cell by cell.

Relative to the results in table 2, they find that workers and consumers update their information every five to six quarters, on average, which is close to the euro area estimates in this paper. They also find much more persistent and volatile monetary policy shocks, such that monetary shocks account for a large share of the volatility of the macroeconomic series. One conjecture for what is behind this discrepancy is that including the high inflation of the 1970s in the sample requires large monetary policy shocks that play a large role in the business cycle.

7. Policy Questions

To begin applying the two estimated models to policy analysis, I explore some questions about monetary policy.

7.1 What Rule Has Best Described Policy?

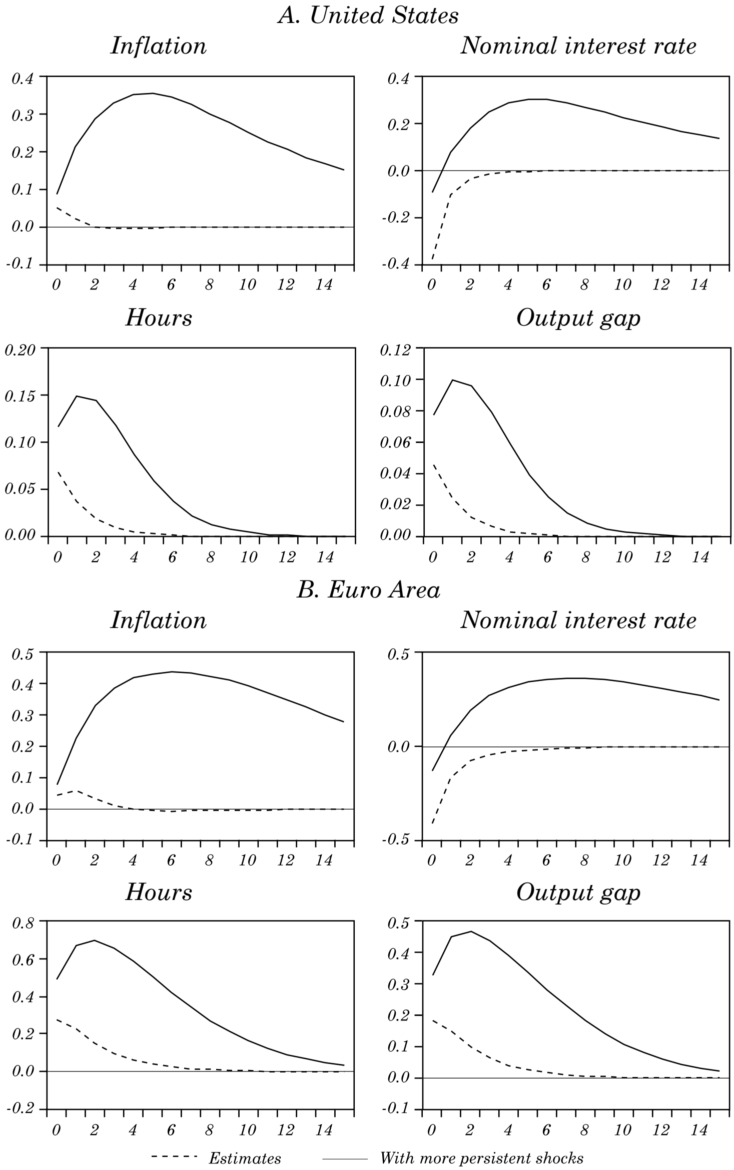
An extensive literature, starting with Taylor (1993), documents that the policy rule in equation (23) provides a good description of policy in the United States and a reasonable description of policy in the euro area. Within this common rule, there is room for differences between the two regions in the parameters of the rule.

According to the estimates in tables 2 and 4, monetary policy has been quite similar in the United States post-1986 and in the euro area post-1993, especially in only modestly responding to real activity. The estimates of ϕ_p and ϕ_y are somewhat lower than the typical result in the literature, but the more surprising posterior mean is the low persistence of monetary policy shocks, especially in the United States.

As noted in section 5, the estimated quick response of most macroeconomic variables to monetary policy shocks is linked to these low estimates of persistence. Figure 3 backs this claim by comparing the impulse responses in the status quo with the responses to raising the persistence of monetary shocks from the posterior means to the prior mean of 0.9. This change reestablishes the conventional delayed hump-shaped responses found in the literature on the post-war United States (Christiano, Eichenbaum, and Evans, 1999).[15]

15. Coibion (2006) first pointed out the role of the persistence of interest rate shocks in delivering hump shapes.

Figure 3. Impulse Response Functions to a More Persistent Monetary Shock



Source: Author's calculations.

7.2 What Is the Role of Policy Announcements?

The past decade has seen an increasing emphasis on transparency in central banking. Part of the argument for transparency is that if the central bank acts predictably, it will reduce confusion and mistakes on the part of private decisionmakers. According to this point of view, if policy shocks must take place, then they should be announced in advance and clearly communicated to the general public. In the context of the SIGE model, this calls for announcing monetary policy shocks a few quarters in advance, so that a large fraction of agents have time to learn of the event in the interim between announcement and action.

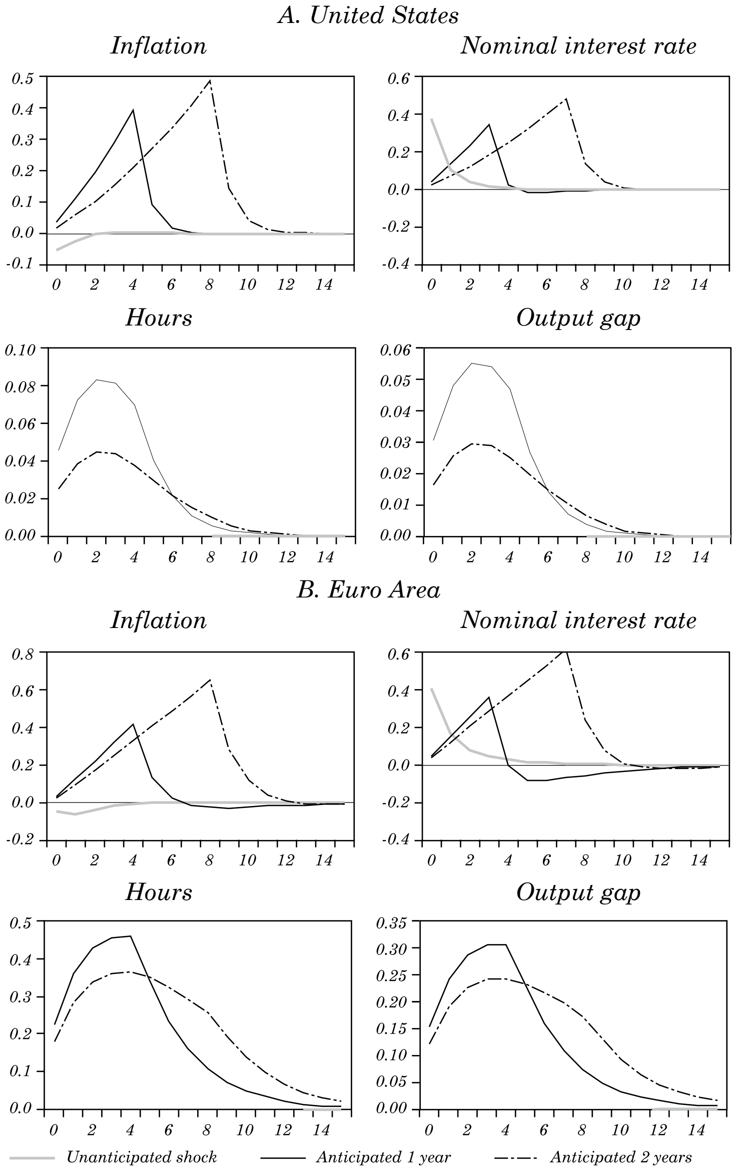
Figure 4 shows the results of announcing a monetary policy shock one or two years ahead in the United States and the euro area. The exercise here consists of learning at date $t = 0$ the value of the monetary shock to occur at dates $t = 4$ or $t = 8$. The announcement is therefore still a shock in the sense of a deviation from the policy rule. The figure reveals that inflation and nominal interest rates move even before the shock materializes because forward-looking agents react instantly to the news of a future shock. The agents that update their information learn about the shocks before it happens and adjust their actions in response. In both regions, announcements lower the initial impact of monetary policy shocks on hours worked and the output gap, while significantly increasing the overall impact on inflation.

7.3 What Is the Result of Having Interest Rates Move Gradually?

As described by Bernanke (2004), the FOMC tends to change interest rates gradually. Academic arguments in favor of such actions typically involve financial stability, the gradual revelation of news, or the desire to move long-term interest rates. Woodford (2003c) notes that in forward-looking models like SIGE, gradualism involves combining policy responses with announcements of future policy changes.

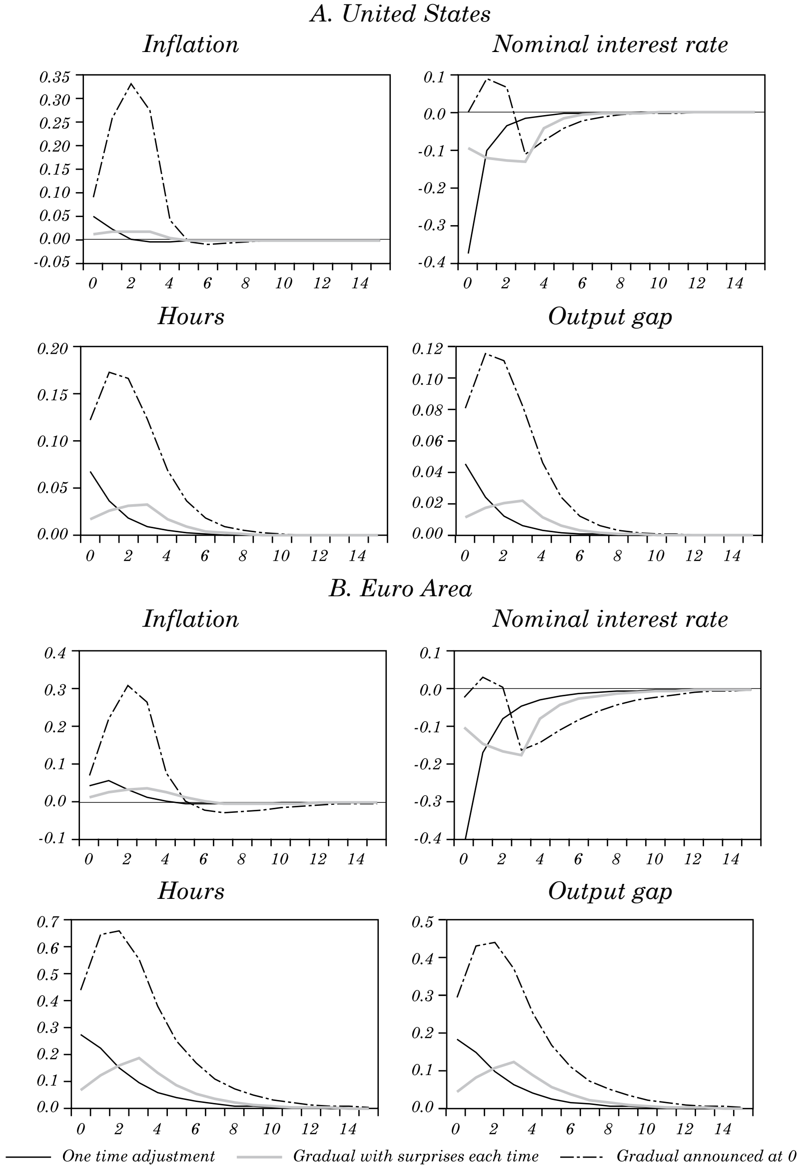
Figure 5 compares three different patterns of shocks for the two regions. In the first case, there is a one-standard-deviation shock to interest rates at date 0. In the second case, there are four consecutive shocks, each of size $\sigma_{\varepsilon}/4$ and each coming as a surprise to the agents. In the third scenario, the sequence of four shocks is announced at date 0. The results indicate that an anticipated gradual cut in interest rates has a much stronger impact than an expected cut of the same size. If the gradual cut is unexpected, however, the impact is actually smaller. Therefore, gradual policy changes can be quite effective according to the SIGE model, but only if they are announced and credible.

Figure 4. Impulse Response Functions to Policy Announcements



Source: Author's calculations.

Figure 5. Impulse Response Functions to Gradual Movements in Policy



Source: Author's calculations.

7.4 How Would Taylor's Proposal Compare?

Taylor (1993) originally suggested that the interest rate responses to inflation and output should be 1.5 and 0.5, respectively. Figure 6 compares this rule with the one estimated here for the impulse responses of inflation and hours worked to productivity and aggregate demand shocks. For both shocks and both regions, Taylor's more aggressive policy rule leads to a smaller response in the output gap to the shock. The unconditional variance of hours worked would fall by 1.3 percent (2.7 percent) if the United States (euro area) moved to this rule, and welfare would be 4 (6) basis points of steady-state consumption higher.

7.5 How Does a Price-Level Target Compare?

Ball, Mankiw, and Reis (2005) show that in an economy with inattentive firms, the optimal policy is an "elastic price standard" that keeps the price level close to a deterministic target K_t, allowing for deviations of the price level from the target in response to deviations of output from the Pareto-optimal level:

$$p_t = K_t - \phi\left(y_t - y_t^o\right). \tag{44}$$

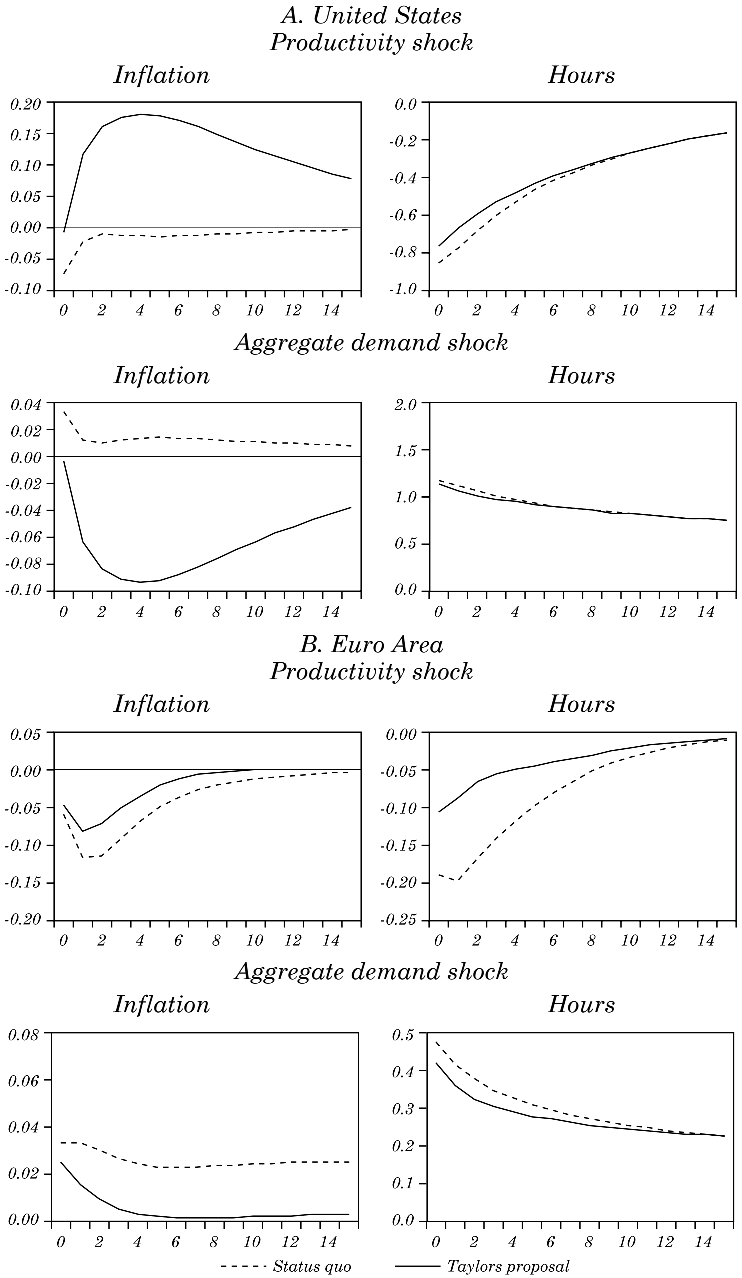
Under this rule, positive deviations of inflation from the target are not bygones, but must be accompanied by future negative deviations to revert the price level back to target.

Figure 7 shows the impulse responses to productivity and aggregate demand shocks of having a strict rule with $\phi = 0$. In the United States, fully stabilizing inflation has little impact on the response of hours worked. The response of hours worked to the markup shocks (not reported) becomes significantly more pronounced, though, so the rule has a negative effect on welfare of 4 basis points on impact. For the euro area, the welfare loss from this rule would be a substantial 17 basis points.

Figure 8 graphs the responses to an elastic rule, where ϕ is set following the guidelines of Ball, Mankiw, and Reis (2005).[16] The ϕ for the United States is 0.12, while that for the euro area is 3.08. Both lead to a slight loss in welfare relative to the Taylor rule with the estimated coefficients.

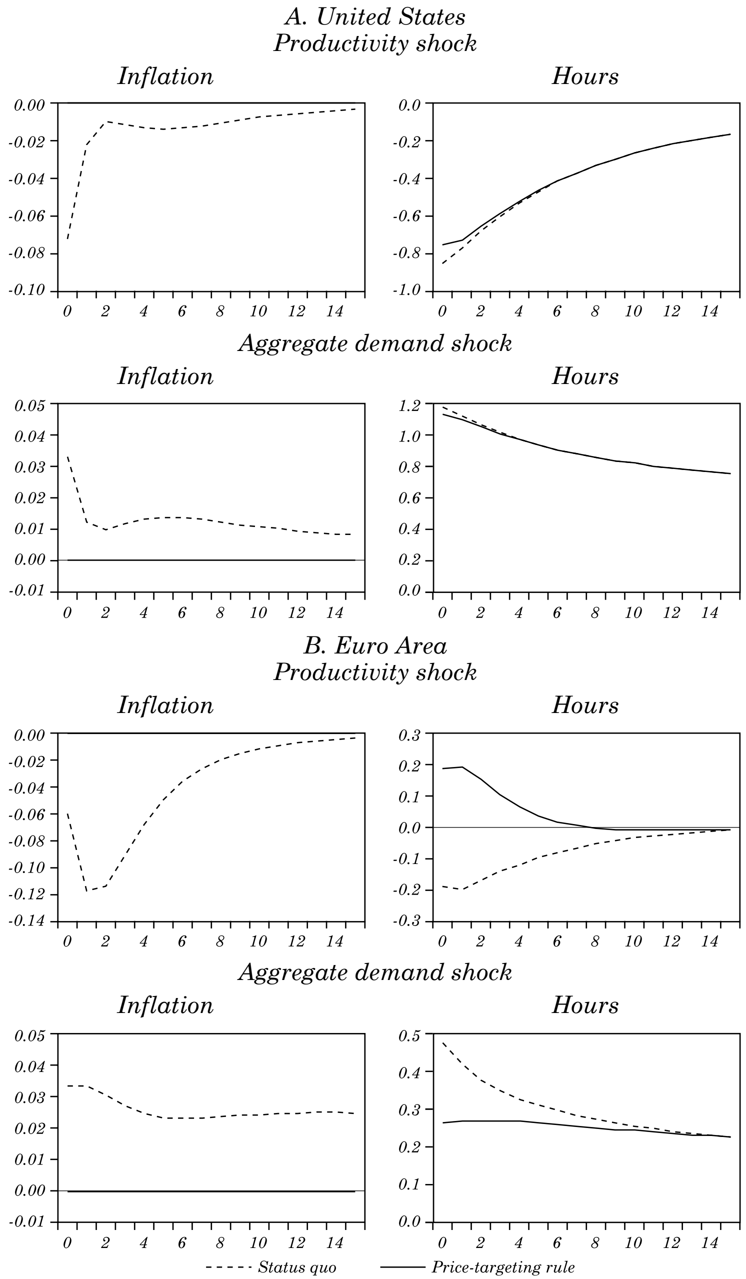
16. More concretely, Ball, Mankiw, and Reis (2005) show that the optimal ϕ is the inverse of the product of $(1 + \psi)/(1 + \psi\nu)$ and the relative weight of relative-price distortions and output-gap fluctuations in the policymaker's objective function. I approximate this relative weight by the ratio of the change in the volatility of the output gap and the change in the volatillity of inflation, both in response to a one-basis-point increase in the standard deviation of all shocks.

Figure 6. Impulse Response Functions with a Taylor Rule



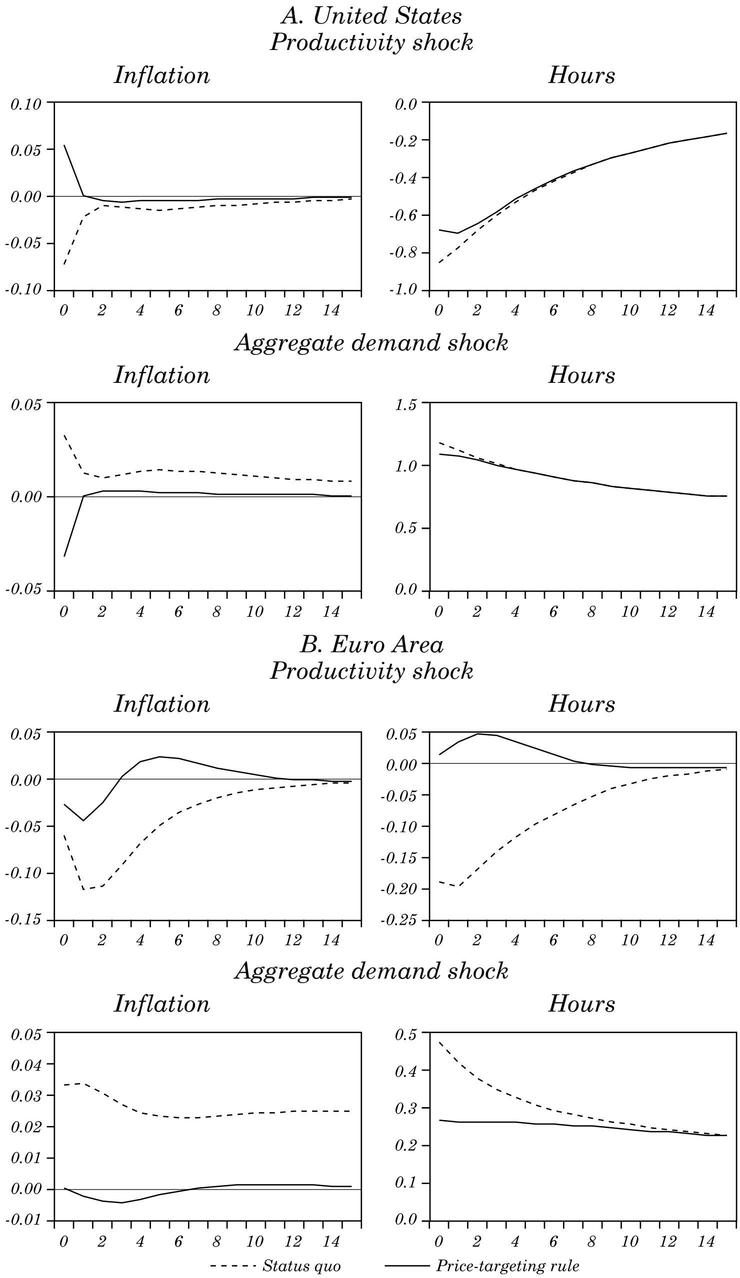
Source: Author's calculations.

Figure 7. Impulse Response Functions with a Strict Price-Level Target



Source: Author's calculations.

Figure 8. Impulse Response Functions with an Elastic Price-Level Target



Source: Author's calculations.

8. Conclusion

The aim of this paper was to build one particular model of the macroeconomy that can be used to give systematic policy advice. The two guiding principles behind the construction of the model were, first, that inattentiveness is a feature of behavior that affects all markets and decisions and, second, that it is the only feature that leads to a deviation from an otherwise classical equilibrium. In reality, many frictions are probably at play, but insisting on a single friction allows one to explore how far inattentiveness alone affects macroeconomic dynamics and policy, while staying within a coherent theoretical framework where in which all details are explicitly stated.

Many of the details of the model, as well as the way in which the parameters were picked, may be open to debate, and there is room for disagreement on how well the model fits the data. I have tried throughout the paper to highlight the theoretical gaps in the model, the different views on how to set its parameters, and the ways in which it succeeded and failed at explaining the data. In the model's defense, it did not seem to perform noticeably worse than some popular alternatives, like the models in Christiano, Eichenbaum, and Evans (2005), Levin, Onatski, Williams and Williams (2006), or Smets and Wouters (2003, 2007).

While the model's performance is probably still far from the level of success one should demand to confidently give precise policy recommendations, the exercise did provide some policy lessons. First, the persistence of monetary policy shocks has been low, and this is a crucial determinant of the speed at which inflation and output respond to these shocks. Second, announcements and gradualism, through their effects on the expectations of forward-looking agents, can have a large impact on the effects of monetary policy. Third, Taylor's suggested policy rule parameters would lead to better outcomes than the status quo, while an elastic price standard has a disappointing performance when inattentiveness is pervasive.

Appendix

A1. Inattentive Actions

Planner-savers, who every period face a probability δ of revising their plans, have a value function $V(M_t)$ conditional on date t being a planning date. They choose a plan for current and future consumption all the way into infinity $\{C_{t+l,l}\}_{l=0}^{\infty}$ since with a vanishingly small probability they may never update again:

$$V\left(M_t\right) = \max_{\{C_{t+l,l}\}} \left\{ \begin{array}{l} \sum_{l=0}^{\infty} \xi^l \left(1-\delta\right)^l \frac{C_{t+l,l}^{1-1/\theta}}{1-1/\theta} \\ +\xi\delta \sum_{l=0}^{\infty} \xi^l \left(1-\delta\right)^l E_t\left[V\left(M_{t+1+l}\right)\right] \end{array} \right\}, \tag{45}$$

subject to the sequence of budget constraints in equation (9) and a no-Ponzi condition.

The optimality conditions are

$$\xi^l \left(1-\delta\right)^l C_{t+l,l}^{-1/\theta} = \xi\delta \sum_{k=l}^{\infty} \xi^k \left(1-\delta\right)^k E_t\left[V'\left(M_{t+1+k}\right)\bar{\Pi}_{t+l,t+1+k}\right] \tag{46}$$

and

$$V'\left(M_t\right) = \xi\delta \sum_{l=0}^{\infty} \xi^l \left(1-\delta\right)^l E_t\left[V'\left(M_{t+1+l}\right)\bar{\Pi}_{t,t+1+l}\right], \tag{47}$$

where

$$\bar{\Pi}_{t+l,t+1+k} = \prod_{z=t+l}^{t+k} \Pi_{z+1}$$

is the the compound return between $t + l$ and $t + 1 + k$ for $k > l$. Now, for $l = 0$, the right-hand side of equation (46) is the same as the right-hand side of equation (47). Therefore, $C_{t,0}^{-1/\theta} = V'\left(M_t\right)$, or the marginal utility of an extra unit of consumption equals the marginal value of an extra unit of wealth. Using this result to replace the $V'(M_{t+1+l})$ terms in equation (47) and writing the equation recursively gives the Euler equation in equation (10). The second Euler equation in equation (11) then follows.

The workers face a similar problem:

$$\hat{V}(M_t) = \max_{\{W_{t+l,l}\}} \left\{ \begin{array}{l} -\sum_{l=0}^{\infty} \xi^l (1-\omega)^l \kappa E_t \left(\dfrac{L_{t+l,l}^{1+1/\psi} + 1}{1+1/\psi} \right) \\ +\xi\omega \sum_{l=0}^{\infty} \xi^l (1-\omega)^l E_t \left[\hat{V}\left(M_{t+1+l} \right) \right] \end{array} \right\}, \tag{48}$$

subject to the sequence of budget constraints in equation (9), a no-Ponzi scheme condition, and the demand for the variety of labor *j* in equation (14), which each worker supplies monopolistically. The optimality conditions are

$$\begin{aligned} &\frac{\xi^l (1-\omega)^l \kappa E_t (\tilde{\gamma}_{t+l} L_{t+l,l}^{1+1/\psi})(1-\tau_w)}{W_{t+l,l}} = \\ &\xi\omega \sum_{k=l}^{\infty} \xi^k (1-\omega)^k E_t \left[\frac{V'(M_{t+1+k}) \bar{\Pi}_{t+l,t+1+k} (\tilde{\gamma}_{t+l} - 1) L_{t+l,l}}{P_{t+l}} \right] \end{aligned} \tag{49}$$

and

$$\hat{V}'\left(M_t\right) = \xi\omega \sum_{k=0}^{\infty} \xi^k (1-\omega)^k E_t \left[\hat{V}'\left(M_{t+1+k}\right) \bar{\Pi}_{t,t+1+k} \right]. \tag{50}$$

Now, as in the consumer problem, combining equation (49) for $l = 0$ with equation (50) leads to the following conclusion:

$$\frac{\hat{V}_t'(M_t) W_{t,0}}{P_t} = \frac{(1-\tau_w) \tilde{\gamma}_t \kappa L_{t,0}^{1/\psi}}{\tilde{\gamma}_t - 1}. \tag{51}$$

This expression shows that ψ is the Frisch elasticity of labor supply for attentive agents and that the marginal disutility of working is equated to the real wage rate times the marginal value of wealth times a markup taking into account the elasticity of demand for the good. Using it in the optimality condition leads to the two Euler equations in equations (16) and (17).

A2. The Log-Linear Equilibrium for the Full Model

At the nonstochastic steady state, the five exogenous processes are constant. Using the conditions defining the optimum, it follows that output is $Y = AL^{\beta}$, consumption is $C = Y/G$, and labor is

$$\kappa L^{1+1/\psi} = \frac{\beta G(\nu - 1)(\gamma - 1)}{(1-\tau_w)(1-\tau_p)\nu\gamma}. \tag{52}$$

I log-linearize the equilibrium conditions around this point. Small caps denote the log-deviations of the respective large-cap variable from the steady state, with the exceptions of the following: v_t and γ_t, which are the log-deviations of $\tilde{\nu}_t$ and $\tilde{\gamma}_t$; r_t, which is the log-deviation of the short rate $E_t[\Pi_{t+1}]$; and R_t, which is the log-deviation of the long rate $\lim_{k\to\infty} E_t[\overline{\Pi}_{t,t+1+k}]$.

Starting with the goods market, log-linearizing the demand for good j by combining equations (3) and (6) gives

$$y_{t,i} = y_t - \nu\,(p_{t,i} - p_t). \tag{53}$$

The production function (5) and the firm's optimality condition (7) become

$$y_{t,i} = \alpha_t + \beta l_{t,i} \tag{54}$$

and

$$p_{t,i} = E_{t-i}\left[p_t + \frac{\beta(w_t - p_t) + (1-\beta)y_t - a_t - \nu_t\beta/(\nu - 1)}{\beta + \nu(1-\beta)}\right]. \tag{55}$$

Turning to the bond market, the consumer's Euler equations in equations (10) and (11) become

$$c_{t,0} = E_t(c_{t+1,0} - \theta r_t) \tag{56}$$

and

$$c_{t,j} = E_{t-j}(c_{t,0}) \tag{57}$$

Next, in the labor market, the demand for a labor variety in equation (14), together with the market clearing condition in this market, leads to:

$$l_{t,k} = l_t - \gamma\left(w_{t,k} - w_t\right), \tag{58}$$

and the optimality conditions in the workers' problem become

$$w_{t,0} - p_t - \frac{l_{t,0}}{\psi} + \frac{\gamma_t}{\gamma - 1} = E_t\left(-r_t + w_{t+1,0} - p_{t+1} - \frac{l_{t+1,0}}{\psi} + \frac{\gamma_{t+1}}{\gamma - 1}\right) \tag{59}$$

and

$$w_{t,k} = E_{t-k}(w_{t,0}) \tag{60}$$

Finally, the static price indices and aggregate quantity are

$$p_t = \lambda\sum_{i=0}^{\infty}\left(1 - \lambda\right)^i p_{t,i}, \tag{61}$$

$$w_t = \omega\sum_{k=0}^{\infty}\left(1 - \omega\right)^k w_{t,k}, \tag{62}$$

and

$$y_t = g_t + \delta\sum_{j=0}^{\infty}\left(1 - \delta\right)^j c_{t,j}. \tag{63}$$

These eleven equations over time characterize the equilibrium solution for the set of twelve variables ($y_{t,i}$, y_t, $c_{t,0}$, $c_{t,j}$, $l_{t,0}$, $l_{t,k}$, l_t, $w_{t,k}$, w_t, p_t, $p_{t,i}$, r_t) as a function of the five exogenous processes (Δa_t, g_t, γ_t, ν_t, ε_t). There is one equation missing, namely, the policy rule in equation (23).

A3. The Reduced-Form Aggregate Relations

Integrating equation (54) over i gives the aggregate production function in equation (19).

For the Phillips curve, starting with equation (61), replace $y_{t,j}$ using equation (53) and $p_{t,i}$ using equation (55). Rearrange to obtain equation (20).

Moving to the IS curve, iterate equation (56) forward and take the limit as time goes to infinity. Then, the facts that there is complete insurance and that eventually all agents become aware of the shocks imply that $\lim_{\tau\to\infty} E_t\left(c_{t+\tau,0}\right) = \lim_{\tau\to\infty} E_t\left(y_{t+\tau}\right) \equiv y_t^{\infty}$. Using the definition of the long rate R_t and replacing for $c_{t,0}$ in equations (57) and (63) gives an expression for output. Using the fact that $\lim_{\tau\to\infty} E_t\left[g_{t+\tau}\right] = 0$ gives the IS curve in equation (21).

Finally, for the wage curve, take very similar steps as in the IS curve: iterate equation (59) forward and use the solution to replace $w_{t,0}$ in equation (60). Combining the $w_{t,j}$ in the aggregator for w_t in equation (62) and replacing out $l_{t,j}$ using equation (58) gives the wage curve in equation (22).

A4. Proof of Propositions 1 and 2

Take the case of $s = a$. By a method of undetermined coefficients, equations (19) through (22) imply[17]

$$\hat{y}_n = \hat{a}_n + \beta \hat{l}_n; \tag{64}$$

$$\hat{p}_n = \Lambda_n \left[\hat{p}_n + \frac{\beta \hat{w}_n + (1-\beta)\hat{y}_n - \hat{a}_n - \beta \nu_n / (\nu - 1)}{\beta + \nu(1-\beta)} \right]; \tag{65}$$

$$\hat{r}_n = \frac{\hat{y}_{n+1}}{\theta \Delta_{n+1}} - \frac{\hat{y}_n}{\theta \Delta_n}; \tag{66}$$

$$(\gamma + \psi)\hat{w}_n = \Omega_n \left[(\psi + \gamma)\hat{p}_n + \gamma(\hat{w}_n - \hat{p}_n) + \hat{l}_n + \frac{\psi \hat{y}_n}{\theta \Delta_n} \right]. \tag{67}$$

Rearranging the first three equations immediately proves proposition 2. Using the first two expressions to replace $\hat{l}_n$ and $\hat{w}_n$ in the fourth expression proves proposition 1. The case of the other four shocks follows along the same lines.

17. I have omitted the (s) arguments to save space.

A5. Proof of Proposition 3

Taking again the case $s = a$, combining the Taylor rule with the Fisher equation, and again omitting the (s) arguments, the undetermined coefficients are

$$\hat{r}_n + \hat{p}_{n+1} - \hat{p}_n = \phi_p(\hat{p}_n - \hat{p}_{n-1}) + \phi_y\,(\hat{y}_n - \Xi_n\hat{s}_n).$$

Using the results in propositions 1 and 2 to replace $\hat{r}_n$ and $\hat{y}_n$ and rearranging delivers the proposition. The other cases are similar.

A6. Proof of Proposition 4

Since the K_t is known to all agents, real variables are neutral with respect to it, and it only induces a deterministic component in prices. Focusing on the stochastic component, in terms of moving-average coefficients, the policy rule implies that

$$\hat{p}_n = \phi\,(\hat{y}_n - \tilde{\Xi}_n\hat{s}_n).$$

Using the expression in proposition 1 to replace $\hat{y}_n$ delivers the result.

A7. Solutions for Other Interest Rate Rules

The proofs for the case of these rules follow along the same lines as propositions 3 and 4 so they are omitted. First, consider alternative interest rate rules:

Proposition 7. If policy follows the interest rate rules below, the undetermined coefficients for the price level satisfy the following second-order difference equation:

$$A_{n+1}\hat{p}_{n+1}(s) - B_n\hat{p}_n(s) + C_{n-1}\hat{p}_{n-1}(s) = D_n(s) \quad \text{for } n = 0,1,2,... \tag{68}$$

with $A_n = 1 + \Psi_n / \theta\Delta_n$ and $D_n(\varepsilon) = -1$ for all cases. The remaining coefficients are as follows:

—For the employment rule, $i_t = \phi_p\Delta p_t + \phi_y l_t$,

$$B_n = A_n + \phi_p + \frac{\phi_y\Psi_n}{\beta},\ C_n = \phi_p \tag{69}$$

and

$$D_n(s)=\begin{cases}\dfrac{\Upsilon_n(s)}{\theta\Delta_n}-\dfrac{\Upsilon_{n+1}(s)\hat{s}_{n+1}}{\theta\Delta_{n+1}\hat{s}_n}+\dfrac{\phi_y\left[\Upsilon_n(s)-1\right]}{\beta} & \text{for } s=a\\ \dfrac{(\Upsilon_n(s)-1)}{\theta\Delta_n}-\dfrac{\left[\Upsilon_{n+1}(s)-1\right]\hat{s}_{n+1}}{\theta\Delta_{n+1}\hat{s}_n}+\dfrac{\phi_y\Upsilon_n(s)}{\beta} & \text{for } s=g\\ \dfrac{\Upsilon_n(s)}{\theta\Delta_n}-\dfrac{\Upsilon_{n+1}(s)\hat{s}_{n+1}}{\theta\Delta_{n+1}\hat{s}_n}+\dfrac{\phi_y\Upsilon_n(s)}{\beta} & \text{for } s=\gamma,\nu\end{cases}\tag{70}$$

—For the speed-limit rule, $i_t = \phi_p \Delta p_t + \phi_y \Delta(y_t - y_t^c)$,

$$B_n = A_n + \phi_p + \phi_y \Psi_n,\ C_n = \phi_p + \phi_y \Psi_n\tag{71}$$

and

$$D_n(s)=\begin{cases}\dfrac{\Upsilon_n(s)}{\theta\Delta_n}-\dfrac{\Upsilon_{n+1}(s)\hat{s}_{n+1}}{\theta\Delta_{n+1}\hat{s}_n} & \\ +\phi_y\left\{\Upsilon_n(s)-\Xi(s)-\dfrac{\left[\Upsilon_{n-1}(s)-\Xi(s)\right]\hat{s}_{n-1}}{\hat{s}_n}\right\} & \text{for } s=a,\gamma,\nu\\ \dfrac{\left[\Upsilon_n(s)-1\right]}{\theta\Delta_n}-\dfrac{\left[\Upsilon_{n+1}(s)-1\right]\hat{s}_{n+1}}{\theta\Delta_{n+1}\hat{s}_n} & \\ +\phi_y\left\{\Upsilon_n(s)-\Xi(s)-\dfrac{\left[\Upsilon_{n-1}(s)-\Xi(s)\right]\hat{s}_{n-1}}{\hat{s}_n}\right\} & \text{for } s=g\end{cases}\tag{72}$$

—For the inertial rule, $i_t = (1 - \phi_i)[\phi_p \Delta p_t + \phi_y (y_t - y_t^c)] + \phi_i\, i_{t-1}$,

$$B_0 = A_0 + (1-\phi_i)\phi_p + (1-\phi_i)\phi_y\Upsilon_0(s),\tag{73}$$

$$B_n = A_n(1+\phi_i) + (1-\phi_i)\phi_p + (1-\phi_i)\phi_y\Upsilon_0(s),\quad n \geq 1,\tag{74}$$

$$C_n = (1-\phi_i)\phi_p + \phi_i A_n,\tag{75}$$

and

$$D_n(s)=\begin{cases}\dfrac{\Upsilon_n(s)(1-\phi_i)}{\theta\Delta_n}-\dfrac{\Upsilon_{n+1}(s)\hat{s}_{n+1}}{\theta\Delta_{n+1}\hat{s}_n}-\dfrac{\phi_i\Upsilon_{n-1}(s)\hat{s}_{n-1}}{\theta\Delta_{n-1}\hat{s}_n} & \text{for } s=a,\gamma,\nu\\ \quad+\phi_y(1-\phi_i)(\Upsilon_n(s)-\Xi(s)) & \\ \dfrac{[\Upsilon_n(s)-1](1-\phi_i)}{\theta\Delta_n}-\dfrac{[\Upsilon_{n+1}(s)-1]\hat{s}_{n+1}}{\theta\Delta_{n+1}\hat{s}_n}-\dfrac{\phi_i[\Upsilon_{n-1}(s)-1]\hat{s}_{n-1}}{\theta\Delta_{n-1}\hat{s}_n} & \text{for } s=g\\ \quad+\phi_y(1-\phi_i)[\Upsilon_n(s)-\Xi(s)] & \end{cases} \tag{76}$$

—For the wage-inflation rule $i_t=\phi_p\Delta w_t+\phi_y(y_t-y_t^c)$,

$$\begin{aligned}B_n &= A_n+\phi_p\left[\left[1+\nu\left(\frac{1}{\beta}-1\right)\right]\left(\frac{1}{\Lambda_n}-1\right)+\left(1-\frac{1}{\beta}\right)\Psi_n\right]+\phi_y\Psi_n,\\ C_n &= \phi_p\left[\left[1+\nu\left(\frac{1}{\beta}-1\right)\right]\left(\frac{1}{\Lambda_n}-1\right)+\left(1-\frac{1}{\beta}\right)\Psi_n\right]\end{aligned} \tag{77}$$

and

$$D_n(s)=\begin{cases}\dfrac{\Upsilon_n(s)}{\theta\Delta_n}-\dfrac{\Upsilon_{n+1}(s)\hat{s}_{n+1}}{\theta\Delta_{n+1}\hat{s}_n}+\phi_y[\Upsilon_n(s)-\Xi(s)]+\phi_p\left(1-\dfrac{1}{\beta}\right)\left[\Upsilon_n(s)-\dfrac{\Upsilon_{n-1}(s)\hat{s}_{n-1}}{\hat{s}_n}\right] & \text{for } s=a\\ \dfrac{[\Upsilon_n(s)-1]}{\theta\Delta_n}-\dfrac{[\Upsilon_{n+1}(s)-1]\hat{s}_{n+1}}{\theta\Delta_{n+1}\hat{s}_n}+\phi_y[\Upsilon_n(s)-\Xi(s)] & \\ \quad+\phi_p\left(1-\dfrac{1}{\beta}\right)\left[\Upsilon_n(s)-\dfrac{\Upsilon_{n-1}(s)\hat{s}_{n-1}}{\hat{s}_n}\right] & \text{for } s=g\\ \dfrac{\Upsilon_n(s)}{\theta\Delta_n}-\dfrac{\Upsilon_{n+1}(s)\hat{s}_{n+1}}{\theta\Delta_{n+1}\hat{s}_n}+\phi_y[\Upsilon_n(s)-\Xi(s)]+\phi_p\left(1-\dfrac{1}{\beta}\right)\left[\Upsilon_n(s)-\dfrac{\Upsilon_{n-1}(s)\hat{s}_{n-1}}{\hat{s}_n}\right] & \text{for } s=\gamma\\ \dfrac{\Upsilon_n(s)}{\theta\Delta_n}-\dfrac{\Upsilon_{n+1}(s)\hat{s}_{n+1}}{\theta\Delta_{n+1}\hat{s}_n}+\phi_y[\Upsilon_n(s)-\Xi(s)]+\phi_p\left(1-\dfrac{1}{\beta}\right)\left[\Upsilon_n(s)-\dfrac{\Upsilon_{n-1}(s)\hat{s}_{n-1}}{\hat{s}_n}\right] & \\ \quad+\dfrac{\phi_p}{\nu-1}\left(1-\dfrac{\hat{s}_{n-1}}{\hat{s}_n}\right) & \text{for } s=\nu\end{cases} \tag{78}$$

Finally, consider alternative price-targeting rules,

Proposition 8. If the policy rule follows other price-level standards, the undetermined coefficients for the price level are as follows:

—With an employment rule, $p_t=K_t-\phi l_t$,

$$\hat{p}_n(s) = \begin{cases} \dfrac{\phi\left[1-\Upsilon_n(s)\right]\hat{s}_n}{\beta+\phi\Psi_n} & \text{for } s = a \\ \dfrac{-\phi\Upsilon_n(s)\hat{s}_n}{\beta+\phi\Psi_n} & \text{for } s = g,\gamma,\nu \end{cases} \tag{79}$$

for n = 0, 1, 2,….

—With a speed-limit rule, $p_t = K_t - \phi\Delta(y_t - y_t^o)$,

$$(1+\phi\Psi_n)\hat{p}_n(s) - \phi\Psi_{n-1}\hat{p}_{n-1}(s) = \phi\left\{\begin{matrix}\left[\Xi(s)-\Upsilon_n(s)\right]s_n \\ -\left[\Xi(s)-\Upsilon_{n-1}(s)\right]s_{n-1}\end{matrix}\right\}, \tag{80}$$

for n = 0, 1, 2,… and with $\hat{p}_{-1}(s) = 0$.

—With an inertial rule, $p_t = K_t - \phi(y_t - y_t^o) + \phi_p p_{t-1}$,

$$(1+\phi\Psi_n)\hat{p}_n(s) - \phi_p\hat{p}_{n-1}(s) = \phi[\Xi(s) - \Upsilon_n(s)]s_n \tag{81}$$

for n = 0, 1, 2,… and with $\hat{p}_{-1}(s) = 0$.

—With a wage-targeting rule, $w_n = K_n - \phi(y_t - y_t^o)$,

$$\hat{p}_n(s) = \frac{\left\{\phi\left[\tilde{\Xi}(s)-\Upsilon_n(s)\right] - \left(\dfrac{1}{\beta}-1\right)\Upsilon_n(s)\right\}\hat{s}_n - \begin{cases} \dfrac{\hat{s}_n}{\beta} & \text{for } s = a \\ \dfrac{\hat{s}_n}{\nu-1} & \text{for } s = \nu \\ 0 & \text{for } s = g,\gamma \end{cases}}{1+\left[1+\nu\left(\dfrac{1}{\beta}-1\right)\right]\left(\dfrac{1}{\Lambda_n}-1\right)+\left(1-\dfrac{1}{\beta}\right)\Psi_n+\phi\Psi_n} \tag{82}$$

for n = 0, 1, 2,….

A8. Proof of Proposition 5

Since $\mathbf{X}_t$ is a sum of multivariate normal distributions it is also multivariate normal. Its mean is a column vector of zeros, and its variance-covariance matrix is $\boldsymbol{\Omega}(\mathbf{I}_N \otimes \boldsymbol{\Sigma})\boldsymbol{\Omega}'$. Using the formula for the density of a multivariate normal, the result in the proposition follows immediately.

A9. Proof of Proposition 6

Taking the unconditional expectation through the arguments of expression (38), the goal is to maximize the folowing expression:

$$\int_0^1 \left\{ E\left[\ln\left(C_{t,j}\right)\right] - \frac{\kappa E\left(L_{t,j}^{1+1/\psi}\right)}{1+1/\psi} \right\} dj. \tag{83}$$

With the definition of the log-linearized values, $c_{t,j} = \ln(C_{t,j}) - \ln(C)$ and $l_{t,j} = \ln(L_{t,j}) - \ln(L)$, this becomes

$$\ln\left(C\right) + \int_0^1 \left[E\left(c_{t,j}\right) - \frac{\kappa L^{1+1/\psi} E\left(e^{(1+1/\psi) l_{t,j}}\right)}{1+1/\psi} \right] dj. \tag{84}$$

Recall that the model assumed that the tax on prices exactly offsets the monopoly distortion in the goods market: $1 - \tau_p = \nu/(\nu - 1)$; the tax on wages exactly offsets the monopoly distortion in the goods market: $1 - \tau_w = \gamma/(\gamma - 1)$; and the distortion from government spending is, on average, zero: $G = 1$. In this case, the nonstochastic steady state is an efficient equilibrium without uncertainty. These assumptions lead to focusing monetary policy on the task of stabilizing economic activity (Woodford, 2003b). From equation (52), they imply that $\kappa L^{1+1/\psi} = \beta$.

In the log-linear solution of the model, both $c_{t,j}$ and $l_{t,j}$ are normal variables with a zero mean. Therefore, social welfare is

$$\ln\left(C\right) - \frac{\beta}{1+1/\psi} \int_0^1 \exp\left[0.5\left(1+\frac{1}{\psi}\right)^2 \operatorname{var}\left(l_{t,j}\right)\right] dj. \tag{85}$$

Because $l_{t,j}$ is a normal variable, var($l_{t,j}$) is a linear function of the variance of the exogenous shocks. These are small in the data, so approximating exp[var($l_{t,j}$)] by 1 + var($l_{t,j}$) involves little numerical error. Social welfare then becomes:

$$\ln(C)+\beta\left(1+\frac{1}{\psi}\right)-0.5\beta\left(1+\frac{1}{\psi}\right)\int_0^1 \text{var}\left(l_{t,j}\right)dj. \tag{86}$$

Using the distribution of workers according to when they last updated, this becomes

$$\ln(C)+\beta\left(1+\frac{1}{\psi}\right)-0.5\beta\left(1+\frac{1}{\psi}\right)\omega\sum_{j=0}^{\infty}(1-\omega)^j \text{var}\left(l_{t,j}\right) \tag{87}$$

Next, combining equation (58) with equations (59) and (60) to replace $w_{t,0}$ gives the following expressions:

$$l_{t,j} = l_t - \gamma\left(w_{t,j} - w_t\right) \tag{88}$$

and

$$w_{t,j} = E_{t-j}\left(p_t + \frac{l_{t,j}}{\psi} - \frac{\gamma_t}{\gamma-1} - R_t + y_n^{\infty}\right). \tag{89}$$

Using a method of undetermined coefficients, make the guess that $l_{t,j} = \sum_{s\in S}\left[\sum_{n=0}^{j-1}\varsigma_n(s) + \sum_{n=j}^{\infty}\zeta_n(s)\right]e_{t-n}^s$ and solve to find the expressions in equations (41) and (42). From this, it follows that

$$\text{var}\left(l_{t,j}\right) = \sum_{s\in S}\left[\sum_{n=0}^{j-1}\varsigma_n(s)^2 + \sum_{n=j}^{\infty}\zeta_n(s)^2\right]\sigma^2(s). \tag{90}$$

Finally, some grouping shows that

$$\omega\sum_{j=0}^{\infty}(1-\omega)^j\left[\sum_{n=0}^{j-1}\varsigma_n(s)^2 + \sum_{n=j}^{\infty}\zeta_n(s)^2\right] = \sum_{n=0}^{\infty}\left[\begin{array}{l}(1-\Omega_n)\varsigma_n(s)^2\\ +\Omega_n\zeta_n(s)^2\end{array}\right], \tag{91}$$

where $\Omega_n = \omega\sum_{i=0}^{n}(1-\omega)^i$. Ignoring the terms that are invariant to policy changes, the social welfare function then becomes the expression in equation (40). To evaluate the welfare benefit in percentage units of steady-state consumption of a policy that implies $\boldsymbol{\theta}^{(1)}$ starting from another that implies $\boldsymbol{\theta}^{(0)}$, use equation (87) to obtain equation (39).

A10. MCMC Algorithm

I used a Metropolis-Hastings algorithm to draw from the posterior. In the first step, I looked for the mode of the posterior distribution by using line-search and Newton-Raphson algorithms starting from twenty different points on the parameter space (chosen from previous estimates of similar models and from drawing randomly from either the prior or a uniform on the parameter space). In the second step, I used a mixture of normal approximations around the highest local maxima found, to obtain an approximation of the posterior. This is then used as the proposal function for the Metropolis-Hastings algorithm. In the third step, I took a few sequences of 2,000 draws, scaling the variance-covariance matrix of the proposal function by different values, until the acceptance rates of the Metropolis-Hastings algorithm are 10–20 percent.

In the fourth step, I took 5 independent sequences of 200,000 draws, discarding the first 100,000. Inspecting the 500,000 mixed draws made clear that the algorithm was far from converging, and that the normal approximation of the posterior was poor. I therefore revised the proposal function to a normal distribution with a variance-covariance matrix equal to the scaled estimate of the variance-covariance matrix of the existing 500,000 draws.

In the fifth step, I took five independent sequences of 1,000,000 draws, discarding the first 100,000 draws and keeping only every tenth draw to save on memory space. The Brooks-Gelman scale reduction factors and the plots of the between-chain and within-chain variances indicated that the results were satisfactory in terms of convergence, so I proceeded to mix them to obtain the final 450,000 draws of the posterior, which are used in all the tables.

References

Abel, A.B., J.C. Eberly and S. Panageas. 2007. "Optimal Inattention to the Stock Market." *American Economic Review* 97(2): 244–49.

Adam, K. 2007. "Optimal Monetary Policy with Imperfect Common Knowledge." *Journal of Monetary Economics* 54(2): 276–301.

Amato, J.D. and H.S. Shin. 2006. "Imperfect Common Knowledge and the Information Value of Prices." *Economic Theory* 27: 213–41.

An, S. and F. Schorfheide. 2007. "Bayesian Analysis of DSGE Models." *Econometric Reviews* 26(2–4): 113–72.

Andrés, J., E. Nelson and D. López-Salido. 2005. "Sticky-Price Models and the Natural Rate Hypothesis." *Journal of Monetary Economics* 52(5), 1025–53.

Bacchetta, P. and E. van Wincoop. 2006. "Incomplete Information Processing: A Solution to the Forward Discount Puzzle." Working paper 2006-35. Federal Reserve Bank of San Francisco.

Ball, L., N.G. Mankiw and R. Reis. 2005. "Monetary Policy for Inattentive Economies." *Journal of Monetary Economics* 52(4): 703–25.

Bayoumi, T. 2004. "GEM: A New International Macroeconomic Model." Occasional paper 239. Washington: International Monetary Fund.

Bernanke, B. 2004. "Gradualism." Remarks at an economics luncheon cosponsored by the Federal Reserve Bank of San Francisco (Seattle Branch) and the University of Washington. Seattle, Washington, 20 May 2004.

Boivin, J. and M. Giannoni. 2006. "Has Monetary Policy Become More Effective?" *Review of Economics and Statistics* 88(3): 445–62.

Branch, W.A. 2007. "Sticky Information and Model Uncertainty in Survey Data on Inflation Expectations." *Journal of Economic Dynamics and Control* 31(1): 245–76.

Carroll, C.D. 2003. "Macroeconomic Expectations of Households and Professional Forecasters." *Quarterly Journal of Economics* 118: 269–98.

Carroll, C.D. and J. Slacalek. 2006. "Sticky Expectations and Consumption Dynamics." Johns Hopkins University.

Christiano, L.J., M. Eichenbaum and C.L. Evans. 1999. "Monetary Policy Shocks: What Have We Learned and to What End?" In *Handbook of Macroeconomics,* edited by J. Taylor and M. Woodford. Elsevier.

———. 2005. "Nominal Rigidities and the Dynamic Effects of an Shock to Monetary Policy." *Journal of Political Economy* 113(1): 1–45.

Coibion, O. 2006. "Inflation Inertia in Sticky Information Models." *Contributions to Macroeconomics* 6(1): 1374–74.

———. 2007. "Testing the Sticky Information Phillips Curve." Working paper 61. College of William and Mary.

Dopke, J., J. Dovern, U. Fritsche and J. Slacalek. 2006. "Sticky Information Phillips Curves: European Evidence." Discussion paper 615. DIW Berlin.

———. 2008. "The Dynamics of European Inflation Expectations." *Topics in Macroeconomics* 8(1): 1540.

Erceg, C.J., L. Guerrieri and C. Gust. 2006. "SIGMA: A New Open Economy Model for Policy Analysis." *International Journal of Central Banking* 2(1): 1–50.

Gabaix, X. and D. Laibson. 2002. "The 6D Bias and the Equity Premium Puzzle." In *NBER Macroeconomics Annual 2001*, edited by B. Bernanke and K. Rogoff, 257–312. MIT Press.

———. 2006. "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics* 121(2): 505–40.

Galí, J. 2004. "On the Role of Technology Shocks as a Source of Business Cycles: Some New Evidence." *Journal of the European Economic Association* 2(2–3): 372–80.

Gorodnichenko, Y. 2006. "Monetary Policy and Forecast Dispersion: A Test of the Sticky Information Model." University of Michigan.

Hall, R.E. 1997. "Macroeconomic Fluctuations and the Allocation of Time." *Journal of Labor Economics* 15(1): S223–S250.

Hellwig, C. 2002. "Public Announcements, Adjustment Delays, and the Business Cycle." Economics online papers 208. University of California at Los Angeles.

Hellwig, C. and L. Veldkamp. 2008. "Knowing What Others Know: Coordination Motives in Information Acquisition." *Review of Economic Studies* (forthcoming).

Khan, H. and Z. Zhu. 2006. "Estimates of the Sticky-Information Phillips Curve for the United States." *Journal of Money, Credit, and Banking* 38(1): 195–207.

Kiley, M.T. 2007. "A Quantitative Comparison of Sticky-Price and Sticky-Information Models of Price Setting." *Journal of Money, Credit, and Banking* 39(s1): 101–25.

Klenow, P.J. and J.L. Willis. 2007. “Sticky Information and Sticky Prices.” *Journal of Monetary Economics* 54(sup. 1): 79–99.

Knotek, E.S. 2006. “A Tale of Two Rigidities: Sticky Prices in a Sticky-Information Environment.” Research working paper 06-15. Federal Reserve Bank of Kansas City.

Korenok, O. 2005. “Empirical Comparison of Sticky Price and Sticky Information Models.” Working paper 501. Virginia Commonwealth University, School of Business.

Korenok, O. and N.R. Swanson. 2005. “The Incremental Predictive Information Associated with Using Theoretical New Keynesian DSGE Models vs. Simple Linear Econometric Models.” *Oxford Bulletin of Economics and Statistics*, 67(1): 905–30.

————. 2007. “How Sticky Is Sticky Enough? A Distributional and Impulse Response Analysis of New Keynesian DSGE Models.” *Journal of Money, Credit, and Banking* 39(6): 1481–508.

Kydland, F.E. and E.C. Prescott. 1982. “Time to Build and Aggregate Fluctuations.” *Econometrica* 50(6): 1345–70.

Laforte, J.P. 2007. “Pricing Models: A Bayesian DSGE approach for the US Economy.” *Journal of Money, Credit, and Banking* 39(s1): 127-154.

Levin, A.T., A. Onatski, J.C. Williams and N. Williams. 2006. “Monetary Policy Under Uncertainty in Micro-Founded Macroeconometric Models.” In *NBER Macroeconomics Annual 2005*, edited by M. Gertler and K. Rogoff. MIT Press.

Lorenzoni, G. 2008. “A Theory of Demand Shocks.” Massachusetts Institute of Technology.

Lucas, R.E. Jr. 1980. “Methods and Problems in Business Cycle Theory.” *Journal of Money, Credit, and Banking* 12(4): 696–715.

Mackowiak, B. and M. Wiederholt. 2007. “Optimal Sticky Prices under Rational Inattention.” Northwestern University.

Mankiw, N.G. and R. Reis. 2002. “Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve.” *Quarterly Journal of Economics* 117(4): 1295–328.

————. 2003. “Sticky Information: A Model of Monetary Non-neutrality and Structural Slumps.” In *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, edited by P. Aghion, R. Frydman, J. E. Stiglitz, and M. Woodford. Princeton University Press.

————. 2006. “Pervasive Stickiness.” *American Economic Review* 96(2): 164–69.

———. 2007. "Sticky Information in General Equilibrium." *Journal of the European Economic Association* 2(2–3): 603–13.

Mankiw, N.G., R. Reis and J. Wolfers. 2004. "Disagreement about Inflation Expectations." In *NBER Macroeconomics Annual 2003*, edited by M. Gertler and K. Rogoff, 209–48. MIT Press.

Meyer-Gohde, A. 2007. "Solving Linear Rational Expectations Models with Lagged Expectations Quickly and Easily." SFB-649 discussion paper 2007-069. Berlin: Humboldt University.

Molinari, B. 2007. "The Role of Sticky Information in Inflation Dynamics: Estimates and Findings." Temi di Ricerca 50. Rome: Ente Einaudi.

Morris, S. and H.S. Shin. 2006. "Inertia of Forward-Looking Expectations." *American Economic Review* 96(2): 152–57.

Nunes, R.C. 2006. "On the Epidemiological Microfoundations of Sticky Information." Universitat Pompeu Fabra.

Paustian, M. and E. Pytlarczyk. 2006. "Sticky Contracts or Sticky Information? Evidence from an Estimated Euro Area DSGE Model." Bowling Green State University.

Pickering, A. 2004. "Sticky Information and the Phillips Curve: A Tale of Two Forecasts." University of Bristol.

Reis, R. 2006a. "Inattentive Consumers." *Journal of Monetary Economics* 53(8): 1761–800.

———. 2006b. "Inattentive Producers." *Review of Economic Studies* 73(3): 793–821.

Rich, R.J. and J.S. Tracy. 2006. "The Relationship between Expected Inflation, Disagreement, and Uncertainty: Evidence from Matched Point and Density Forecasts." Staff report 253. Federal Reserve Bank of New York.

Rotemberg, J. and M. Woodford. 1997. "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy." *NBER Macroeconomics Annual 1997*, edited by B. S. Bernanke and J. J. Rotemberg, 297–346. MIT Press.

Rudebusch, G.D. 2002. "Term Structure Evidence on Interest Rate Smoothing and Monetary Policy Inertia." *Journal of Monetary Economics* 49(6): 1161–87.

Sims, C.A. 1998. "Stickiness." *Carnegie-Rochester Conference Series on Public Policy* 49(1): 317–56.

Smets, F. and R. Wouters. 2003. "An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area." *Journal of the European Economic Association* 1(5): 1123–75.

———. 2007. "Shocks and Frictions in U.S. Business Cycles: A Bayesian DSGE Approach." *American Economic Review* 97(3): 586–606.

Sveen, T. and L. Weinke. 2005. "New Perspectives on Capital, Sticky Prices, and the Taylor Principle." *Journal of Economic Theory* 123(1): 21–39.

Svensson, L.E.O. 2003. "What Is Wrong with Taylor Rules? Using Judgment in Monetary Policy through Targeting Rules." *Journal of Economic Literature* 41: 426–77.

Taylor, J.B. 1979. "Estimation and Control of a Macroeconomic Model with Rational Expectations." *Econometrica* 47(5): 1267–86.

———. 1993. "Discretion versus Policy Rules in Practice." *Carnegie-Rochester Conference Series in Public Policy* 39: 195–214.

Trabandt, M. 2007. "Sticky Information vs. Sticky Prices: A Horse Race in a DSGE Framework." Working paper 1369. Kiel Institute for the World Economy.

Woodford, M. 1998. "Doing without Money: Controlling Inflation in a Post-Monetary World." *Review of Economic Dynamics* 1(1): 173–219.

———. 2003a. "Imperfect Common Knowledge and the Effects of Monetary Policy." In *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, edited by P. Aghion, R. Frydman, J. E. Stiglitz, and M. Woodford. Princeton University Press.

———. 2003b. *Interest and Prices*. Princeton University Press.

———. 2003c. "Optimal Interest-Rate Smoothing." *Review of Economic Studies* 70(4): 861–86.

———. 2005. "Firm-Specific Capital and the New Keynesian Phillips Curve." *International Journal of Central Banking* 1(2): 1–46.

Monetary Policy and Key Unobservables: Evidence from Large Industrial and Selected Inflation-Targeting Countries

Klaus Schmidt-Hebbel
Pontificia Universidad Católica de Chile

Carl E. Walsh
University of California at Santa Cruz

In informal terms, we are uncertain about where the economy has been, where it is now, and where it is going.

—Donald Kohn

In recent years, the design of monetary policy has focused on gaps—the output gap, the interest rate gap, and the unemployment rate gap have all played a role in policy discussions. Standard models used for policy analysis are either specified in terms of such gaps or imply important roles for these gap variables in the implementation of monetary policy. In each case, the gap is defined as the difference (often in percentage terms) between an observable variable, such as output or unemployment, and an unobserved variable, such as potential output or the natural rate of unemployment.

At the time of the conference, Klaus Schmidt-Hebbel was affiliated with the Central Bank of Chile.

We thank Nicoletta Batini, John Cochrane, Marco Del Negro, Hans Genberg, John McDermott, Ricardo Reis, John Williams, Noah Williams, and participants at the Annual Conference of the Central Bank of Chile and seminars at the European Central Bank and the Bank of England for very useful comments. We thank Thomas Laubach and John Williams for kindly sharing their codes with us. We also thank Fabián Gredig, Gustavo Leyva, and Juan Díaz for superb research assistance.

Monetary Policy under Uncertainty and Learning, edited by Klaus Schmidt-Hebbel and Carl E. Walsh, Santiago, Chile.

The presence of unobservable variables in the definitions of these gaps poses significant problems for central banks as they implement monetary policy. These problems are both conceptual in nature (what is the right definition of the output gap, potential output or the neutral real interest rate?) and practical (which of many empirical strategies for estimating unobservables should be used?). These problems are compounded by the fact that real-time data used to estimate unobservables will be revised in the future, implying that the best estimates available at the time policy decisions must be taken may, in hindsight, diverge significantly from estimates based on subsequent vintages of data.

To estimate these key unobservables, economists have drawn on a variety of methodologies. Univariate approaches based on statistical methods designed to decompose a time series into trend and cycle have been widely used to estimate variables such as potential output or the natural rate of unemployment. Multivariate approaches, in turn, employ the joint behavior of several variables whose trend or cyclical elements may be related. Multivariate strategies offer the possibility of bringing economic structure to bear on the estimation problem by incorporating the restrictions implied by an economic model. For example, Okun's Law suggests a relationship between the output gap and the gap between unemployment and the natural rate of unemployment. Thus, the joint behavior of output and unemployment may provide information that is useful for estimating both these gaps. However, the results obtained by previous researchers studying different time periods or different economies are difficult to compare across countries since estimation methodologies often differ significantly. This hinders the ability to assess how business cycles might be linked across countries, how potential output or the neutral real interest rate in different countries might be related, and how closely related the various gaps might be across a sample of countries.

While the literature on international business cycles employs common methods to estimate output gaps (Backus, Kehoe, and Kydland, 1992), this work typically uses univariate statistical techniques (such as the Hodrick-Prescott filter) to extract the cyclical component of output. A univariate approach ignores the information that is potentially available if one considers the joint behavior of several macroeconomic variables that are affected by the same set of unobservable variables. Variable definitions, sample periods, and the set of unobservables examined also vary across applications to individual countries. And while individual central banks have undertaken efforts to estimate these unobservable variables, their approaches have generally been country specific and have not provided either systematic estimation or comparison across countries.

Garnier and Wilhelmsen (2009) and Benati and Vitale (2007) adopt a joint estimation approach to uncover important unobservables for several countries. Garnier and Wilhelmsen focus on the United States, the euro area, and Germany, while Benati and Vitale study the United States, the United Kingdom, the euro area, Sweden, and Australia. However, this approach has not been extended to include a larger number of inflation-targeting economies or any emerging or developing economies.

Our objective is to provide a consistent approach to estimating potential output, the neutral interest rate, and the natural rate of unemployment, using data from ten economies: the three largest industrial economies (the United States, the euro area, and Japan) and seven inflation-targeting countries (Australia, Canada, Chile, New Zealand, Norway, Sweden, and United Kingdom). Country-by-country estimation of the three unobservables is based on a parsimonious monetary policy model, extending Laubach and Williams' (2003) sequential-step estimation procedure. This allows us to exploit our ten countries' time-series estimates of unobservables to test for commonalities and differences in their macroeconomic developments.

Section 1 provides a brief discussion of the role of unobservables in the design and implementation of monetary policy. This discussion serves, in part, to motivate the variables on which our empirical analysis focuses—namely, potential output, the neutral real interest rate, and the natural rate of unemployment. Section 2 then briefly sets out our empirical strategy. In section 3, we discuss the monetary policy model, the estimation approach, and the data, and report the country-by-country empirical results for parameter estimates and unobservables' time series. Section 4 extends the model and reports the corresponding results and robustness test results for the United States and Chile. Section 5 then uses our estimated series on the key unobservables to provide evidence of common trends, rising macroeconomic stability (the Great Moderation), comovements across our sample economies, and convergence of observables and unobservables in sample countries toward the United States and the euro area. Section 6 concludes and discusses extensions.

1. The Role and Importance of Unobservables in Monetary Policy

In this section, we discuss the role that key unobservables play in policy design. We then briefly review how errors in estimating potential gross domestic product (GDP) and the natural rate of unemployment have contributed to critical policy mistakes.

1.1 Unobservable Variables and Policy Design

The theoretical foundations both for monetary policy analysis and for the empirical models employed by central banks contain several important variables that are not directly observable. The output gap (the log difference between real GDP and an unobserved time-varying benchmark such as potential GDP) and the unemployment rate gap (the difference between the actual unemployment rate and the unobserved natural rate of unemployment) are typically the driving forces explaining inflation. Central banks may also need to monitor these unobservables out of a direct concern for macroeconomic stability. Both potential GDP and the natural rate of unemployment must be inferred from observable macroeconomic variables. Policymakers must also monitor difficult-to-measure expectations of inflation because they need to ensure that private sector expectations are consistent with the central bank's inflation targets (that is, they need to ensure that expectations are anchored) and because movements in inflation expectations can contribute to fluctuations in actual inflation. They also need to adjust policy interest rates to reflect changes in the economy's neutral real interest rate.

The critical role of these unobservable variables in designing monetary policy can be illustrated using a simple New Keynesian model. This benchmark model consists of a forward-looking Phillips Curve, an expectational IS relationship, and a specification of policy in terms of either an objective function (which the central bank is then assumed to maximize) or a decision rule (see Clarida, Galí, and Gertler, 1999).

If the central bank's objective is to minimize the volatility of inflation and the gap between output and potential output, then optimal policy (under discretion) can be described in terms of what Svensson and Woodford (2005) call a targeting rule. Such a rule involves ensuring that a weighted sum of the output gap and the inflation gap (that is, inflation minus the inflation target) is always kept equal to zero. Intuitively, the output gap should be negative when inflation is above target, as this will tend to produce a fall in inflation and thus bring inflation back to its target level. Similarly, the output gap should be positive when inflation is below target. The Bank of Norway describes such a targeting relationship between the output gap and inflation in its inflation report, in discussing the desirable properties of future interest rate paths. The discussions of interest rate projections in the Reserve Bank of New Zealand's monetary policy statements are consistent with a similar, though implicit targeting rule. In following such a rule, the central bank knows its inflation target, and it has direct measures of both inflation and output (while the latter may be subject to serious real-time measurement errors, it is directly observable in principle), but it must estimate the level of potential output.

Potential output is not the only unobserved variable the central bank must estimate as it implements policy. To actually implement an optimal targeting rule, the central bank must still determine how to move its policy interest rate to maintain the required relationship between the output and inflation gaps. Determining the nominal interest rate that will implement the optimal policy requires knowledge of the relationship between interest rates and real spending, a relationship commonly summarized in New Keynesian models by an expectational IS curve. Using a standard specification of the IS relationship, one finds that the optimal interest rate will satisfy the following relationship (see Clarida, Galí, and Gertler, 1999):

$$i_t = r_t^* + \left[1 + \frac{\sigma\kappa(1-\rho)}{\rho\lambda}\right] E_t \pi_{t+1}, \tag{1}$$

where i is the nominal interest rate, π is the inflation rate, r^* is the neutral real interest rate, the rate consistent with a zero output gap, and E is the conditional expectations operator.[1] The parameters σ, κ, λ, and ρ are, respectively, the inverse of the interest elasticity of aggregate demand, the output gap elasticity of inflation, the relative weight the policymaker places on output gap volatility relative to inflation volatility, and the degree of serial correlation in shocks to

1. There are numerous ways to write this relationship and to define the various unobservables. For example, it would be more in keeping with standard New Keynesian models to define r^* as the real interest rate consistent with output and the flexible-price equilibrium level of output being equal.

the inflation equation. Both the variables on the right-hand side of equation (1) are unobservable or measurable only indirectly—for example, via surveys, asset prices, or the term structure of interest rates.[2]

To solve for the equilibrium under the interest rate rule given by equation (1), the IS and Phillips curve relationships must also be specified. The ones underlying the derivation of equation (1) take the form

$$x_t = E_t x_{t+1} - \left(\frac{1}{\sigma}\right)(i_t - E_t \pi_{t+1} - r^*{}_t) \tag{2}$$

and

$$\pi_t = \beta E_t \pi_{t+1} + \kappa x_t + e_t, \tag{3}$$

where x is the output gap and e is a zero-mean stochastic error term. The parameter β is the inflation-expectations elasticity of inflation.

It is clear from equation (1) that the neutral real interest rate will be of critical importance for getting the level of the policy rate right. Under an interest rate operating procedure for monetary policy, the level of the nominal rate when the inflation rate is equal to its target must be consistent with the economy's equilibrium real rate of return. When inflation is equal to its (constant) target level, the Fisher relationship requires that the nominal interest rate equal the neutral rate plus the target inflation rate. Thus, while most of the recent literature emphasizes the importance of the Taylor Principle—that is, the need to adjust the nominal rate more than one for one with changes in inflation—it is equally important to fully adjust the nominal rate in response to changes in the neutral real interest rate. Woodford (2003) has labeled the equilibrium real interest rate associated with the absence of fluctuations resulting from nominal distortions as the Wicksellian real rate. An optimal monetary policy that maintains zero inflation to "undo" the real distortions created by nominal rigidities would ensure that the gap between the nominal interest rate and the Wicksellian rate remains equal to zero.

2. If the inflation-adjustment relationship incorporates lagged inflation, the targeting rule would also include further terms involving forecasts of future inflation rates and output gaps.

Unfortunately, this Wicksellian or neutral real rate is unobservable. It is, however, closely related to another key unobservable—the output gap. In the context of the simple model used to derive equation (1), the neutral real interest rate is proportional to the growth rate of potential real output. Laubach and Williams (2003) use this relationship between these two unobservable variables to help them estimate the neutral real interest rate for the United States.

Equations (2) and (3) also serve to highlight the key role of unobservable variables. The output gap appears in both, as does expected future inflation, while the neutral real interest rate appears in the IS relationship. Before a central bank can actually use this simple framework for policy analysis, methods need to be developed for estimating potential output (to obtain an output gap measure), expected inflation, and the neutral real interest rate.

The difficulties in measuring the output gap go, in some sense, beyond the need to measure potential output, because the very definition of the output gap has evolved over the past twenty years. At the conceptual level, three distinct definitions have been employed. The first definition of the output gap is in terms of the relationship between actual GDP and potential GDP, where potential GDP is typically associated with the level of GDP that would be produced at full employment of labor and capital at normal utilization rates. This is the definition most commonly used in models employed by central banks.

In recent years, the development of the New Keynesian Phillips curve has focused attention on a second definition of the output gap, which the underlying theory identifies as the key variable driving inflation. This is the output gap measured as the gap between actual GDP and the level of GDP that would be produced in the absence of nominal wage and price rigidities. This flexible-price output gap provides a measure of economic fluctuations that are due to nominal rigidities. These nominal rigidities allow monetary policy to have real effects, but they also create real distortions. Standard New Keynesian models imply that monetary policy should aim at eliminating these distortions by minimizing fluctuations in the output gap.

However, stabilizing the flexible-price output gap is difficult, not least because the economy's equilibrium output that would arise if there were no nominal rigidities is clearly not observable, and it cannot be estimated using the (often) univariate statistical approaches employed to estimate potential output. Instead, any estimate must come from employing a dynamic stochastic general equilibrium (DSGE) model that can simulate the behavior of an economy that is not subject to nominal rigidities. Since the correct model of the economy is unknown, any estimate of the output gap will be subject to a great deal of uncertainty. Levin and others (2006) provide one example of a DSGE model that is estimated based on U.S. data, which they use to construct a measure of the flexible-price output level and the associated flexible-price output gap. To date, no central banks have employed such a definition of the output gap in their formal policy models.[3] Nevertheless, many central banks are working on developing DSGE models and applying them to estimate flexible-price output levels, as well as other unobservables.

Finally, a third definition of the output gap is the gap between output and the welfare-maximizing level of output. The gap defined in this manner is sometimes called the welfare gap. While this gap may be the most relevant for policy from a conceptual viewpoint, it is also the hardest to measure. The welfare gap and the flexible-price output gap move together in standard New Keynesian models, so stabilizing one is equivalent to stabilizing the other, a property that Blanchard and Galí (2007) label “the divine coincidence.” In general, however, the relationship between the two gap measures holds only under very special conditions. If real wages are sticky or if there are other labor market frictions or fluctuations in distortionary taxes, the flexible-price output gap and the welfare gap will diverge.

In addition to illustrating the general point that hard-to-measure variables are conceptually relevant for policy, equations (1) through (3) highlight the variables that are the primary focus of our study. These are the neutral real interest rate, potential output, and expected inflation. For our purposes, we define the output gap as the log of real GDP minus the log of potential GDP, which is the common definition among central banks. The natural rate of unemployment, which is linked to potential output, does not appear explicitly in equation (1), but we incorporate it into our analysis.

3. A possible exception is models that have developed from the Bank of Canada’s Quarterly Projections Model (QPM), such as the Forecasting and Policy System model of the Reserve Bank of New Zealand. This model distinguishes between a long-run component, a short-run equilibrium component, and a cyclical component to output. The output gap is then defined relative to the short-run equilibrium level and thus might correspond to a flexible price output gap. However, the short-run equilibrium level of output is an estimate of a slow-moving trend, based on a multivariate filter. Variables (in addition to output) included in the trend estimation procedure include capacity utilization, unemployment, and inflation. QPM was replaced recently at the Bank of Canada by a new open economy DSGE model, called the Terms-of-Trade Economic Model (ToTEM); see Murchison and Rennison (2006).

1.2 Unobservable Variables and Policy Mistakes

Unobservable variables play a critical role in the design and implementation of optimal monetary policy, but these same variables have also been center stage in a number of accounts of past policy errors.[4] For example, Orphanides (2002, 2003), Erceg and Levin (2003), Reis (2003), and Primiceri (2006) all argue that errors by either policymakers or the public in estimating key macroeconomic variables were central to an understanding of critical episodes in the inflation history of the United States over the past forty years.

Orphanides focuses on the Federal Reserve's real-time overestimation of potential (trend) output following the productivity slowdown of the early 1970s. Simply put, overestimation of potential GDP implied an underestimation of the output gap. This led to a policy stance that was, in retrospect, too expansionary and contributed to producing the Great Inflation of the 1970s. Orphanides and Van Norden (2002) document the difficulties of estimating the output gap when, for policy purposes, this must be done using real-time data.[5] McCallum (2001) draws the conclusion that policymakers should not respond strongly to movements in the estimated output gap.[6]

Primiceri (2006) argues that the Fed's failure to correctly estimate potential output is only part of the story behind the Great Inflation.[7] He argues that if that were the only mistake, inflation would not have risen so much or for so long. The second factor contributing to the persistence of high inflation was the Fed's underestimation of the persistence of inflation. Initial increases in inflation were not expected to persist, so policy did not react strongly. Because potential output was overestimated, economic slowdowns that were

4. See Sargent (2008) for an overview and discussion.

5. The Reserve Bank of New Zealand provides a figure comparing their real-time quarterly output gap estimates and estimates prepared using final data (as of November 2002) for the period 1997–2002 (Reserve Bank of New Zealand, 2004, figure 9, page 15). There are sizable differences between the two: for instance, the final series changes sign four times during the period shown, while the real time series changes sign three times *and never in the same quarter* as the final estimate series.

6. Orphanides and Williams (2002) find that policy rules that respond to the change in the unemployment rate gap or the output gap perform well. One reason might be that differencing eliminates much of the error in measuring the level of the output gap.

7. Primiceri's model is actually expressed in terms of the natural rate of unemployment rather than potential output.

thought to be associated with negative output gaps did not seem to lower inflation. Policymakers thus concluded that inflation was unresponsive to economic activity and that a major recession would be needed to lower inflation. Perceiving that they faced a large sacrifice ratio if they tried to lower inflation, policymakers hesitated to try to bring inflation down. Primiceri develops a simple general equilibrium model in which the policymaker learns about the natural rate and the degree of inflation persistence, and his model accounts for both the policy mistakes of the 1970s, as the Fed underestimated the natural rate of unemployment and overestimated the sacrifice ratio associated with lowering inflation, and the disinflationary shift in policy under Volcker. Primiceri's analysis shows that both the difficulties in estimating unobservable variables and the fact that central banks do not know the true structure of the economy can contribute to policy errors.

The public also faces the need to estimate unobservable variables. Erceg and Levin (2003) focus on shifts in the Fed's implicit inflation target when these shifts are not publicly announced. In this case, the public becomes aware of the shift in target only gradually. Erceg and Levin characterize the Volcker disinflation as the result of a fall in the Fed's target inflation rate. Since this target change was not made explicit through any public announcement, agents overestimated inflation, which led to a significant contraction in real economic activity. While our focus is on estimating unobservable variables for use in designing monetary policy, the work of Erceg and Levin provides a reminder of the consequences that can occur when the central bank's inflation target is, from the perspective of the public, an unobservable.

2. Alternative Approaches to Estimating the Neutral Real Rate, the Output Gap, and the Natural Rate of Unemployment

There is a vast literature that uses a range of empirical techniques to estimate unobservable macroeconomic variables. Our survey is therefore brief and highly selective, focusing on contributions that are the most directly relevant for our own empirical approach. For example, while a large amount of work employs univariate methods to estimate potential output or the natural rate of unemployment, we do not focus on these approaches. We follow multivariate approaches that incorporate information from other macroeconomic variables, usually employing theory to guide the relationship between the variables or employing structural equations motivated by theory. We focus on multivariate approaches that are directly relevant for the methods we use to obtain estimates of key unobservable variables. These approaches generally combine statistical representations borrowed from the literature on identifying trend and cyclical components of a time series with relationships among variables implied by an economic model.

The general methodology we employ involves a multivariate Kalman filter to extract estimates of unobserved components from observed time series. The basic framework can be represented in quite general terms of a specification for the dynamic evolution of a vector $\mathbf{Z}_t$ of unobserved factors and a vector of observed variables $\mathbf{Y}_t$ that are related to $\mathbf{Z}_t$. The evolution of the unobserved variables is given in state-space form by

$$\mathbf{Z}_{t+1} = \mathbf{A}\mathbf{Z}_t + \mathbf{u}_{t+1}. \tag{4}$$

The measurement equations linking $\mathbf{Y}_t$ to $\mathbf{Z}_t$ take the form

$$\mathbf{Y}_t = \mathbf{B}\mathbf{Y}_{t-1} + \mathbf{C}\mathbf{Z}_t + \mathbf{D}\mathbf{Z}_{t/t} + \mathbf{G}\mathbf{X}_t + \mathbf{v}_t, \tag{5}$$

where $\mathbf{Z}_{t/t}$ is the time t estimate of the state vector $\mathbf{Z}_t$ and $\mathbf{X}_t$ is a vector of exogenous and observable variables. Both $\mathbf{u}_{t+1}$ and $\mathbf{v}_t$ are zero-mean stochastic error terms. In section 3, we specify the formulations of equations (4) and (5) that we use in our empirical analysis.

Time t estimates of $\mathbf{Z}_t$ are updated using the Kalman filter. Since

$$\mathbf{Y}_t - \mathbf{B}\mathbf{Y}_{t-1} - (\mathbf{C} + \mathbf{D})\mathbf{Z}_{t/t-1} - \mathbf{G}\mathbf{X}_t$$

is the new information available from observing $\mathbf{Y}_t$ in period t, the equation for updating estimates of $\mathbf{Z}$ is given by

$$\mathbf{Z}_{t/t} = \mathbf{Z}_{t/t-1} + \mathbf{K}\,[\mathbf{Y}_t - \mathbf{B}\mathbf{Y}_{t-1} - (\mathbf{C} + \mathbf{D})\mathbf{Z}_{t/t-1} - \mathbf{G}\mathbf{X}_t]. \tag{6}$$

The basic structure given by equations (4) through (6) has been used extensively to estimate a range of unobservable variables. Data on the observables $\mathbf{Y}_t$ and $\mathbf{X}_t$ are used to estimate the parameter matrices $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$, and $\mathbf{G}$.

An early application of the Kalman filter approach to estimating potential GDP for the United States is provided by Kuttner (1994).[8] Kuttner lets $\mathbf{Z}_t$ consist of trend and cyclical components of output, with the trend following a random walk with drift and the cyclical component described by a second-order autoregressive, or AR(2), process. The vector $\mathbf{Y}_t$ consists of real output and inflation and reflects a Phillips curve relationship. Output is the sum of its trend and cyclical components, and inflation is a function of lagged output growth and the cyclical component of output.

Basistha and Nelson (2007) take a related approach to estimating potential GDP and output in the United States. Like Kuttner, they adopt a latent variable approach and incorporate a Phillips curve relationship. They also include the unemployment rate and allow trend and cyclical components of output to be correlated.

Laubach and Williams (2003) extend the Kuttner framework to incorporate the neutral real interest rate, r^*, as an additional unobserved variable. They assume that r^* is a function of the growth rate of potential GDP and a stochastic component that follows an autoregressive process. They expand the set of measurement equations to include an IS relationship linking the output gap to the gap between the real and neutral interest rates.[9] While this specification allows for an integrated approach to estimating potential GDP and the neutral real interest rate, Laubach and Williams employ a separate univariate inflation-forecasting equation to obtain the estimate of expected inflation they need to construct the real interest rate.

Fuentes, Gredig, and Larraín (2008) further extend the approach of Laubach and Williams by incorporating the unemployment rate and Okun's Law linking the output gap and the gap between the unemployment rate and the natural rate of unemployment. The latter is assumed to follow a random walk. They compare the resulting measures of the output gap for Chile with gap estimates obtained from structural vector autoregressions (VARs) and production function approaches. Interestingly, the estimates based on the Kalman filter provided the best out-of-sample forecasts for inflation.

8. Orphanides and Williams (2002) provide an overview of the literature that estimates the natural rates of unemployment and the neutral real interest rates for the United States.

9. They also allow the growth rate of potential GDP to follow a random walk.

Each of these examples from the literature focuses on a single country; the United States in the cases of Kuttner (1994), Basistha and Nelson (2007), and Laubach and Williams (2003) and Chile in the case of Fuentes, Gredig, and Larraín (2008). The closest formulation to our approach is by Benati and Vitale (2007). They, too, focus on multiple unobservables (namely, potential output, the natural unemployment rate, the neutral real interest rate, and expected inflation), and they obtain estimates of each unobservable for five economies (Australia, the euro area, Sweden, the United Kingdom, and the United States). Benati and Vitale allow for time variation in the model parameters. We restrict our attention to constant coefficient models.

Björksten and Karagedikli (2003) report estimates of the neutral real interest rate for seven countries (namely, Australia, Canada, New Zealand, Sweden, Switzerland, the United Kingdom, and the United States), using a methodology based on long- and short-term interest rates. To extract real interest rates, however, they assume that expected inflation is equal to actual inflation. They find a marked decline since 1998 in neutral real rates for all seven countries.[10] Similarly, Fuentes and Gredig (2008) find evidence of a trend decline in Chile's neutral interest rate.

3. Empirical Results

Our approach, following the preceding literature, is based on a parsimonious New Keynesian specification. We use the core relationships in the New Keynesian model to guide our specification of the linkages between observable variables and the key unobservables as summarized in equation (5). The two relationships from the New Keynesian model that we draw on are the IS equation and the Phillips curve. We also use a Taylor rule to represent monetary policy and Okun's Law to link the unemployment gap and the output gap.

3.1 The Model

We start with a simple backward-looking IS relationship, as in Rudebusch and Svensson (1999), where the output gap (x) is determined by its own lag, the lagged real interest rate gap (the

10. See also Basdevant, Björksten, and Karagedikli (2004).

difference between the observed ex ante real interest rate, r, and the unobserved neutral real interest rate, r^*), and a serially uncorrelated error term (ε_1):

$$x_t = \alpha_1 x_{t-1} + \alpha_2 (r_{t-1} - r^*_{t-1}) + \varepsilon_{1,t}. \tag{7}$$

The output gap is defined as the difference between actual output (y) and unobserved potential output or the natural level of output (y^*), both in logs:

$$x_t = y_t - y^*_t. \tag{8}$$

The second relationship is a standard Phillips curve specification for inflation. We specify this equation in terms of the inflation gap rather than the level of inflation, where the inflation gap, $\overline{\pi}_t$, is the difference between actual inflation and either trend inflation (in the case of non-inflation-targeting countries) or between actual inflation and the target inflation rate (for inflation targeters). The inflation gap is determined by its own lag, the expected inflation gap, the lagged output gap, and a serially uncorrelated error term (ε_2):

$$\overline{\pi}_t = \beta_1 \overline{\pi}_{t-1} + \beta_2 \overline{\pi}^e_t + \beta_3 x_{t-1} + \varepsilon_{2,t}. \tag{9}$$

The inflation gap is an observable variable, given by

$$\overline{\pi}_t = \pi_t - \pi_t^{\ T}, \tag{10}$$

where π_t is actual inflation and $\pi_t^{\ T}$ is the trend or target rate. Similarly, the inflation expectations gap is defined as the difference between observed (estimated) inflation expectations and trend or target inflation:

$$\overline{\pi}^e_t = \pi^e_t - \pi_t^{\ T}. \tag{11}$$

We specify a standard Taylor rule that relates the observed ex ante real interest rate to the ex ante real natural rate, the real interest rate lag, the inflation expectations gap, the lagged output gap, and a serially uncorrelated error term (ε_3):

$$r_t = r^*_t + \delta_1 (r_{t-1} - r^*_{t-1}) + \delta_2 \overline{\pi}^e_t + \delta_3 x_{t-1} + \varepsilon_{3,t}. \tag{12}$$

Equations (7) through (12) comprise our basic model. As an extension of this model, we add Okun's Law that relates the observed unemployment rate (u) to the unobserved natural rate of unemployment (u^*), the lagged gap between the observed unemployment rate and the natural rate of unemployment, the output gap, and a serially uncorrelated error term (ε_4):

$$u_t = u_t^* + \gamma_1(u_{t-1} - u_{t-1}^*) + \gamma_2 x_{t-1} + \varepsilon_{4,t}. \tag{13}$$

Now we turn to the transition equations of the model corresponding to equation (4) in the schematic formulation of section 2. As in Laubach and Williams (2003), potential output is taken to follow a second-order integrated, or I(2), process and unobserved potential output growth (g) follows a random walk:

$$y_t^* = y_{t-1}^* + g_{t-1} + \varepsilon_{5,t} \tag{14}$$

and

$$g_t = g_{t-1} + \varepsilon_{6,t}, \tag{15}$$

where ε_5 and ε_6 are serially uncorrelated error terms.

To close the model, we specify random-walk processes for both the neutral real interest rate and the natural rate of unemployment:

$$r_t^* = r_{t-1}^* + \varepsilon_{7,t} \tag{16}$$

and

$$u_t^* = u_{t-1}^* + \varepsilon_{8,t}, \tag{17}$$

where ε_7 and ε_8 are serially uncorrelated error terms.

3.2 Estimation Method

We closely follow Laubach and Williams' (2003) procedure in estimating our model, adapting it to our specification. As they note, maximum-likelihood estimates of the standard deviations of the innovations to the transition equations of the unobservables, as in equations (14) through (17), are likely to be biased toward zero because of the pile-up problem discussed by Stock (1994). We therefore also use the Stock and Watson (1998) median unbiased estimator to obtain estimates of the signal-to-noise ratios reflected by the ratios of the corresponding residual variances $\lambda_g = \sigma_6/\sigma_5$, $\lambda_r = (1 - \delta_1)\ \sigma_7/\sigma_3$, and $\lambda_u = (1 - \gamma_1)\ \sigma_8/\sigma_4$, where σ_i ($i = 1,\dots 8$) denote the corresponding variances of the error terms, ε_i. We impose the latter ratios when estimating the remaining model parameters by maximum likelihood.

We also follow Laubach and Williams (2003) closely in the subsequent sequential-step estimation procedure. In the first step (following Kuttner, 1994), we apply the Kalman filter to estimate jointly the IS relationship—after substituting equation (8) into (7)—and the Phillips curve—after substituting equations (10) and (11) into (9). In this stage we omit the real interest rate gap from the IS equation and assume that potential output growth (g) is constant. From the latter preliminary estimation, we obtain a preliminary potential output level series from which we compute an estimate of the (preliminary) constant potential output growth. We then estimate equation (14) to test for structural breaks in the level of g. Using Stock and Watson (1998, table 3), we determine a positive value for λ_g when the null of no structural break is rejected.

In the second step, we apply the Kalman filter to estimate jointly the IS relationship, the Phillips curve, the Taylor rule (equation 12), and the transition equations for potential output level (equation 14) and potential output growth (equation 15). At this stage, we impose a preliminary constant neutral interest rate (r^*) in the IS relation and the Taylor rule. We also impose the λ_g estimate obtained in the first step. From the latter preliminary estimation, we obtain an estimate of the (preliminary) constant neutral rate interest rate. We then estimate equation (12) to test for structural breaks in the level of r^*. Using Stock and Watson (1998, table 3), we determine a positive value for λ_r when the null of no structural break is rejected.

In step 3, we estimate jointly the IS relationship, the Phillips curve, the Taylor rule, and Okun's Law (equation 13), in addition to transition equations (14), (15), and (16). We impose a preliminary constant natural unemployment rate in Okun's Law. We also impose the λ_g and λ_r estimates obtained in the first and second steps. From the latter preliminary estimation, we obtain an estimate of the (preliminary) constant neutral unemployment rate. We then estimate equation (13) to test for structural breaks in the level of u^*. Using Stock and Watson (1998, table 3), we determine a positive value for λ_u when the null of no structural break is rejected.

Final step 4 comprises Kalman filter estimation of the full model, imposing the estimates for λ_g, λ_r, and λ_u obtained sequentially in the preceding steps. This yields the final estimates for our model coefficients and time series of unobservables. As in Laubach and Williams, we compute confidence intervals and standard errors for the parameters and unobservables applying Hamilton's (1986) Monte Carlo method.

3.3 Data

Our sample covers ten economies: the three largest industrial economies (namely, the United States, the euro area, and Japan), all of which have central banks that do not explicitly or exclusively target inflation; a group of six industrial countries with inflation-targeting central banks, comprised of New Zealand, Canada, United Kingdom, Australia, Sweden, and Norway; and Chile, an emerging economy with an inflation-targeting central bank.[11]

Time coverage of each country sample is determined by availability of quarterly data. Our standard sample covers the 1970–2006 period. One exception on the long side is the United States (1960–2007) and on the short side exceptions are New Zealand (1974–2006), Norway (1979–2006), and, in particular, Chile (1986–2006).[12] Data sources and definitions are reported in a data appendix.

3.4 Estimation Results

Here we report estimation results for our state-space model in its basic version (without Okun's Law) for all countries. This implies omitting step 3 of the estimation method described above and modifying step 4 accordingly. The model thus consists of equations (7) through (12) and (14) through (16). In section 4 below, we report empirical results based on the extended model that includes equations (13) and (17) for the United States and Chile and the corresponding full four-step estimation procedure.[13]

11. We attempted to include Israel (with 1986–2006 data), but we were not able to attain convergence of our estimation model.

12. We were restricted to using smaller samples owing to the lack of data on monetary policy rates or short-term deposit rates for New Zealand (before 1974) and Norway (before 1979) and the lack of quarterly data for most series for Chile before 1986.

13. We have experimented with two alternative specifications. The first includes one additional lag in both the IS and Phillips curves. In the second, we impose the restriction that the coefficients associated with inflation expectations and lagged inflation sum to unity. We did not obtain successful results applying either of these changes. In the first, we were not able to run the third step, while in the second, we encountered numerical problems.

Tables 1 through 5 report country estimates for the two key ratios of the standard deviations of the residuals (λ_g and λ_r), all structural model parameters, and standard deviations of the equation residuals. We report results for the full sample available for each country and a shorter sample extending from 1986 to 2006 for nine countries, except the United States, where it extends through 2007:2. Figures 1–10 depict the estimated time series of observables and unobservables for each country, consistent with the full-sample estimations.

Our estimation strategy is the following. When obtaining estimation results from the last step (that is, the modified fourth stage of the generalized model), we report them directly. If estimation results were not obtained at either the second or third stages, we conduct a grid search over an interval of values for the standard deviation ratios (λ_g and λ_r), as reported in the footnotes of the tables. We therefore report a varying number of results for each country. For example, for

Figure 1. Inflation, Output, and the Interest Rate in the United States, 1960:1–2007:2 and 1986:1–2007:2[a]

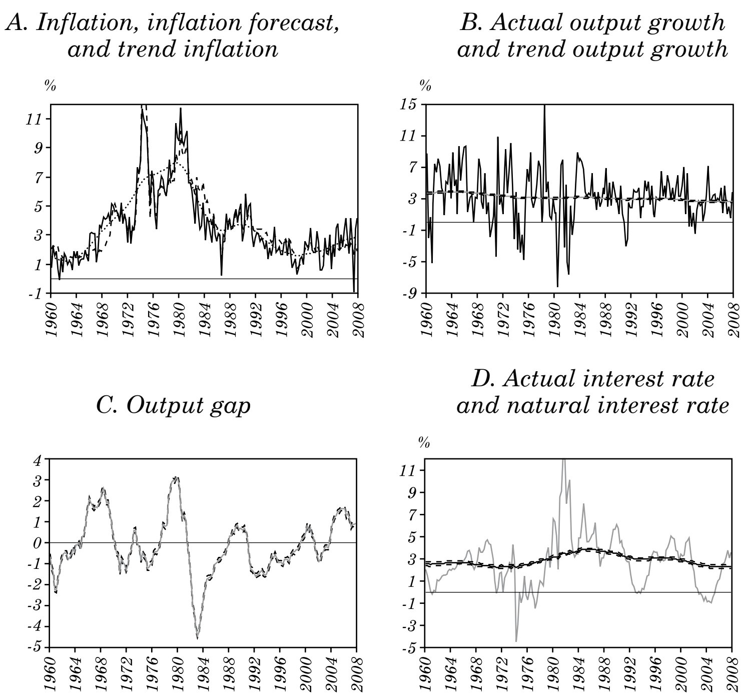
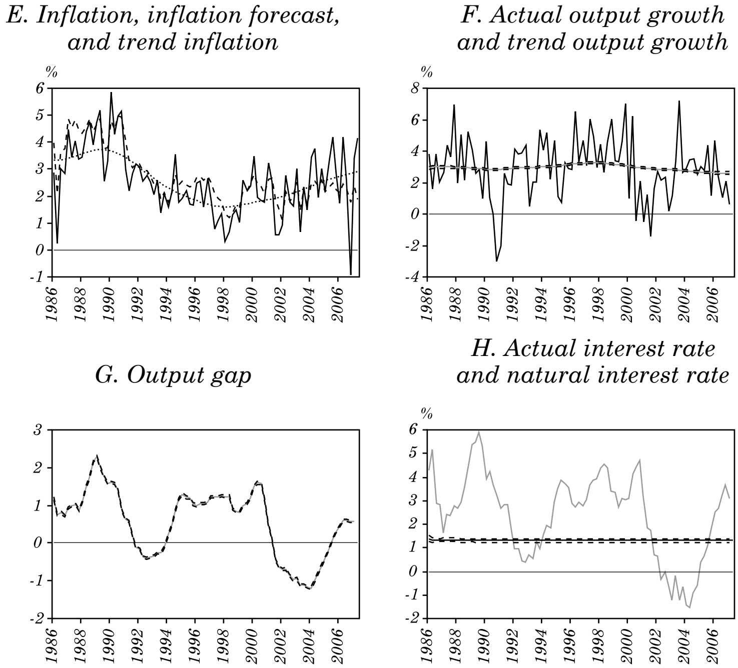


Figure 1. (continued)



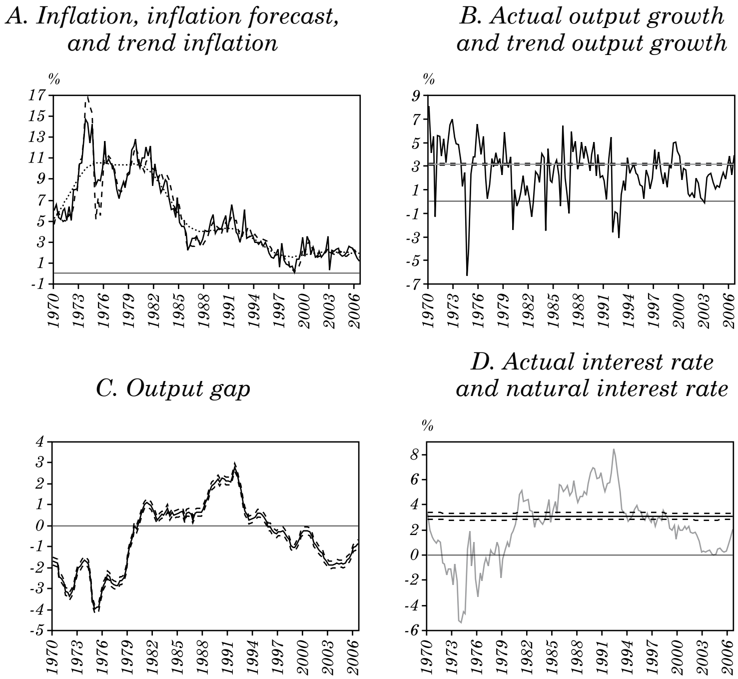
Source: Authors' calculations.

a. In panels A and E actual inflation is the solid line, inflation forecast the dashed line and inflation trend the dotted line. Panels A through D correspond to data from 1960:1−2007:2 and panels from E through H correspond to data from 1986:1−2007:2.

the United States (table 1), we report only one set of results for each sample period, as we obtained estimates for all model parameters. In contrast, we experienced estimation problems in the case of Japan (table 1), so we report a second set of results for each sample period, based on predetermined median values for λ_g and λ_r, corresponding to an interval of values over which we conducted a grid search.

While estimation results differ in significant ways across the ten countries, we point out the following general findings (abstracting from country-specific exceptions), reported in tables 1–5 and figures 1–10. First, the potential growth rate and the neutral real interest rate are typically not constant—not even for the shorter 1986–2006 sample—as reflected by nonzero values of λ_g and λ_r reported in the tables and depicted in the figures. This has implications for the

Figure 2. Inflation, Output, and the Interest Rate in the Euro Area, 1970:2–2006:4[a]



Source: Authors' calculations.
a. In panel A actual inflation is the solid line, inflation forecast the dashed line and inflation trend the dotted line.

construction of output gap measures as well as for the specification of Taylor rules.

Second, point values and significance levels of structural parameter estimates vary from country to country and sometimes from sample to sample for a given country. For example, most parameter estimates conform to our priors in the full-sample estimations for Canada, Chile, and the United States. At the other extreme is Japan, where parameter estimates were hard to obtain and, when estimated over a grid search, often did not conform to expected signs or significance levels.

Third, the IS equation generally reflects very large output gap inertia (reflected in the large and significant parameter estimate of its own lag). However, the sensitivity of the output gap to the lagged real interest rate gap ranges from negative and significant to positive and significant.

Fourth, the Phillips curve generally reflects small but significant inflation gap reversion, suggesting partial reversal of quarterly inflation shocks. (The exception is Chile, which reflects positive inflation gap persistence.) Expected inflation shocks affect inflation gaps positively, significantly, and by a large magnitude in many countries. The lagged output gap raises inflation significantly, positively, and by a sizable magnitude in most countries.

Fifth, the Taylor rule reflects significant inertia in central bank real interest rate innovations in all countries, with the exception of Japan. Most central banks raise nominal interest rates in response to a lagged inflation shock ($\delta_2 \geq -1$), but not enough to satisfy the Taylor principle. (Because we have specified the Taylor rule for the real interest rate, the Taylor principle requires that $\delta_2 \geq 0$.) The exception is Chile, where the coefficient estimate was found to be not

Figure 3. Inflation, Output, and the Interest Rate in Japan, 1970:2–2006:4[a]

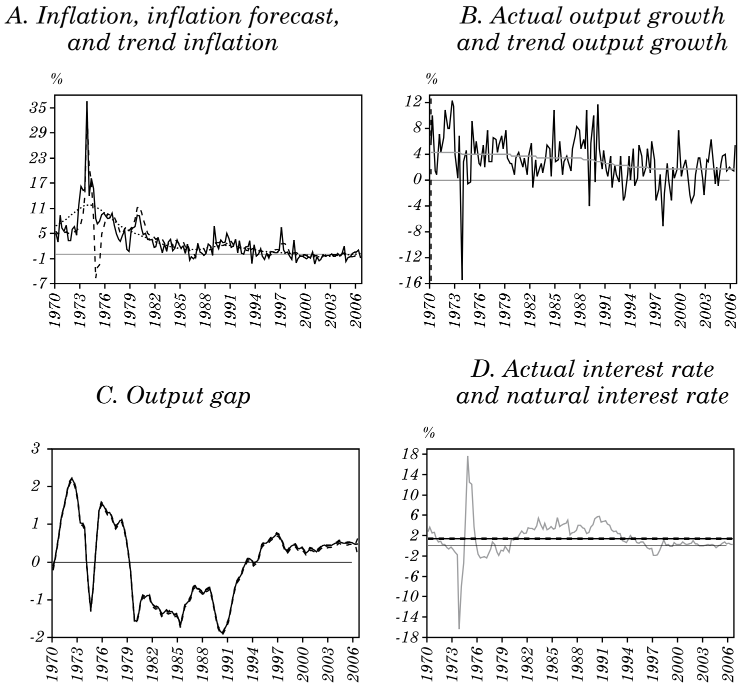
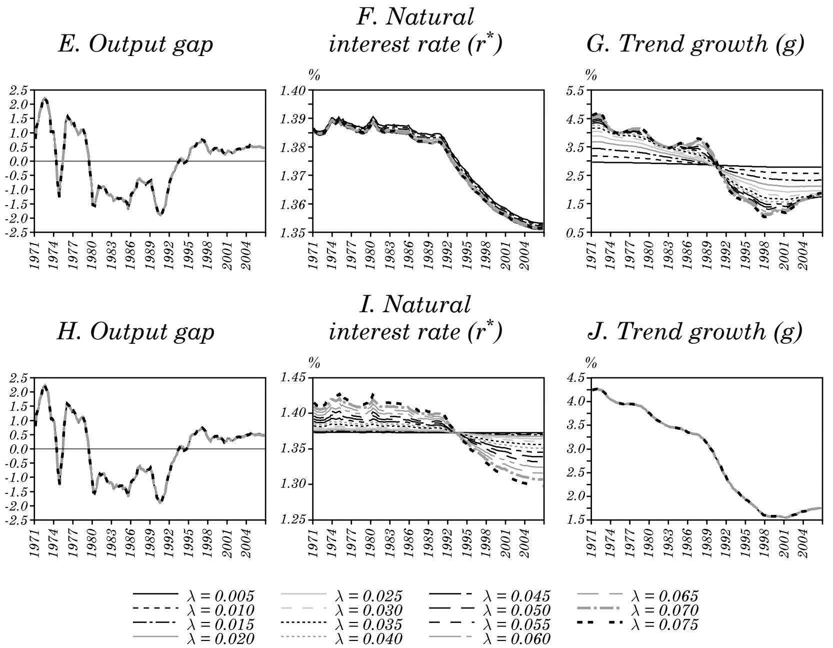


Figure 3. (continued)



Source: Authors' calculations.

a. In panel A actual inflation is the solid line, inflation forecast the dashed line and inflation trend the dotted line. Panels E, F, and G show the unobservables for different grid values for λ_g, while panels H, I, and J show the unobservables for different grid values for λ_r.

significantly different from zero.[14] We obtain a wide range for the interest rate gap response to a lagged output gap shock: monetary policy ranges from countercyclical (United States) to acyclical (Sweden) and to procyclical (Japan).

Finally, judging by conformity of parameter point estimates and significance levels to priors, the best country results were obtained for the United States (1960–2007) and Chile (1986–2006).

Our estimates for unobservables reveal the following results. First, the estimated time series for potential output growth displays

14. This may reflect that Chile's Central Bank responded to a rise in inflation expectations by maintaining its indexed policy rate when it was indexed to past inflation (1986–2000) and raising its nominal rate by the same magnitude of the shock in inflation expectations when the policy rate was set in nominal terms (2001–06).

smooth behavior, but g changes over time in most countries (except the euro area and Australia), consistent with positive country estimates for λ_g. Second, with relatively stable potential output growth, the variance of country output gaps is largely determined by the variance in actual output growth rates. Third, similar to potential output growth, the neutral real interest rate follows a smooth pattern in all countries, in line with positive country estimates for λ_r. Fourth, we generally obtained precise estimates for our three unobservables, as reflected by the narrow confidence intervals depicted in the figures. Fifth, we obtain similar estimates for potential output growth and the neutral real interest rates across the long and short samples for most countries. The exceptions are Australia and Norway, for which we obtain neutral interest rates well above actual levels in the shorter samples. Finally, we also obtain similar estimates for output gaps across the long and short samples in many countries. However,

Figure 4. Inflation, Output, and the Interest Rate in New Zealand, 1974:2–2006:4 and 1986:2–2006:4[a]

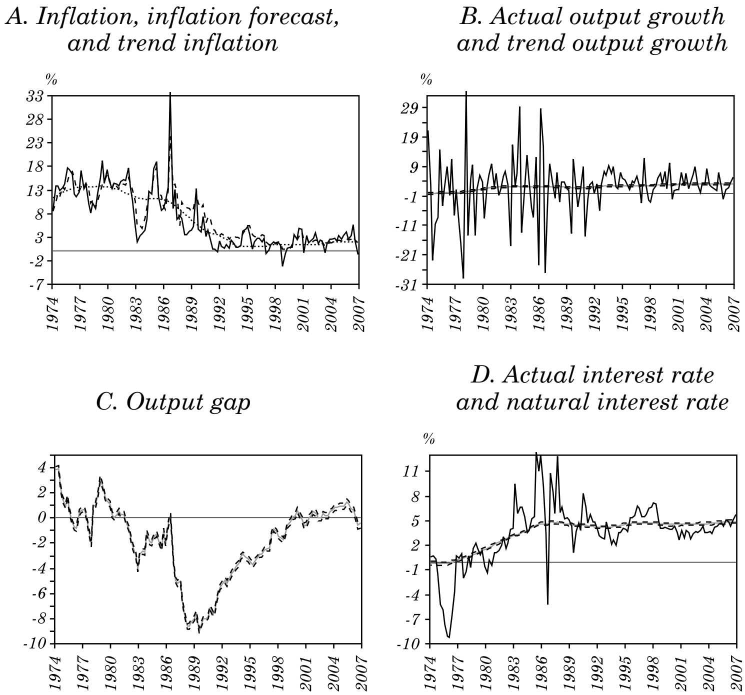
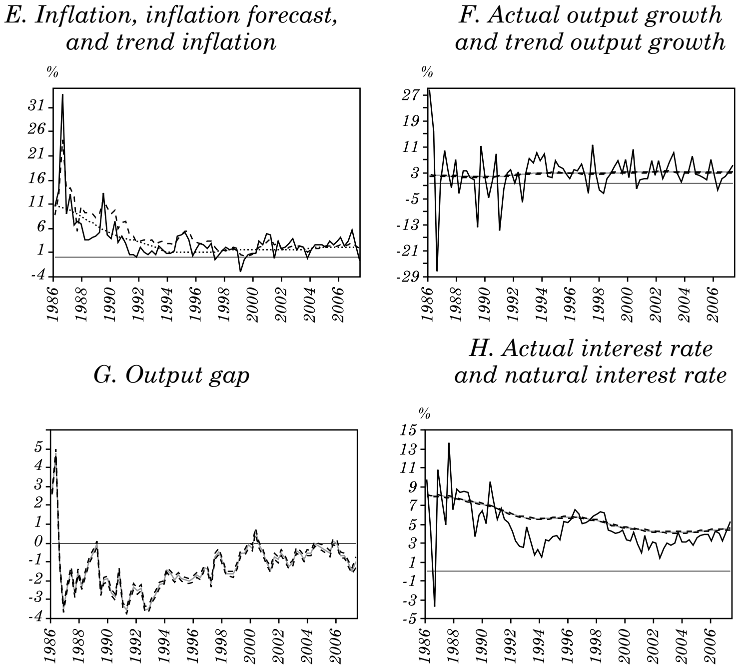


Figure 4. (continued)



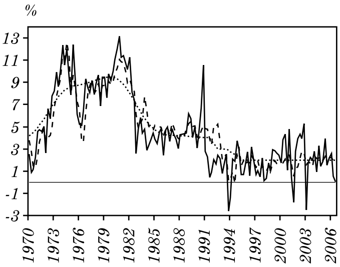
Source: Authors' calculations.

a. In panels A and E actual inflation is the solid line, inflation forecast the dashed line and inflation trend the dotted line. Panels A through D correspond to data from 1974:2−2006:4 and panels from E through H correspond to data from 1986:2−2006:4.

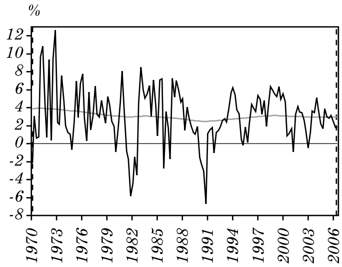
in Australia, New Zealand, Sweden, and the United Kingdom, the dynamic pattern, sign, and/or magnitude of output gap estimates differ significantly in the 1986–2006 sample from those obtained for the larger samples. This may reflect small-sample bias. We thus conduct our tests of the Great Moderation, comovements, and convergence across countries based on our large-sample estimates of unobservables.

Figure 5. Inflation, Output, and the Interest Rate in Canada, 1970:2–2006:4 and 1986:2–2006:4[a]

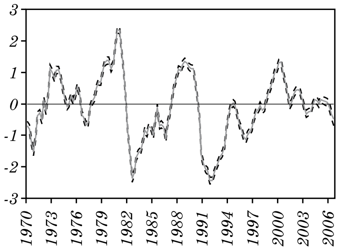
A. Inflation, inflation forecast, and trend inflation



B. Actual output growth and trend output growth



C. Output gap



D. Actual interest rate and natural interest rate

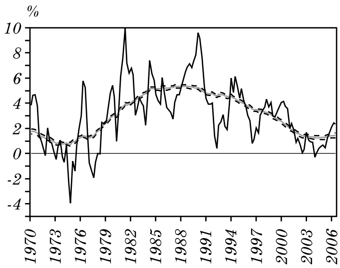
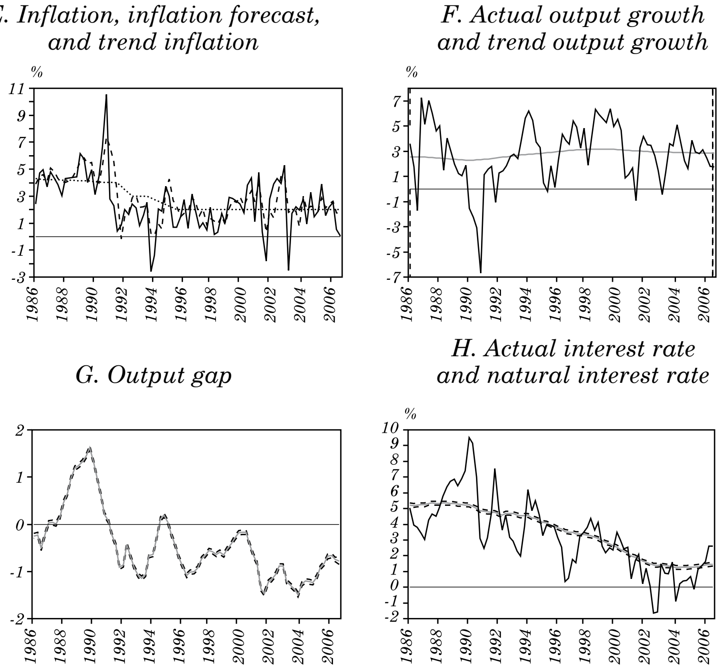


Figure 5. (continued)

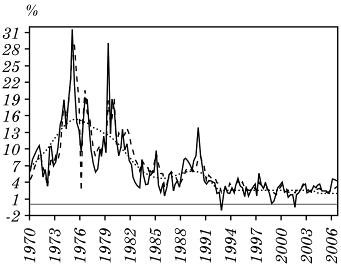


Source: Authors' calculations.

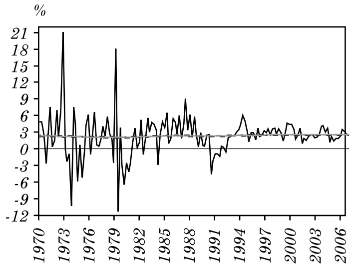
a. In panels A and E actual inflation is the solid line, inflation forecast the dashed line and inflation trend the dotted line. Panels A through D correspond to data from 1970:2−2006:4 and panels from E through H correspond to data from 1986:2−2006:4.

Figure 6. Inflation, Output, and the Interest Rate in the United Kingdom, 1970:2–2006:4 and 1986:2–2006:4[a]

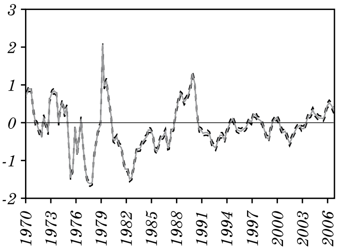
A. Inflation, inflation forecast, and trend inflation



B. Actual output growth and trend output growth



C. Output gap



D. Actual interest rate and natural interest rate

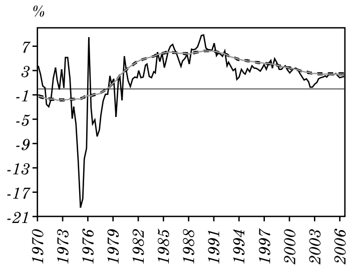
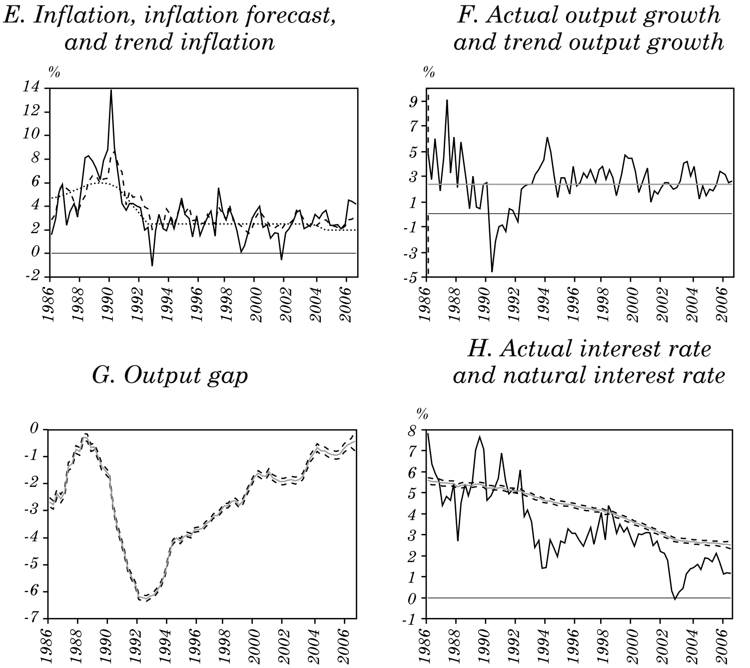


Figure 6. (continued)



Source: Authors' calculations.

a. In panels A and E actual inflation is the solid line, inflation forecast the dashed line and inflation trend the dotted line. Panels A through D correspond to data from 1970:2−2006:4 and panels from E through H correspond to data from 1986:2−2006:4.

Figure 7. Inflation, Output, and the Interest Rate in Australia, 1970:2–2006:4 and 1986:2–2006:4[a]

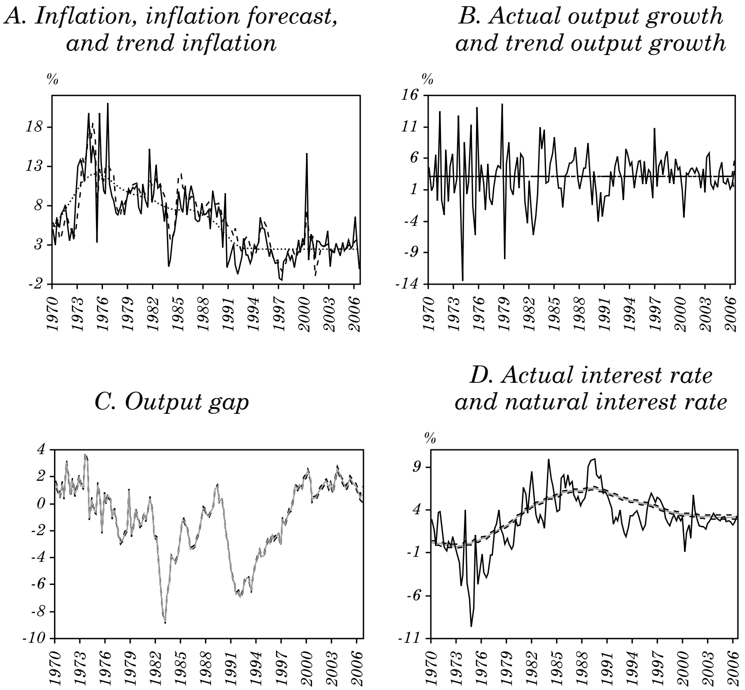
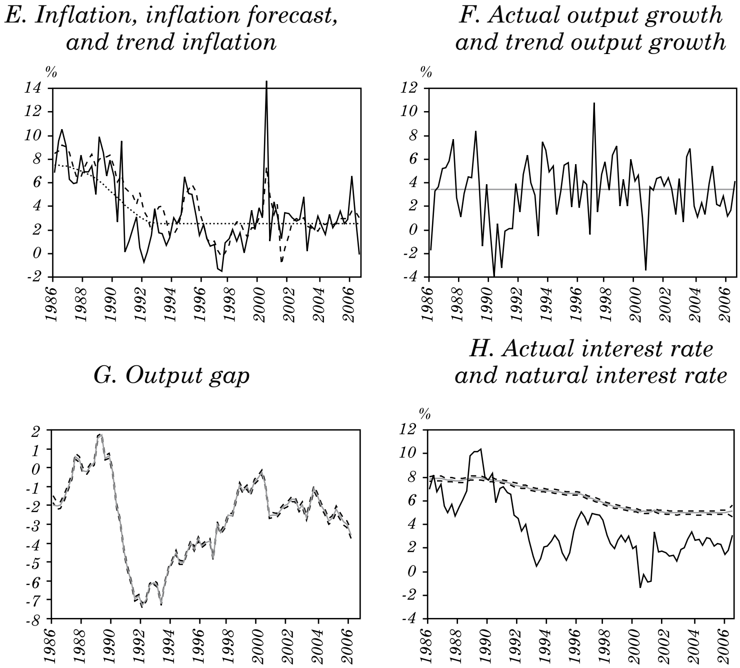


Figure 7. (continued)



Source: Authors' calculations.

a. In panels A and E actual inflation is the solid line, inflation forecast the dashed line and inflation trend the dotted line. Panels A through D correspond to data from 1970:2−2006:4 and panels from E through H correspond to data from 1986:2−2006:4.

Figure 8. Inflation, Output, and the Interest Rate in Sweden, 1970:2–2006:4 and 1986:2–2006:4[a]

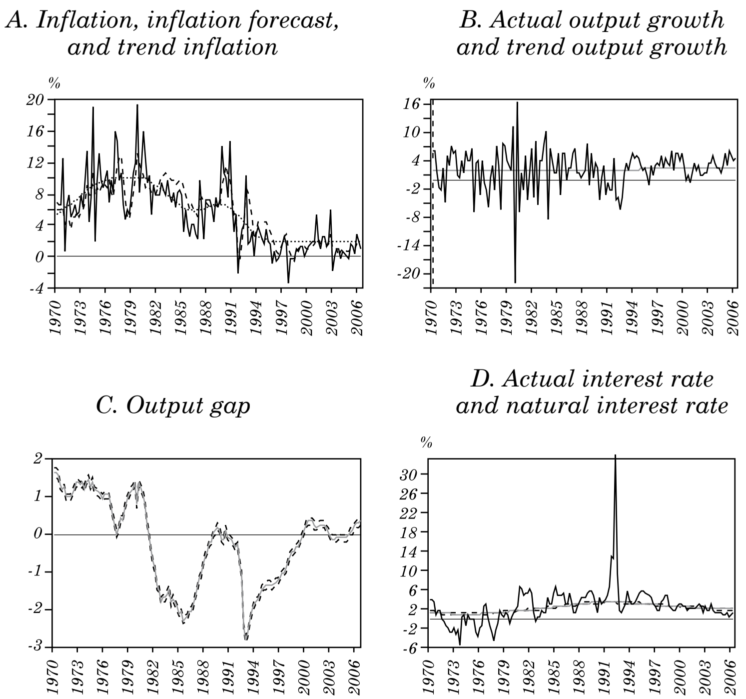
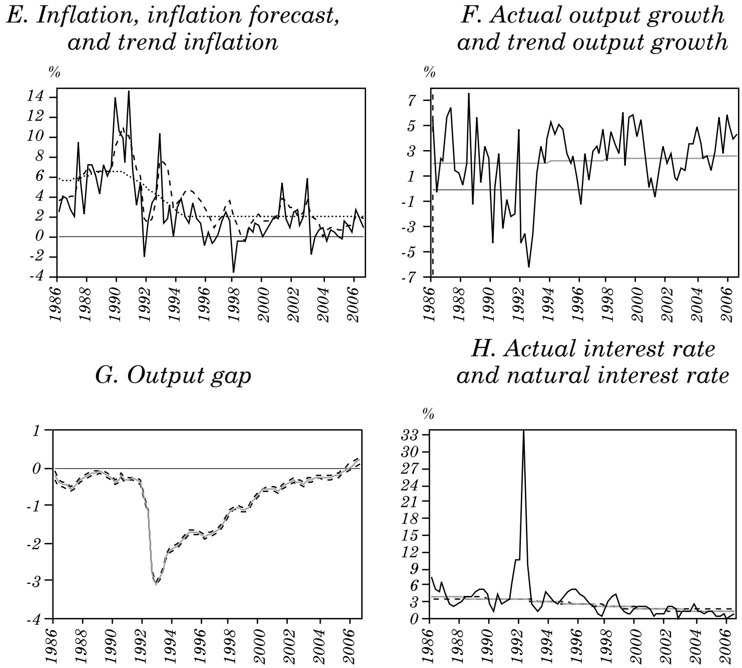


Figure 8. (continued)



Source: Authors' calculations.

a. In panels A and E actual inflation is the solid line, inflation forecast the dashed line and inflation trend the dotted line. Panels A through D correspond to data from 1970:2−2006:4 and panels from E through H correspond to data from 1986:2−2006:4.

Figure 9. Inflation, Output, and the Interest Rate in Norway, 1979:2–2006:4 and 1986:2–2006:4[a]

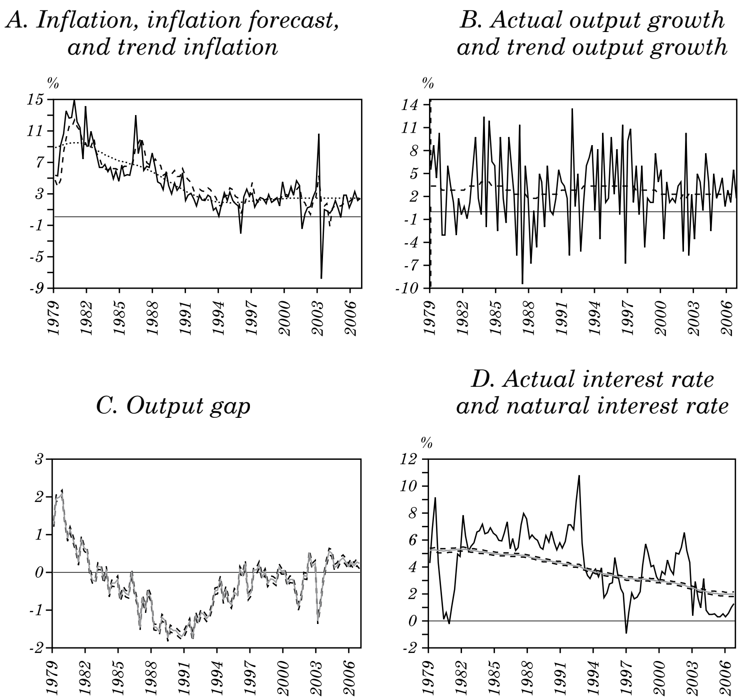
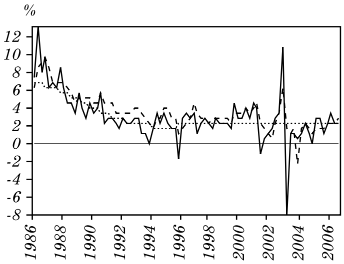
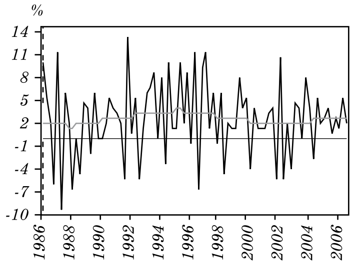


Figure 9. (continued)

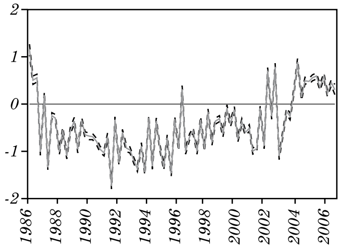
E. Inflation, inflation forecast, and trend inflation



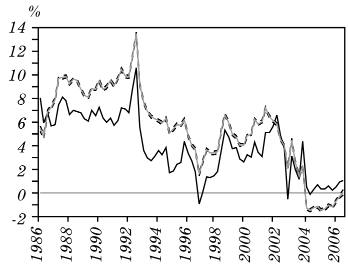
F. Actual output growth and trend output growth



G. Output gap



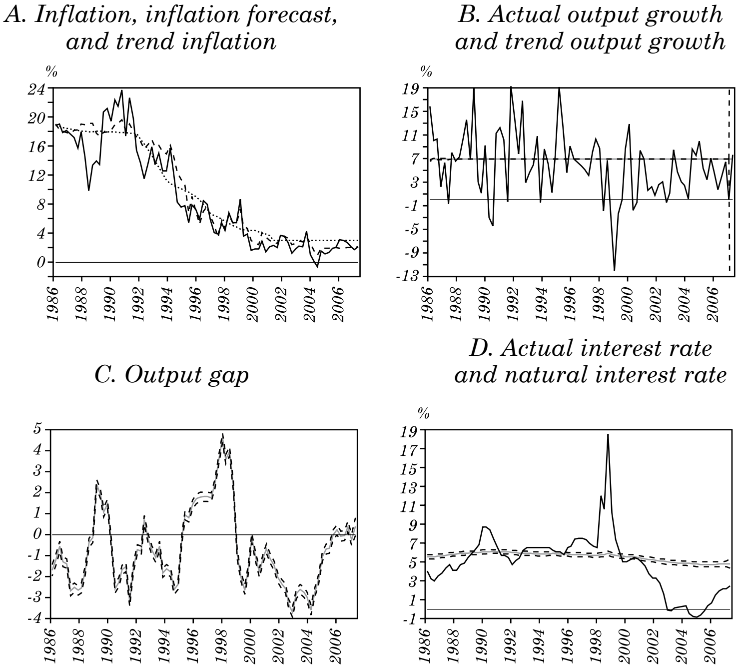
H. Actual interest rate and natural interest rate



Source: Authors' calculations.

a. In panels A and E actual inflation is the solid line, inflation forecast the dashed line and inflation trend the dotted line. Panels A through D correspond to data from 1979:2−2006:4 and panels from E through H correspond to data from 1986:2−2006:4.

Figure 10. Inflation, Output, and the Interest Rate in Chile, 1986:2–2006:4[a]



Source: Authors' calculations.

a. In panel A actual inflation is the solid line, inflation forecast the dashed line and inflation trend the dotted line.

Table 1. Parameter Estimates for the Euro Area, the United States, and Japan[a]

	United States		*Euro Area*		*Japan*			
	1960:01–2007:02	*1986:01–2007:02*	*1970:02–2006:04*	*1986:02–2006:04*	*1970:02 – 2006:04*		*1986 – 2006:04*	
Parameters	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
λ_g	0.0475	0.0612	0.0000	0.0000	0.0000	0.0400	0.0000	0.0400
λ_r	0.0215	0.1399	0.0214	–	–	0.0400	–	0.0400
α_1	0.9492	1.2285	0.9365	0.9740	0.8227	1.0603	0.9753	1.0784
	(0.0351)	(0.1193)	(0.0582)	(0.0183)	(0.0707)	(0.0285)	(0.0077)	(0.0446)
α_2	−0.0710	−0.1355	0.0264	–	–	0.0562	–	0.1030
	(0.0226)	(0.0844)	(0.0325)	(–)	(–)	(0.0282)	(–)	(0.0494)
β_1	−0.0838	−0.0502	0.0144	−0.2482	−0.2137	0.0557	−0.4258	−0.0802
	(0.0565)	(0.0849)	(0.0650)	(0.0794)	(0.0478)	(7711.2291)	(0.0920)	(1557.7195)
β_2	0.8039	1.2426	0.6498	1.1070	0.6607	0.1374	1.3892	−0.0728
	(0.0486)	(0.1173)	(0.0459)	(0.0899)	(0.0309)	(0.0672)	(0.1317)	(0.1139)
β_3	0.4172	−0.3384	−0.0279	0.0481	2.2984	0.5016	0.0563	0.4485
	(0.1189)	(0.1346)	(0.0272)	(0.0593)	(0.4361)	(0.0583)	(0.0308)	(0.1613)
δ_1	0.8632	0.0251	0.3652	–	–	0.0236	–	0.0616
	(0.0233)	(0.1427)	(0.0490)	(–)	(–)	(0.0238)	(–)	(0.0761)

Table 1. (continued)

	United States		*Euro Area*		*Japan*			
	1960:01–2007:02	*1986:01–2007:02*	*1970:02–2006:04*	*1986:02–2006:04*	*1970:02 – 2006:04*		*1986 – 2006:04*	
Parameters	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
δ_2	−0.1329	−0.9141	−0.5706	–	–	−0.7107	–	−0.8420
	(0.0289)	(0.1119)	(0.0506)	(–)	(–)	(0.0336)	(–)	(0.0616)
δ_3	0.1272	2.2387	1.0071	–	–	−2.2838	–	−1.2997
	(0.0752)	(0.5900)	(0.1251)	(–)	(–)	(0.9590)	(–)	(0.3804)
σ_y	0.4831	0.1947	0.3581	0.4267	0.4647	0.2167	0.7196	0.2091
	(0.0951)	(0.0462)	(0.0498)	(0.3034)	(0.1000)	(0.0924)	(0.4655)	(0.0762)
σ_π	0.6790	0.7292	0.7362	0.4680	1.3389	2.2620	1.0289	1.3858
	(0.0319)	(0.0406)	(0.0468)	(0.0401)	(0.1248)	(0.1502)	(0.0859)	(0.1207)
σ_r	1.1502	0.0000	0.6101	–	–	0.3874	–	0.1678
	(0.0317)	(5081.2000)	(0.0384)	(–)	(–)	(0.0688)	(–)	(0.0396)
σ_{y^*}	0.6543	0.4367	0.4776	0.1833	0.8164	0.8946	0.3170	0.8532
	(0.1044)	(3687.0000)	(0.1334)	(0.1583)	(0.1592)	(6510.0068)	(0.6622)	(1304.8673)

Source: Authors' calculations.

a. The estimations presented in columns 1, 2, and 3 are from the third step. Column 4 estimations are from the first step; we did not obtain estimations after the first step due to the matrix singularity problem. For Japan, the estimations in columns 5 and 7 are from the first step, since λ_r is not estimated in the second step when we impose $\lambda_g = 0$ due to the matrix singularity problem. In columns 6 and 8, the estimations are from the third step, where λ_r and λ_g are obtained across a grid search in the interval [0.005; 0.075]. Standard errors are in parentheses.

Table 2. Parameter Estimates for New Zealand and Canada[a]

	New Zealand				*Canada*			
	1974:02 – 2006:04		*1986:02 – 2006:04*		*1970:02 – 2006:04*	*1986:02 – 2006:04*		
Parameters	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
λ_g	0.0544	0.0544	0.0757	0.0757	0.0484	0.0000	0.0484	0.0484
λ_r	0.0000	0.0544	0.0871	0.0757	0.0698	–	0.0698	0.8198
α_1	0.914	0.9462	0.7153	0.6256	0.9598	0.9916	0.8788	0.8773
	(0.0643)	(0.0505)	(0.1345)	(0.0927)	(0.0582)	(0.0187)	(0.0946)	(0.0813)
α_2	–0.0091	0.0203	0.2821	0.2577	–0.0790	–	0.0305	0.0369
	(0.0281)	(0.0396)	(0.0729)	(0.0643)	(0.0291)	(–)	(0.0342)	(0.0464)
β_1	–0.1923	–0.1983	–0.2158	–0.1067	0.0844	–0.3020	–0.2260	–0.2260
	(0.0703)	(0.0685)	(0.221)	(44246.3385)	(14868.7)	(0.0759)	(10298.66)	(10274.93)
β_2	1.4305	1.4288	1.2403	–0.1816	0.0223	1.2527	–0.1199	–0.2318
	(0.0897)	(0.0834)	(0.214)	(0.2006)	(0.0747)	(0.1191)	(0.0886)	(0.0878)
β_3	0.5697	0.5743	0.9306	1.1411	0.6890	0.0739	0.7680	0.8807
	(0.2459)	(0.2219)	(0.2942)	(0.1346)	(0.1242)	(0.2246)	(0.1301)	(0.1433)
δ_1	0.7038	0.5875	0.1262	0.1475	0.7370	–	0.2825	0.1968
	(0.0491)	(0.0472)	(0.0621)	(0.0651)	(0.0420)	(–)	(0.0697)	(0.0684)

Table 2. (continued)

	New Zealand				Canada			
	1974:02 – 2006:04		*1986:02 – 2006:04*		*1970:02 – 2006:04*	*1986:02 – 2006:04*		
Parameters	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
δ_2	−0.3204 (0.0857)	−0.3742 (0.0779)	−0.6219 (0.1614)	−0.5968 (0.1567)	−0.2602 (0.0635)	– (–)	−0.9390 (0.0883)	−0.9290 (0.0796)
δ_3	−0.2211 (0.142)	−0.1838 (0.1383)	−0.1313 (0.1412)	−0.2096 (0.1815)	0.3684 (0.1469)	– (–)	2.2223 (0.4015)	1.5811 (0.3244)
σ_y	1.1969 (0.3918)	1.183 (0.3701)	1.0281 (0.1749)	1.0015 (0.1928)	0.4408 (0.0978)	0.5978 (0.9679)	0.2605 (0.0624)	0.2982 (0.0724)
σ_π	1.5029 (0.1417)	1.4946 (0.1309)	1.5014 (0.2658)	1.5179 (0.2073)	1.3423 (0.0707)	1.1695 (0.0833)	1.3798 (0.1163)	1.2553 (0.1032)
σ_r	2.1501 (0.0847)	2.0427 (0.0697)	1.5995 (0.1426)	1.6071 (0.1417)	1.0691 (0.0576)	– (–)	0.4273 (0.0557)	0.3548 (0.0415)
σ_{y^*}	1.9964 (0.7595)	2.0137 (0.6157)	0.9803 (0.2825)	0.9577 (37739.0760)	0.5649 (11505.45)	0.0000 (185845.55)	0.5019 (7749.17)	0.4724 (7613.58)

Source: Authors' calculations.

a. For New Zealand, the estimations in column 1 are from the second step; we did not obtain estimations in the third step due to the matrix singularity problem. The estimations in column 2 are from the third step, where λ_r is obtained across a grid search in the interval [0.0444; 0.1244]. The estimations in column 3 are from the third step. The estimations in column 4 are from the third step, where λ_r and λ_g are obtained across a grid search in the interval [0.0275; 0.9775]. For Canada, the estimations in column 5 are from the third step. The estimations in column 6 are from the first step, since λ_r is not estimated in the second step when we impose $\lambda_g = 0$, due to the matrix singularity problem. The estimations in column 7 are from the third step, where λ_g and λ_r are obtained in the estimation with the sample 1970-2006. The estimations in column 8 are from the third step, where λ_r is obtained across a grid search in the interval [0.0098; 2.0198]. Standard errors are in parentheses.

Table 3. Parameter Estimates for the United Kingdom and Australia[a]

	United Kingdom				*Australia*			
	1970:02 – 2006:04		*1986:02 – 2006:04*		*1970:02–2006:04*	*1986:02 – 2006:04*		
Parameters	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
λ_g	0.0275	0.0275	0.0000	0.0275	0.0000	0.0069	0.0069	0.0569
λ_r	–	0.0900	0.0000	0.0600	0.0522	0.0000	0.0522	0.0522
α_1	0.8796	0.6669	0.9776	0.9854	0.9363	0.9669	0.9906	0.9291
	(0.0575)	(0.1249)	(0.0345)	(0.0156)	(0.0415)	(0.05)	(0.0432)	(0.1031)
α_2	–	0.0407	−0.036	−0.0490	0.0022	−0.0237	−0.0036	0.0062
	(–)	(0.0195)	(0.0388)	(0.0427)	(0.0321)	(0.0303)	(0.0453)	(0.0247)
β_1	−0.1142	−0.169	−0.2266	−0.2245	−0.2231	−0.4366	−0.4316	−0.2872
	(0.0601)	(0.0759)	(0.1021)	(0.1017)	(0.0553)	(0.1275)	(0.1214)	(15262.2512)
β_2	0.9532	0.8837	1.3391	1.3271	1.0026	1.3629	1.3581	−0.4366
	(0.0391)	(0.0545)	(0.1148)	(0.0984)	(0.0979)	(0.117)	(0.1111)	(0.1165)
β_3	1.0792	2.4103	0.2045	0.2063	0.3114	0.3246	0.3191	1.2311
	(0.3806)	(0.7842)	(0.1227)	(0.0891)	(0.114)	(0.1883)	(0.1497)	(0.1387)
δ_1	–	0.4519	0.8953	0.7431	0.7168	0.8507	0.7758	0.7554
	(–)	(0.0331)	(0.0555)	(0.0694)	(0.0481)	(0.0526)	(0.0706)	(0.0773)

Table 3. (continued)

	United Kingdom				Australia			
	1970:02 – 2006:04		*1986:02 – 2006:04*		*1970:02–2006:04*	*1986:02 – 2006:04*		
Parameters	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*
δ_2	–	–0.7096	–0.1097	–0.0995	–0.3327	–0.2668	–0.2945	–0.3612
	(–)	(0.046)	(0.0935)	(0.0805)	(0.0496)	(0.1081)	(0.0932)	(0.0946)
δ_3	–	0.6368	0.0282	0.0523	0.0438	0.1157	0.1345	0.6577
	(–)	(0.2713)	(0.0403)	(0.0497)	(0.0752)	(0.0976)	(0.0931)	(0.3732)
σ_y	0.6381	0.4554	0.1404	0.4713	1.0046	0.616	0.6615	0.3051
	(0.1161)	(0.1017)	(0.159)	(0.1190)	(0.1178)	(0.1922)	(0.1262)	(0.0995)
σ_π	1.77	1.5288	0.8628	0.8644	2.0193	1.4473	1.4495	1.4616
	(0.1337)	(0.1767)	(0.0611)	(0.0599)	(0.1177)	(0.1283)	(0.126)	(0.1383)
σ_r	–	1.6097	0.788	0.7557	1.6796	0.9827	0.9362	0.8986
	(–)	(0.0818)	(0.0568)	(0.0563)	(0.0757)	(0.083)	(0.0731)	(0.0756)
σ_{y^*}	0.6383	0.7789	0	0.0000	0.0000	0.2168	0.0008	0.5638
	(0.1174)	(0.284)	(4092)	(3802.0530)	(12158.6)	(0.3725)	(157.37)	(11804.8755)

Source: Authors' calculations.

a. For the United Kingdom, the estimations in column 1 are from the first step, since λ_r is not estimated in the second step when we impose $\lambda_g = 0.0275$ due to the matrix singularity problem. The estimations in column 2 are from the third step, where λ_r is obtained across a grid search in the interval [0.0444; 0.1244]. The estimations in column 3 are from the second step. We did not obtain estimations in the third step due to the matrix singularity problem. The estimations in column 4 are from the third step, where λ_g is obtained in the estimation with the sample 1970-2006 and λ_r is obtained across a grid search in the interval [0.055; 0.065]. For Australia, the estimations in column 5 are from the third step. The estimations in column 6 are from the second step. We did not obtain estimations in the third step due to the matrix singularity problem. The estimations in column 7 are from the third step, where λ_r is obtained in the estimation with the sample 1970-2006. The estimations in column 8 are from the third step, where λ_r and λ_g are obtained across a grid search in the interval [0.0275; 0.9775]. Standard errors are in parentheses.

Table 4. Parameter Estimates for Sweden and Norway[a]

	Sweden			*Norway*			
	1970:02–2006:04	*1986:02 – 2006:04*		*1979:02 – 2006:04*		*1986:02 – 2006:04*	
Parameters	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
λ_g	0.0262	0.0000	0.0262	0.0677	0.0677	0.1186	0.1186
λ_r	0.0315	–	0.0315	0.000	0.040	0.000	0.040
α_1	0.9177	0.9913	0.9403	0.9236	0.9375	0.0072	–0.7573
	(0.0478)	(0.0274)	(0.0522)	(0.0405)	(0.0613)	(0.2289)	(0.1780)
α_2	–0.0452	–	–0.0558	–0.0958	–0.0050	–0.1925	0.5243
	(0.0190)	(–)	(0.0110)	(0.0658)	(0.0208)	(0.1273)	(0.2039)
β_1	–0.1680	–0.3390	–0.0646	–0.3339	–0.1700	–0.3609	–0.3064
	(16775.9)	(0.0888)	(11442.1623)	(0.0444)	(16845.7489)	(0.0572)	(25668.8387)
β_2	–0.3429	1.3353	–0.2998	1.4904	–0.2921	1.2578	–0.3553
	(0.0594)	(0.1098)	(0.1031)	(0.0928)	(0.0500)	(0.0891)	(0.0531)
β_3	1.3183	0.2620	1.3436	0.3326	1.5101	0.2158	1.1926
	(0.1133)	(0.3898)	(0.1289)	(0.1267)	(0.0997)	(0.2943)	(0.0766)
δ_1	0.5615	–	0.3929	0.7958	0.6415	0.8777	0.9868
	(0.0292)	(–)	(0.0581)	(0.0485)	(0.0615)	(0.0708)	(0.0115)

Table 4. (continued)

	Sweden			Norway			
	1970:02–2006:04	*1986:02 – 2006:04*		*1979:02 – 2006:04*		*1986:02 – 2006:04*	
Parameters	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
δ_2	−0.4751	–	−0.7365	−0.3852	−0.5778	−0.4053	−0.4583
	(0.1683)	(–)	(0.4081)	(0.0842)	(0.0919)	(0.1509)	(0.1064)
δ_3	−0.4555	–	−0.5290	−0.1139	−1.1227	−0.2346	−0.4790
	(0.3784)	(–)	(1.0947)	(0.0560)	(0.2599)	(0.2826)	(0.1677)
σ_y	0.3447	0.6823	0.1191	0.9041	0.2402	0.7054	0.3312
	(0.1196)	(0.4974)	(0.1642)	(0.2227)	(0.0890)	(0.1376)	(0.1086)
σ_π	1.9639	1.6336	1.7076	1.3759	1.4408	1.4839	1.2810
	(0.1272)	(0.1287)	(0.1579)	(0.0770)	(0.0894)	(0.1064)	(0.0977)
σ_r	2.6759	–	3.1712	1.1974	1.0270	1.2259	0.3762
	(0.0620)	(–)	(0.1470)	(0.0727)	(0.0693)	(0.0822)	(0.2991)
σ_{y^*}	0.9841	0.0000	0.5951	0.7633	1.1710	0.5428	0.8557
	(14365.4)	(106856.8391)	(8956.6283)	(0.4370)	(16930.1425)	(3.2854)	(21516.3255)

Source: Authors' calculations.

a. For Sweden, the estimations in column 1 are from the third step. The estimations in column 2 are from the first step, since λ_r is not estimated in the second step when we impose $\lambda_g = 0$ due to the matrix singularity problem. The estimations in column 3 are from the third step, where λ_r is obtained in the estimation with the sample 1970-2006. For Norway, the estimations in column 4 are from the second step. We did not obtain estimations in the third step due to the matrix singularity problem. The estimations in column 5 are from the third step, where λ_r is obtained across a grid search in the interval [0.0050; 0.0750]. The estimations in column 6 are from the second step. We did not obtain estimations in the third step due to the matrix singularity problem. The estimations in column 7 are from the third step, where λ_r is obtained across a grid search in the interval [0.0050; 0.0750]. Standard errors are in parentheses.

Table 5. Parameter Estimates for Chile[a]

	Chile	
	1986:02 – 2006:04	
Parameters	*(1)*	*(2)*
λ_g	0.0000	0.0820
λ_r	0.0000	0.0800
α_1	1.0771 (0.0540)	0.9412 (0.1074)
α_2	−0.2461 (0.1245)	-0.1076 (0.0961)
β_1	0.4639 (0.0697)	0.4325 (0.0946)
β_2	0.5078 (0.1612)	0.5940 (0.1959)
β_3	0.0142 (0.0251)	0.2756 (0.2216)
δ_1	0.6996 (0.1242)	0.6552 (0.0861)
δ_2	−0.0151 (0.2658)	0.1188 (0.2049)
δ_3	0.0733 (0.0809)	0.3680 (0.2525)
σ_y	1.2847 (0.9877)	1.0436 (0.2924)
σ_π	1.8274 (0.1110)	1.7188 (0.1230)
σ_r	1.3993 (0.0750)	1.2777 (0.0833)
σ_{y^*}	0.0001 (8810.1)	0.7456 (0.3177)

Source: Authors' calculations.

a. The estimations in column 1 are from the second step; we did not obtain estimations in the third step due to the matrix singularity problem. The estimations in column 2 are from the third step, where λ_g and λ_r are obtained across a grid search in the intervals [0.062; 0.102] and [0.06; 0.10], respectively. Standard errors are in parentheses.

4. Extensions for the United States and Chile

In this section, we extend our basic model to include the unemployment gap (Okun's Law) and apply it to the United States and Chile, for which we obtained the best results for the basic model. We also test for robustness of the basic model results for the United States by replacing four-step-ahead inflation forecasts with eight-step-ahead forecasts.[15]

4.1 Results for the United States

For the extended model with Okun's Law for the United States, we proceed in the following way. When freely estimating all parameter values and unobservables, λ_u was estimated in the fourth step at a value of zero, implying a constant 5.6 percent natural rate of unemployment for the United States in 1960–2007. Following the approach adopted for countries in section 3, we next pursue a grid search over alternative preset values of λ_u. The model parameter estimates consistent with $\lambda_u = 0$ and $\lambda_u = 0.4$ (the median value of our grid search) are reported in columns 1 and 2 of table 6. Figure 11 depicts the grid-search results for the unobservables. The findings can be summarized as follows. The parameter estimates are generally similar for the extended model (in both columns 1 and 2 of table 6) to those reported for the basic model (column 1, table 1). In the IS curve, the output gap becomes more sensitive to the lagged interest rate gap, while the coefficient of lagged inflation in the Phillips curve turns positive, with a corresponding reduction in size of the two other Phillips curve coefficients. For the newly introduced Okun's Law, the parameter estimates exhibit the expected signs and are highly significant. The parameter estimate for the lagged unemployment gap reflects large unemployment inertia. The coefficient estimate of the lagged output gap is very large (–0.95) when the natural unemployment rate is estimated as constant and declines to –0.35 when the natural unemployment rate is variable, consistent with a value of λ_u set at 0.4.

Figure 11 depicts estimation ranges for unobservables for λ_u varying between 0.08 and 0.72. The estimates for both potential

15. We did not obtain model convergence when using eight-step-ahead inflation forecasts for Chile. We also conducted sensitivity analyses for the Phillips curve in both countries, by replacing one-period inflation lags with four-quarter lags; the results were almost unchanged.

Table 6. Parameter Estimates for Alternative Models for the United States and Chile[a]

	United States 1960:1 – 2007:2			*Chile 1986:2 – 2006:4*	
	Extended Model (with Okun's law)		*Eight-step-ahead inflation forecasts*	*Extended Model (with Okun's law)*	
Parameters	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*
λ_g	0.0475	0.0475	0.0586	0.0000	0.0820
λ_r	0.0215	0.0215	0.0304	0.0000	0.0800
λ_u	0.0000	0.4000	–	0.0000	0.4000
α_1	0.9539 (0.0302)	0.9558 (0.0331)	0.9503 (0.0441)	1.0033 (0.0515)	1.0329 (0.0433)
α_2	−0.0252 (0.0100)	−0.0681 (0.0213)	−0.0546 (0.0216)	−0.0644 (0.0425)	−0.1583 (0.0685)
β_1	0.1097 (0.0599)	0.0602 (0.0593)	0.0680 (0.1031)	0.4501 (0.0803)	0.4533 (0.0842)
β_2	0.6525 (0.0525)	0.7032 (0.0474)	0.4514 (0.0482)	0.5191 (0.1703)	0.5182 (0.1687)
β_3	0.3926 (0.1876)	0.2820 (0.0968)	0.4337 (0.1427)	0.1474 (0.1614)	0.1173 (0.1420)
γ_1	0.4956 (0.0999)	0.5635 (0.0879)	– (–)	0.2501 (0.1791)	0.2045 (0.2190)
γ_2	−0.9466 (0.3234)	−0.3523 (0.1010)	– (–)	−0.6591 (0.3348)	−0.5356 (0.2237)

Table 6. (continued)

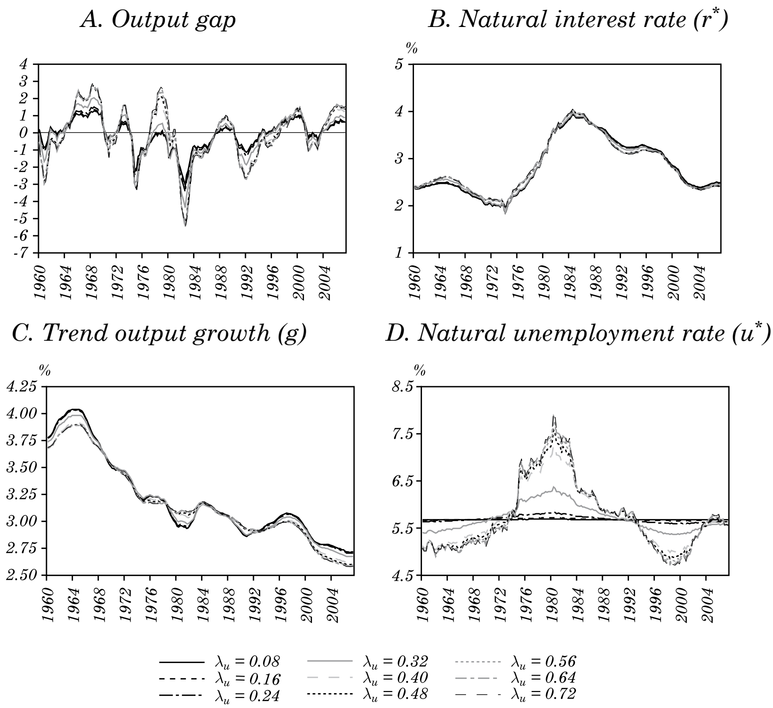
	United States 1960:1 – 2007:2			*Chile 1986:2 – 2006:4*	
	Extended Model (with Okun's law)		*Eight-step-ahead inflation forecasts*	*Extended Model (with Okun's law)*	
Parameters	*(1)*	*(2)*	*(3)*	*(4)*	*(5)*
δ_1	0.8756 (0.0316)	0.8697 (0.0256)	0.7880 (0.0262)	0.7821 (0.0600)	0.6996 (0.0724)
δ_2	−0.1478 (0.0286)	−0.1353 (0.0298)	−0.2193 (0.0201)	0.0205 (0.2750)	−0.0073 (0.2139)
δ_3	0.1731 (0.1587)	0.1250 (0.0825)	0.1910 (0.1075)	0.3329 (0.2328)	0.2654 (0.1804)
σ_y	0.2411 (0.0780)	0.4731 (0.1053)	0.5176 (0.1060)	0.5644 (0.2246)	0.5899 (0.1810)
σ_π	0.8223 (0.0385)	0.7750 (0.0340)	0.8250 (0.0411)	1.8052 (0.1135)	1.8071 (0.1175)
σ_u	0.0442 (0.0643)	0.1253 (0.0144)	– (–)	0.1935 (0.0971)	0.2151 (0.0671)
σ_r	1.1552 (0.0283)	1.1498 (0.0316)	1.2768 (0.0352)	1.3852 (0.0743)	1.3135 (0.0704)
σ_{y^*}	0.7969 (0.3020)	0.6656 (0.1485)	0.6293 (0.1288)	1.2730 (0.6356)	1.1429 (0.5518)

Source: Authors' calculations.

a. For the United States, the estimations in column 1 are from the fourth step of the extended model with Okun's Law. The estimations in column 2 are from the fourth step, where λ_u is obtained from a grid search in the interval [0.08; 0.72]. The estimations in column 3 are from the third step of the modified standard model with eight-step-ahead inflation forecast. For Chile, the estimations in column 4 are from the fourth step. The estimations in column 5 are from the fourth step, where λ_g, λ_r and λ_u are obtained from a grid search in the intervals [0.062; 0.102], [0.06; 0.10], and [0.08; 0.72], respectively. Standard errors are in parentheses.

output growth (which declines from 3.8 percent in the early 1960s to 2.8 percent in the early 2000s) and the natural interest rate (which varies between 2 percent and 4 percent between 1960 and 2006) are robust to changes in λ_u, reflected in their narrow ranges. Moreover, the estimated values and dynamics of both potential growth and the natural interest rate for the extended model are very close to those depicted for the basic model (upper panel, figure 1). However, the range of estimates for the output gap for different values of λ_u is larger. In addition, the median value for the new output gap estimate is not as close to the estimate for the basic model. This should not come as a surprise, since the extended model imposes a close relation between the output gap and the unemployment

Figure 11. Grid-Search Results for the Extended Model for the United States, 1960:1–2007:2[a]



Source: Authors' calculations.

a. The panels show the unobservables for different grid values of λ_u.

gap. Okun's Law implies that the latter gaps are almost a mirror image of each other.

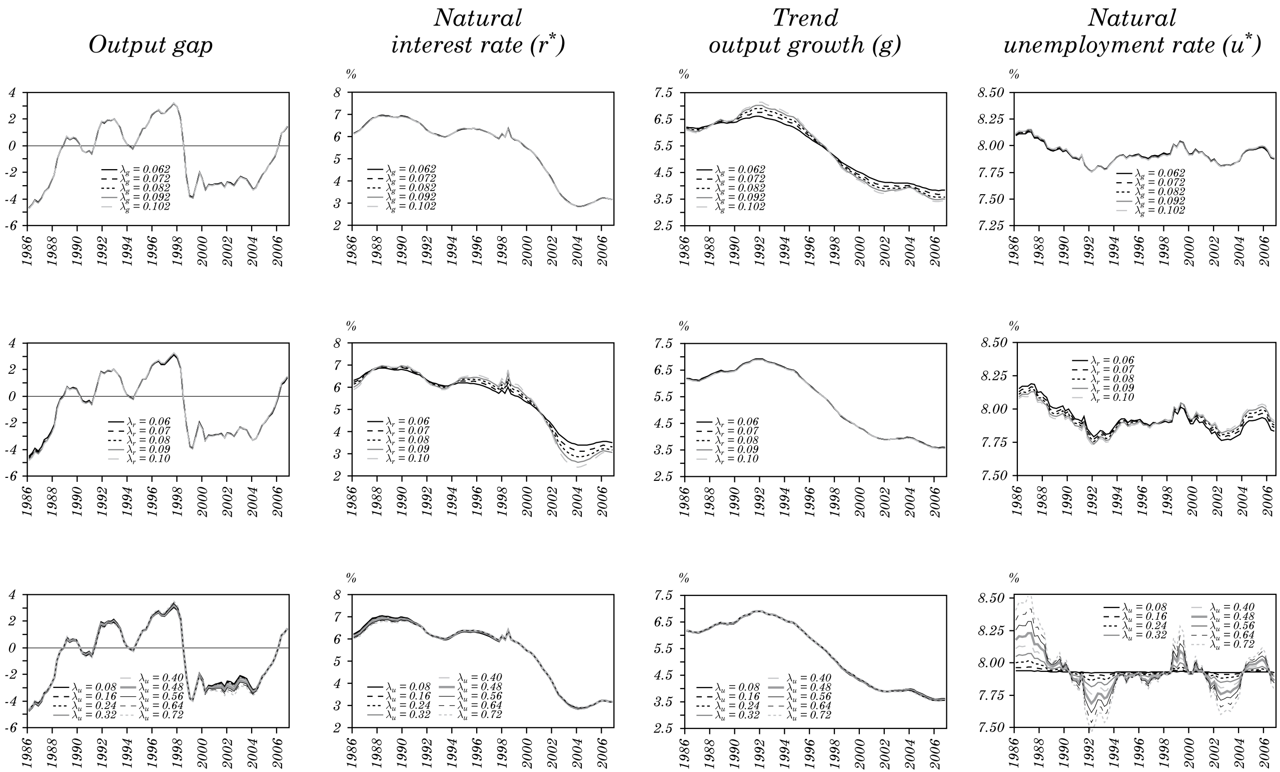
The largest range of estimates depicted in figure 11 is the one for the newly estimated natural rate of unemployment. For the median value of λ_u, the natural rate varies over time between 5.1 percent and 7.2 percent. Over the full range of λ_u values, the natural rate varies over time between 4.8 percent and 8.1 percent. This is consistent with recent findings of King and Morley (2007), who estimate the natural rate as the steady-state of a VAR and attribute most of the volatility in observed unemployment to movements in the natural rate.

We now return to the parsimonious model, replacing the four-step-ahead inflation forecast for the United States with an eight-step-ahead forecast. This change affects the measurement of inflation expectations in the three structural model equations. We obtain the following results for parameter estimates (column 3, table 6). First, the IS curve parameter estimates are not modified much (for comparison, see column 1, table 1). The parameter estimate for the inflation expectations gap in the Phillips curve declines almost by half, but it remains significant. The parameter estimate for the inflation forecast gap in the Taylor rule is still significant, but it is somewhat more negative, implying a corresponding decline in the nominal interest rate reaction to an inflation expectations shock, from +0.87 to +0.78. Both results—for the Phillips curve and the Taylor rule—may suggest that four-quarter-ahead inflation expectations describe inflation and interest rate setting better than eight-quarter-ahead inflation expectations. Finally, with regard to unobservables, the output gap, the neutral interest rate, and potential output growth exhibit similar patterns and values as those based on four-step-ahead inflation forecasts.

4.2 Results for Chile

For the extended model with Okun's Law for Chile, we proceed in a way similar to our approach with the United States. However, the difference is that when freely estimating all parameter values and unobservables, λ_g, λ_r, and λ_u are estimated at zero in the fourth-stage estimation. Therefore, we conduct separate grid searches over alternative preset values of the three signal-to-noise coefficients. The model parameter estimates consistent with $\lambda_g = \lambda_r = \lambda_u = 0$, and with $\lambda_g = 0.082$, $\lambda_r = 0.080$, and $\lambda_u = 0.4$ (the median values of

Figure 12. Grid-Search Results for the Extended Model for Chile, 1986:2–2006:4[a]



Source: Authors' calculations.

a. The panels show the unobservables for different grid values of λ_g (first row), λ_r (second row), and λ_u (third row).

our grid searches) are reported in columns 4 and 5, respectively, of table 6. Figure 12 depicts the corresponding grid-search results for the unobservables. We find that the parameter estimates for the extended model (columns 4 and 5 in table 6) are generally very similar to those reported for the basic model (corresponding columns 1 and 2 in table 5). The one important exception is the IS curve, where the output gap becomes more sensitive (and significant) to the lagged interest rate gap in the extended model (that is, the lambdas are set at positive values). The coefficient of lagged inflation in the Phillips curve now turns positive, with a corresponding reduction in size of the two other Phillips curve coefficients. For the newly introduced Okun's Law, parameter estimates exhibit the expected signs and are highly significant. The parameter estimates for the lagged unemployment gap reflect moderate unemployment inertia, while the coefficient estimate of the lagged output gap is large (close to -0.6).

The estimation ranges depicted in the three rows of figure 12 are relatively narrow for all unobservable variables. The widest range in each row is for the unobservable over which the grid search is conducted. The general dynamic pattern of three unobservables (namely, potential output growth, the output gap, and the neutral interest rate) estimated for the extended model are similar to those obtained for the basic model. Potential output growth is estimated to have declined from 6.5 percent in the late 1980s and early 1990s to 3.5 percent in the early 2000s. The neutral interest rate follows a very similar pattern, falling from 6.5 percent in the late 1980s and early 1990s to 3 percent in the early 2000s.

As in the case of the extended model applied to the United States, the differences in output gap estimates are not surprising, as the extended model imposes a close relation between the output gap and the unemployment gap. Again, Okun's Law implies that the latter gaps are almost a mirror image of each other. However, in contrast to the United States, the range for the new estimates of the natural rate of unemployment is not as large in Chile. For the median value of λ_u, the natural rate varies over time between 7.7 percent and 8.1 percent. Over the full range of λ_u values, the natural rate varies over time between 7.5 percent and 8.5 percent. This is consistent with recent findings by Restrepo (2008) based on different models of estimation for the NAIRU in Chile.

5. The Great Moderation, Comovements, and Convergence in Industrial Economies

The period of low inflation and low volatility in key macroeconomic variables beginning in the late 1980s, following the high inflation and real instability of the mid-1970s and early 1980s, is sometimes called the Great Moderation. It has been documented fairly extensively in academic research and policy evaluations.[16] At the same time, there is a presumption that rising world trade and financial integration should lead to stronger business cycle comovement across countries, as well as stronger convergence in real variables, like growth and real interest rates, particularly among industrial countries. In this section, we exploit our country time-series estimates of unobservables, in addition to the series of selected observables, to test for the Great Moderation, comovements, and convergence in our sample of nine industrial countries, using quarterly data for 1970–2006.[17]

5.1 Common Trends in Key Unobservables

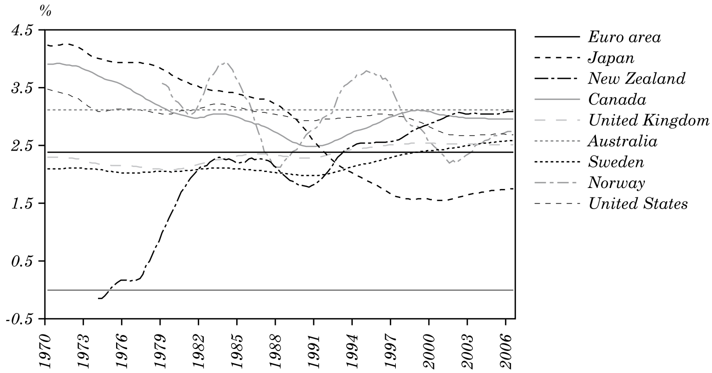
We start by describing the trends in potential output growth rates (figure 13) and neutral real interest rates (figure 14) across the nine countries. The most striking feature of the potential output growth estimates is the large reduction in cross-country variation observed between 1970 and 2006. Leaving out Japan, country point estimates of potential growth ranged from zero (New Zealand) to 4 percent (Canada) in the early 1970s. In contrast, the range of potential growth estimates for 2006 was quite narrow, delimited by Japan's low potential growth rate (1.8 percent) and Australia's constant rate (3.2 percent). The most striking increase in potential growth is New Zealand, where potential growth rose from zero to 3.2 percent in the last four decades; this stands in sharp contrast to Japan's reduction from 4.5 percent to 1.8 percent. Sweden and the United Kingdom exhibit a slight trend rise in potential growth, with the opposite pattern observed in Canada, Norway, and the United States.

Similar to the case of growth, the cross-country dispersion in neutral real interest rates has declined strongly in the last four

16. For example, the International Monetary Fund's October 2006 W*orld Economic Outlook* devotes a well-documented chapter to the Great Moderation.

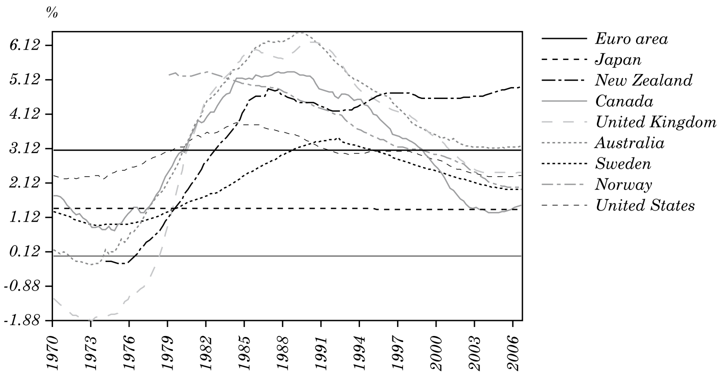
17. We use our shorter time series for New Zealand and Norway, and we drop Chile due to the lack of quarterly data before 1986.

Figure 13. Potential Output Growth in Nine Economies, 1970:2–2006:4[a]



Source: Authors' calculations.
a. The sample period starts in 1974:2 for New Zealand and 1979:2 for Norway.

Figure 14. Neutral Real Interest Rate in Nine Economies, 1970:2–2006:4[a]



Source: Authors' calculations.
a. The sample period starts in 1974:2 for New Zealand and 1979:2 for Norway.

decades (figure 14). In the early 1970s, neutral real rates ranged from –1.9 percent (United Kingdom) to 3.1 percent (euro area). By 2006, the range had narrowed to an interval from 1.5 percent (Japan) to 3.1 percent (euro area), except for New Zealand. Six countries exhibit an inverted-U-shaped pattern of their neutral real interest rates. This reflects strong monetary adjustment in response to the

Great Inflation of the late 1970s, with real policy rates peaking in the 1980s and early 1990s at levels of up to 6.5 percent (Australia in 1990). The stabilization success of the 1980s and 1990s that led to the Great Moderation allowed for subsequent lower neutral rates in the 1990s and 2000s. The exception to the latter trend is New Zealand, where the neutral real interest rate continued to rise, reaching 4.8 percent in 2006.

5.2 The Great Moderation

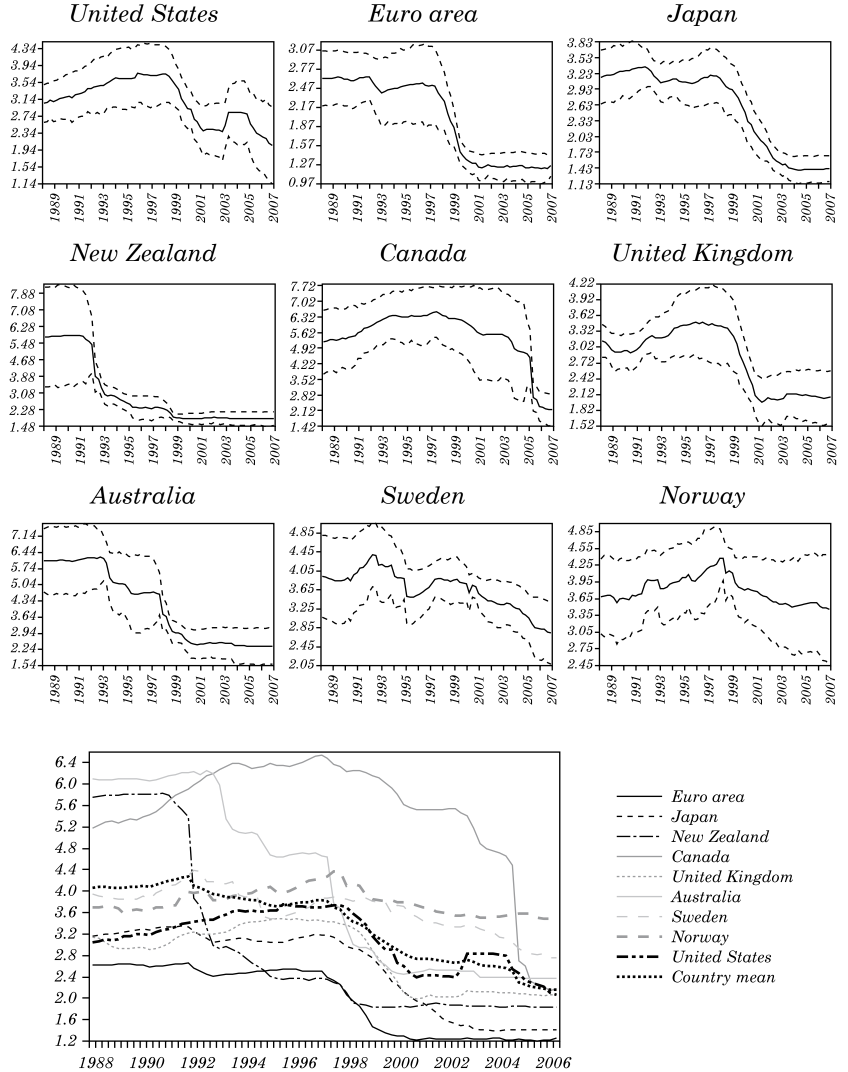
To investigate the Great Moderation, we focus on volatility trends of seven key variables in our nine sample countries. Three variables are observables (inflation, output growth, and the real interest rate) and four are unobservables (potential output growth, the output gap, the natural real interest rate, and the interest rate gap). We compute rolling standard deviations for the latter variables using a window of seventy-four quarters.[18] We then report the associated confidence intervals obtained by bootstrap techniques.[19]

This approach is informative about the Great Moderation, reflected in increased stability of key macroeconomic variables. We focus on both the level of the rolling standard deviation and the varying width of the confidence interval. The results are depicted separately for each variable in figures 15 through 21. The nine smaller panels in each figure show rolling point estimates of the standard deviation and their estimated time-varying confidence intervals for each country, while the larger bottom panel depicts the nine point estimates for each country and the corresponding country mean to better represent the common volatility trend across our sample countries. We find that the volatility of inflation has declined in all countries, except Norway; the mean volatility of inflation fell from 4.0 percent in 1970–87 to 2.2 percent

18. We use a window size of seventy-four quarters (or eighteen and a half years), which is half our thirty-seven-year sample coverage from 1970 to 2006. We choose this rather large window to show more clearly long-term volatility trends, avoiding excessive noise in standard deviations that shows up when using conventional forty-quarter (ten-year) rolling windows.

19. We apply a bootstrap technique for estimating time-varying confidence intervals because of its superior asymptotic properties in small samples, in comparison with standard confidence intervals. Hall's confidence intervals are calculated using the stationary bootstrap method of Politis and Romano (1994). This technique guarantees stationary artificial series by allowing a random block size (indeed, it follows a geometric distribution) when resampling the data. We set the mean of the block size at three and perform 2,000 replications.

Figure 15. Inflation Volatility Trends in Nine Economies, 1970:2–2006:4[a]



Source: Authors' calculations.

a. The sample period starts in 1974:2 for New Zealand and 1979:2 for Norway. The window size for the rolling estimations is seventy-four quarters. For instance, the first point estimate corresponds to 1988:3, which is based on the period 1970:2–1988:3.

in 1988–2006 (figure 15).[20] Moreover, this trend is also significant as reflected by the narrowing confidence intervals. The exception is again Norway, where point estimates decline while confidence intervals rise after 1988. The largest reductions in inflation volatility are observed in Australia, Canada, and New Zealand, roughly from 6.0 percent to about 2.2 percent. The euro area exhibits the lowest inflation volatility during most of the sample period.

The reduction of the volatility of output growth in all nine countries is remarkable, reflected by both declining point estimates and narrowing confidence intervals (figure 16). The average country level of output growth volatility fell roughly by half, from 5.0 percent in 1970–87 to 2.7 percent in 1988–2006. The largest growth stabilization was recorded in New Zealand, where growth volatility fell from 14 percent in the 1970s and 1980s to 5 percent in the 1990s and 2000s. Australia, Sweden, and the United Kingdom also exhibit large reductions in growth volatility. Again the euro area exhibits the highest level of stability throughout the last thirty-seven years.

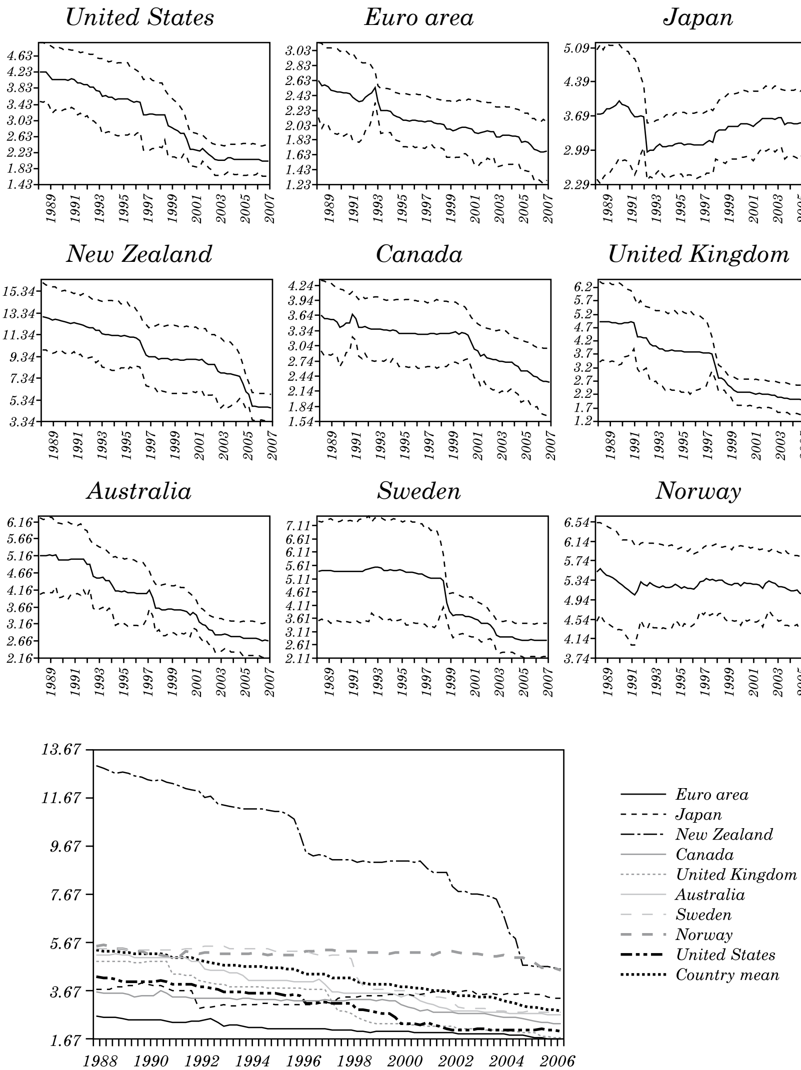
We now turn to our first unobservable, potential output growth.[21] Like all estimated unobservables, potential growth either is estimated as a constant (in the euro area and Australia) or, if variable (in the other countries), exhibits a smooth pattern over time, without high-frequency volatility (figure 17). Therefore, its volatility—like that of the neutral interest rate, reported below—is lower by an order of magnitude than the volatilities exhibited by observable variables. The average country volatility (for the seven countries where potential output varies over time) declines only marginally over time. Opposite trends are observed in different countries. For example, New Zealand records a strong trend decline in potential growth volatility, while a growing trend is observed in Japan up to 2000, which is partially reversed thereafter.

The average country volatility of the output gap (our second unobservable) falls slightly, from 1.6 percent in 1970–87 to 1.4 percent in 1988–2006 (figure 18). There are moderate to large reductions in the volatility of the output gap in six countries, no clear trends in two countries, and a slight trend rise in one country (Australia). The

20. The correlation between the first and second moments of inflation is known to be very large. Hence, the declining trends in inflation volatility described here are matched by declining trends in inflation levels.

21. The descriptive statistics discussed below for our estimates of unobservable are conditional on our estimates and should thus be taken with caution, in comparison with those reported for observables like inflation, actual growth, and actual interest rates.

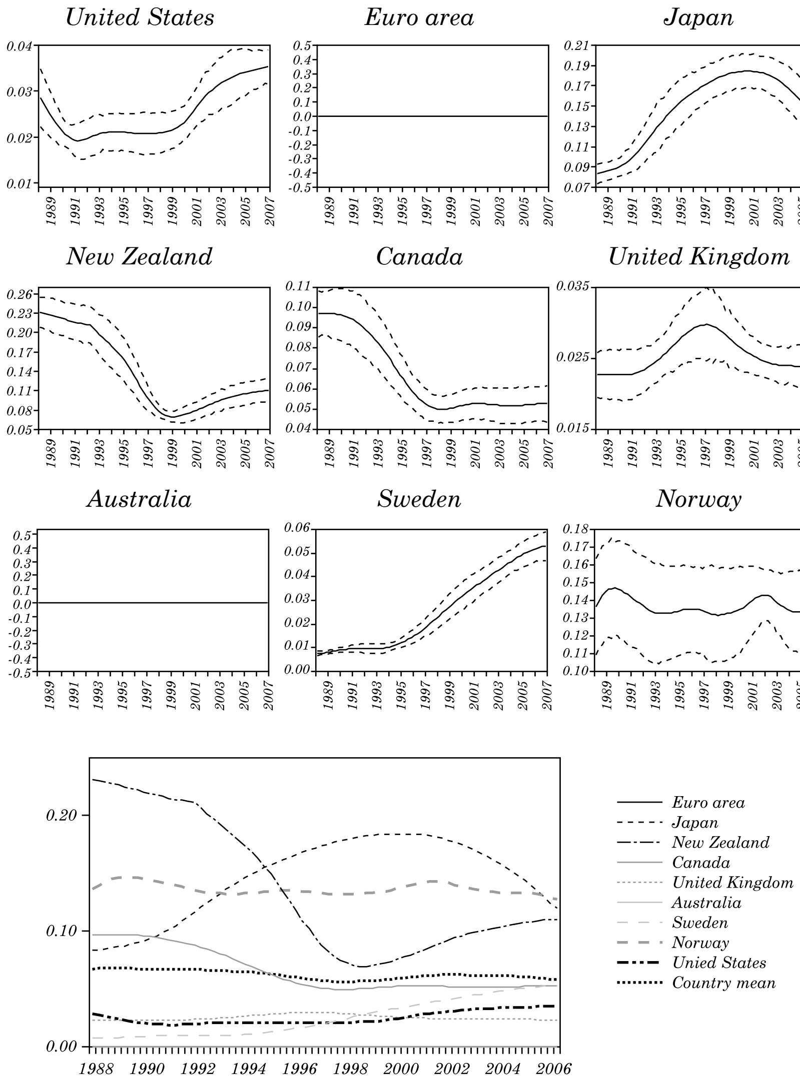
Figure 16. Actual Output Growth Volatility Trends in Nine Economies, 1970:2–2006:4[a]



Source: Authors' calculations.

a. The sample period starts in 1974:2 for New Zealand and 1979:2 for Norway. The window size for the rolling estimations is seventy-four quarters. For instance, the first point estimate corresponds to 1988:3, which is based on the period 1970:2–1988:3.

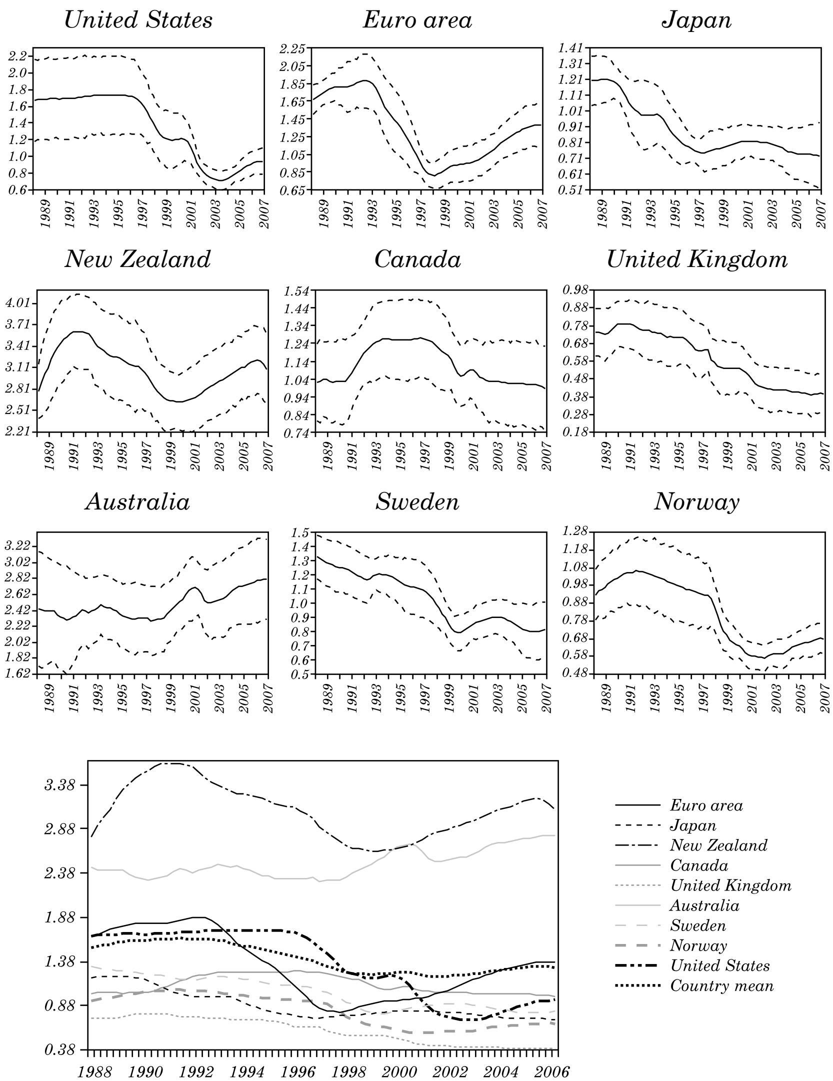
Figure 17. Potential Output Growth Volatility Trends in Nine Economies, 1970:2–2006:4[a]



Source: Authors' calculations.

a. The sample period starts in 1974:2 for New Zealand and 1979:2 for Norway. The window size for the rolling estimations is seventy-four quarters. For instance, the first point estimate corresponds to 1988:3, which is based on the period 1970:2–1988:3.

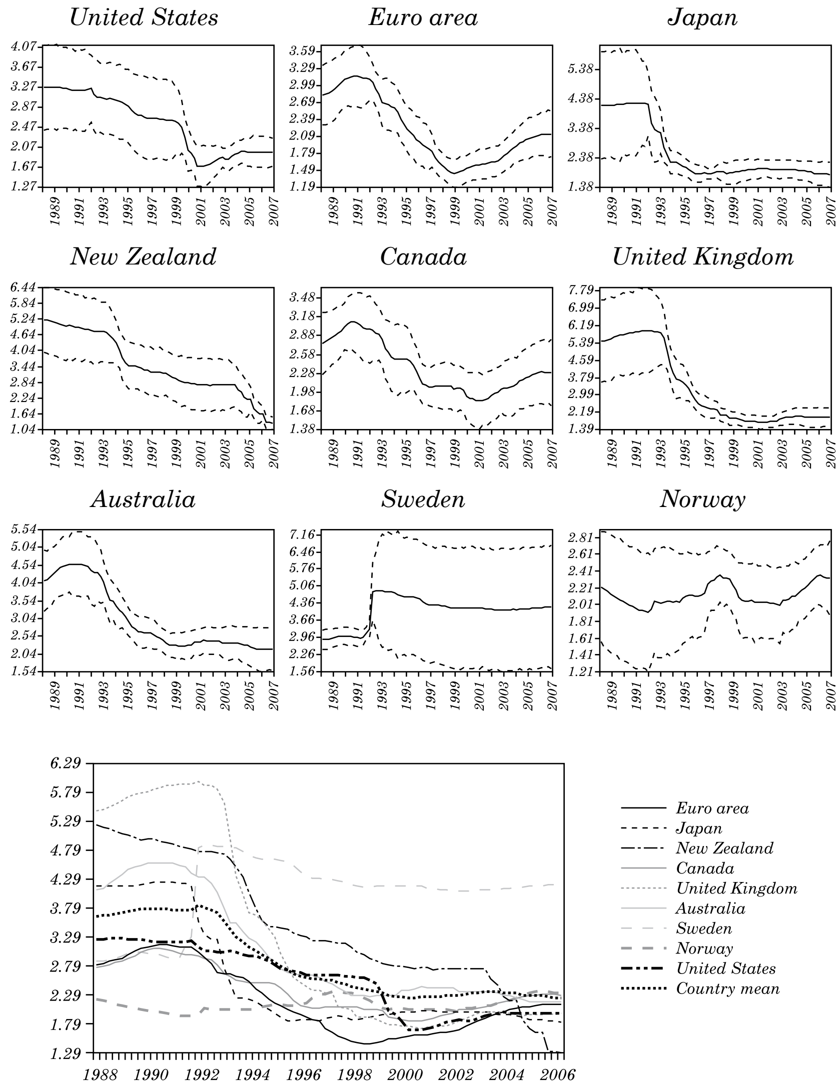
Figure 18. Output Gap Volatility Trends in Nine Economies, 1970:2–2006:4[a]



Source: Authors' calculations.

a. The sample period starts in 1974:2 for New Zealand and 1979:2 for Norway. The window size for the rolling estimations is seventy-four quarters. For instance, the first point estimate corresponds to 1988:3, which is based on the period 1970:2–1988:3.

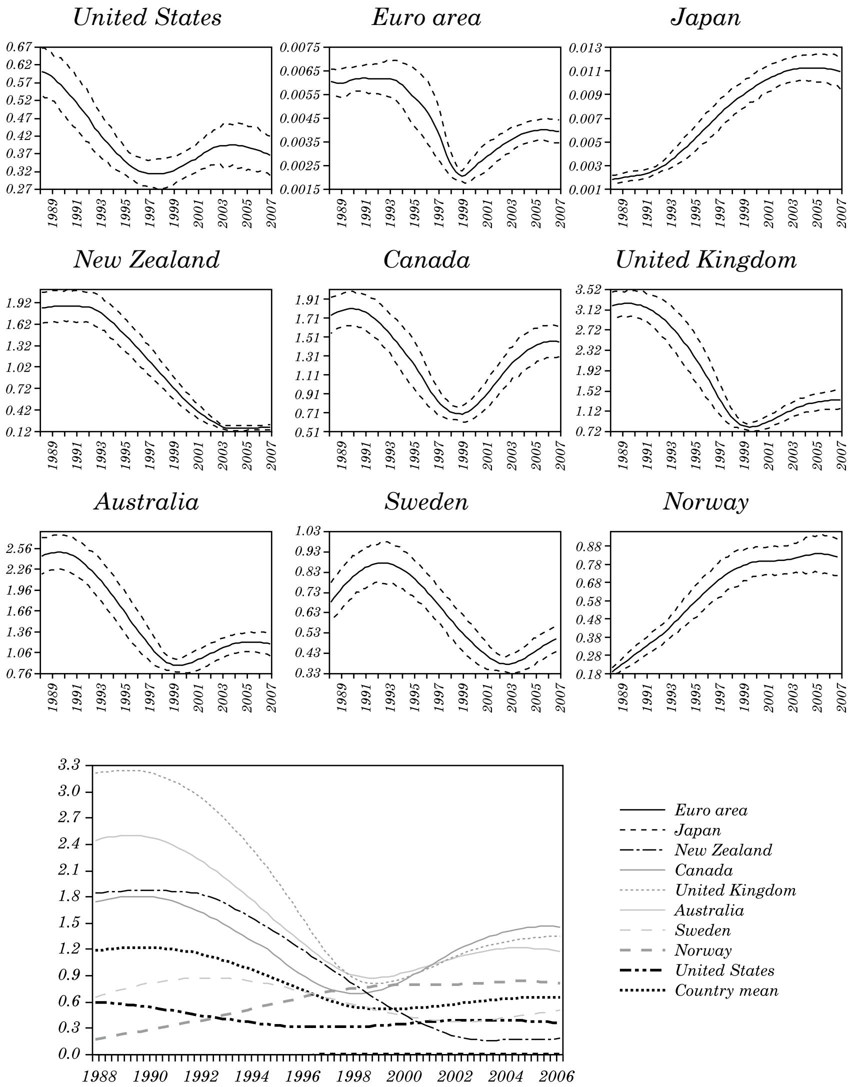
Figure 19. Actual Interest Rate Volatility Trends in Nine Economies, 1970:2–2006:4[a]



Source: Authors' calculations.

a. The sample period starts in 1974:2 for New Zealand and 1979:2 for Norway. The window size for the rolling estimations is seventy-four quarters. For instance, the first point estimate corresponds to 1988:3, which is based on the period 1970:2–1988:3.

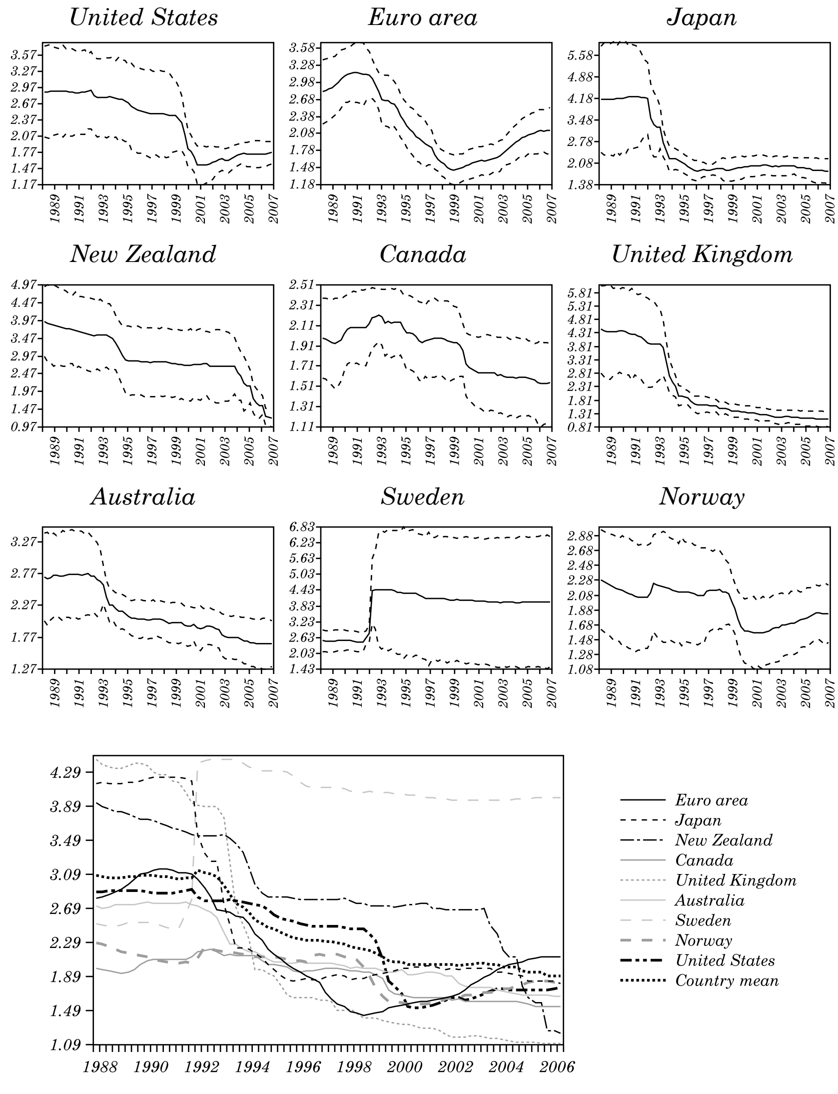
Figure 20. Neutral Interest Rate Volatility Trends in Nine Economies, 1970:2–2006:4[a]



Source: Authors' calculations.

a. The sample period starts in 1974:2 for New Zealand and 1979:2 for Norway. The window size for the rolling estimations is seventy-four quarters. For instance, the first point estimate corresponds to 1988:3, which is based on the period 1970:2–1988:3.

Figure 21. Interest Rate Gap Volatility Trends in Nine Economies, 1970:2–2006:4[a]



Source: Authors' calculations.

a. The sample period starts in 1974:2 for New Zealand and 1979:2 for Norway. The window size for the rolling estimations is seventy-four quarters. For instance, the first point estimate corresponds to 1988:3, which is based on the period 1970:2–1988:3.

United Kingdom exhibits the most stable output gap throughout the full 1970–2006 period.

A general pattern of declining volatility is also found for the real interest rate: the average country volatility falls from 3.8 percent to 2.3 percent (figure 19). The largest reductions in interest rate volatility are recorded in New Zealand and the United Kingdom. Norway does not exhibit a trend reduction because its interest rate volatility is low from the start. The exception is Sweden, which experienced a sharp rise in interest rate volatility in the third quarter of 1992, as a result of a very short but very high interest rate spike.

As with potential output growth, the results for the volatility of our estimated neutral real interest rate are mixed (figure 20). Nevertheless, the average country volatility of the neutral rate declines by half, from 1.2 percent in 1970–87 to 0.6 percent in 1988–2006. The largest decline in neutral rate volatility is recorded by the United Kingdom, while volatility rises in Norway. Japan records the lowest neutral rate volatility, close to zero, throughout the full sample period.

Finally, the results for the interest rate gap largely mimic those of the real interest rate because the natural interest rate exhibits very low variability relative to the real rate (figure 21).

The evidence presented here is strongly supportive of the Great Moderation in key macroeconomic variables in industrial countries. The strong trend reduction in volatilities of three observed variables (namely, inflation, output growth, and the real interest rate) and the moderate decline in volatilities of the unobservable neutral interest rate and the two unobservable gap measures (the output gap and the interest rate gap), as well as the narrowing of the corresponding confidence intervals, are proof of the gains attained in macroeconomic stability during the period from 1988 to 2006. The narrowing of country differences in volatilities that came about with the reduction in country volatilities during the last four decades also suggests stronger comovements across countries, which is our next topic.

5.3 Comovements

This section focuses on comovements of key variables across countries. We look at the same variables as above, less inflation. Cross-country correlations are reported for each variable for the full sample period (the 1970s to 2006) in tables 7 and 8. We focus on pairwise regional patterns. Output growth correlations among the three largest economies are low but significantly different from

Table 7. Cross-Country Correlations of Key Output Variables in Nine Industrial Economies, 1970:2–2006:4[a]

Variable and country	*United States*	*Euro area*	*Japan*	*New Zealand*	*Canada*	*United Kingdom*	*Australia*	*Sweden*	*Norway*
A. Actual output growth									
United States	1.00	**0.24**	**0.19**	**0.24**	**0.50**	**0.29**	**0.30**	0.18	**0.15**
Euro area		1.00	**0.31**	**0.25**	**0.32**	**0.37**	0.10	**0.32**	**0.30**
Japan			1.00	-0.07	0.13	**0.28**	-0.02	-0.01	-0.08
New Zealand				1.00	**0.20**	0.08	0.12	**0.24**	**0.26**
Canada					1.00	**0.27**	**0.31**	0.11	0.08
United Kingdom						1.00	0.05	**0.28**	-0.01
Australia							1.00	0.08	0.01
Sweden								1.00	**0.22**
Norway									1.00
B. Potential output growth[b]									
United States	1.00	0.00	**0.82**	**-0.61**	**0.55**	**-0.64**	0.00	**-0.73**	**0.66**
Euro area		1.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00
Japan			1.00	**-0.83**	**0.58**	**-0.90**	0.00	**-0.75**	0.27
New Zealand				1.00	-0.56	**0.85**	0.00	**0.70**	-0.30
Canada					1.00	**-0.34**	0.00	-0.05	0.16
United Kingdom						1.00	0.00	**0.80**	-0.31
Australia							1.00	0.00	0.00
Sweden								1.00	-0.34
Norway									1.00

Table 7. (continued)

Variable and country	*United States*	*Euro area*	*Japan*	*New Zealand*	*Canada*	*United Kingdom*	*Australia*	*Sweden*	*Norway*
C. Output gap									
United States	1.00	**-0.28**	**0.27**	0.29	**0.66**	**0.50**	**0.53**	**0.47**	**0.32**
Euro area		1.00	**-0.77**	**-0.75**	-0.20	0.04	**-0.56**	**-0.60**	**-0.57**
Japan			1.00	**0.48**	0.11	0.02	**0.48**	**0.42**	0.21
New Zealand				1.00	0.26	-0.12	**0.42**	**0.44**	**0.80**
Canada					1.00	**0.38**	**0.66**	**0.52**	**0.34**
United Kingdom						1.00	**0.40**	**0.34**	0.04
Australia							1.00	**0.65**	**0.27**
Sweden								1.00	**0.38**
Norway									1.00

Source: Authors' calculations.
a. The sample period is 1974:2–2006:4 for New Zealand and 1979:2–2006:4 for Norway. Figures in bold indicate significant correlation coefficients based on Hall's confidence intervals calculated using the stationary bootstrap technique, while underlined figures indicate significant correlation coefficients based on the *t* distribution.
b. The potential output growth estimate is constant for the euro area and Australia.

Table 8. Cross-Country Correlations of Key Interest Rate Variables in Nine Industrial Economies, 1970:2–2006:4[a]

	United States	*Euro area*	*Japan*	*New Zealand*	*Canada*	*United Kingdom*	*Australia*	*Sweden*	*Norway*
A. Actual interest rate									
United States	1.00	**0.49**	0.26	0.24	**0.71**	**0.39**	**0.52**	0.22	0.13
Euro area		1.00	**0.52**	**0.48**	**0.63**	**0.55**	**0.69**	**0.65**	**0.60**
Japan			1.00	0.06	0.22	-0.09	0.18	0.25	**0.62**
New Zealand				1.00	**0.27**	**0.61**	**0.57**	0.29	0.16
Canada					1.00	**0.53**	**0.60**	0.32	**0.39**
United Kingdom						1.00	**0.66**	**0.37**	**0.40**
Australia							1.00	**0.38**	**0.30**
Sweden								1.00	**0.33**
Norway									1.00
B. Natural interest rate									
United States	1.00	**0.64**	**0.37**	0.17	**0.90**	**0.76**	**0.78**	**0.48**	**0.91**
Euro area		1.00	0.14	**0.60**	**0.83**	**0.74**	**0.73**	**0.68**	**0.60**
Japan			1.00	**-0.63**	0.21	-0.20	-0.15	**-0.38**	**0.99**
New Zealand				1.00	**0.45**	**0.78**	**0.74**	**0.82**	**-0.63**
Canada					1.00	**0.90**	**0.91**	**0.77**	**0.77**
United Kingdom						1.00	**0.99**	**0.89**	**0.57**
Australia							1.00	**0.89**	**0.54**
Sweden								1.00	-0.05
Norway									1.00

Table 8. (continued)

	United States	*Euro area*	*Japan*	*New Zealand*	*Canada*	*United Kingdom*	*Australia*	*Sweden*	*Norway*
C. Interest rate gap									
United States	1.00	**0.41**	0.21	0.18	**0.58**	0.16	**0.39**	0.17	-0.24
Euro area		1.00	**0.52**	**0.31**	0.13	0.04	**0.26**	**0.54**	**0.42**
Japan			1.00	0.07	-0.04	-0.39	-0.05	**0.23**	**0.39**
New Zealand				1.00	-0.10	**0.24**	**0.23**	0.11	0.14
Canada					1.00	**0.22**	**0.27**	0.01	-0.11
United Kingdom						1.00	**0.34**	0.01	0.09
Australia							1.00	0.06	-0.09
Sweden								1.00	**0.27**
Norway									1.00

Source: Authors' calculations.
a. The sample period is 1974:2–2006:4 for New Zealand and 1979:2–2006:4 for Norway. Figures in bold indicate significant correlation coefficients based on Hall's confidence intervals calculated using the stationary bootstrap technique, while underlined figures indicate significant correlation coefficients based on the *t* distribution.

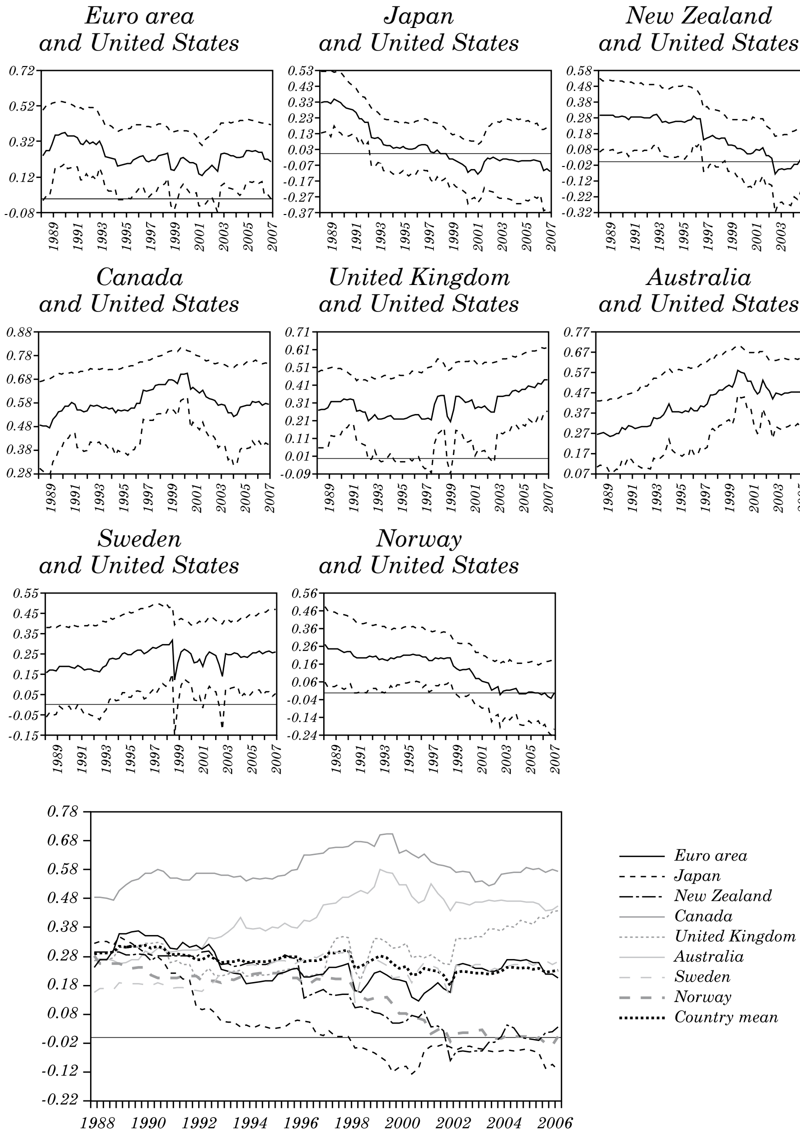
zero. The correlations between the three larger economies and some smaller countries (Canada and European economies) are somewhat larger. Our estimates for potential output growth in the euro area and Australia are constant, so we focus on correlations of third countries with the United States. Canada, Japan, and Norway display large positive correlations with the United States, whereas we find large negative correlations with the United States in New Zealand, Sweden, and the United Kingdom.

Output gap correlations between the euro area and every included country are either largely negative or zero, reflecting highly nonsynchronous business-cycle conditions of the euro area with other industrial countries. This stands in stark contrast to the United States, whose output gap is highly and positively correlated with all economies, except the euro area.

Among the three big economies, real interest rates are positively correlated. The same is true for most pairwise correlations, except Japan's. This reflects the common, long cycle of low-high-low real interest rates observed in most countries during the last four decades. Even stronger correlations are observed in the case of neutral real interest rates, again except Japan, reflecting the common world trend in monetary policy observed in most industrial countries. Cross-country interest rate gap correlations are similar to actual interest rate correlations, but they are often smaller and less significant.

To describe cross-country comovements, we follow the approach adopted above in documenting volatility trends. Here we focus on rolling correlations of key variables between the United States and the eight industrial economies. We report point estimates of correlation coefficients and their confidence intervals for seventy-four-quarter windows during 1970–2006, using the stationary bootstrap technique mentioned above. We find no common trend in output growth correlations with the United States (figure 22). While output growth correlations with the United States rise in Australia, Canada, Sweden, and the United Kingdom, they decline in Japan, New Zealand, and Norway. Potential output growth correlations turn from positive (and mostly significant) to negative (and significant) in New Zealand, Canada, United Kingdom, and Sweden (figure 23). Except for the euro area and Japan, output gap correlations of all other countries with the United States rise over time, confirming increasing cyclical synchronization between small and medium-sized industrial economies and the U.S. economy (figure 24).

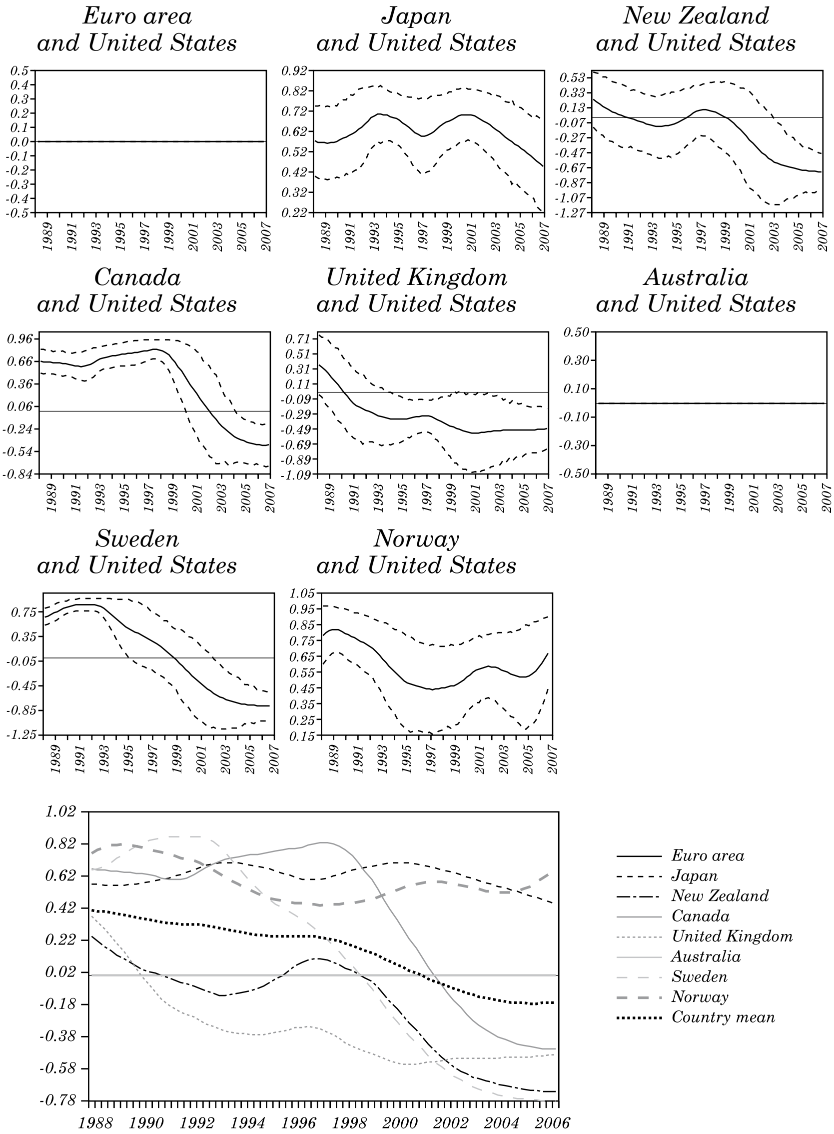
Figure 22. Actual Output Growth Correlation of Eight Industrial Economies with the United States, 1970:2–2006:4[a]



Source: Authors' calculations.

a. The sample period starts in 1974:2 for New Zealand and 1979:2 for Norway. The window size for the rolling estimations is seventy-four quarters. For instance, the first point estimate corresponds to 1988:3, which is based on the period 1970:2–1988:3.

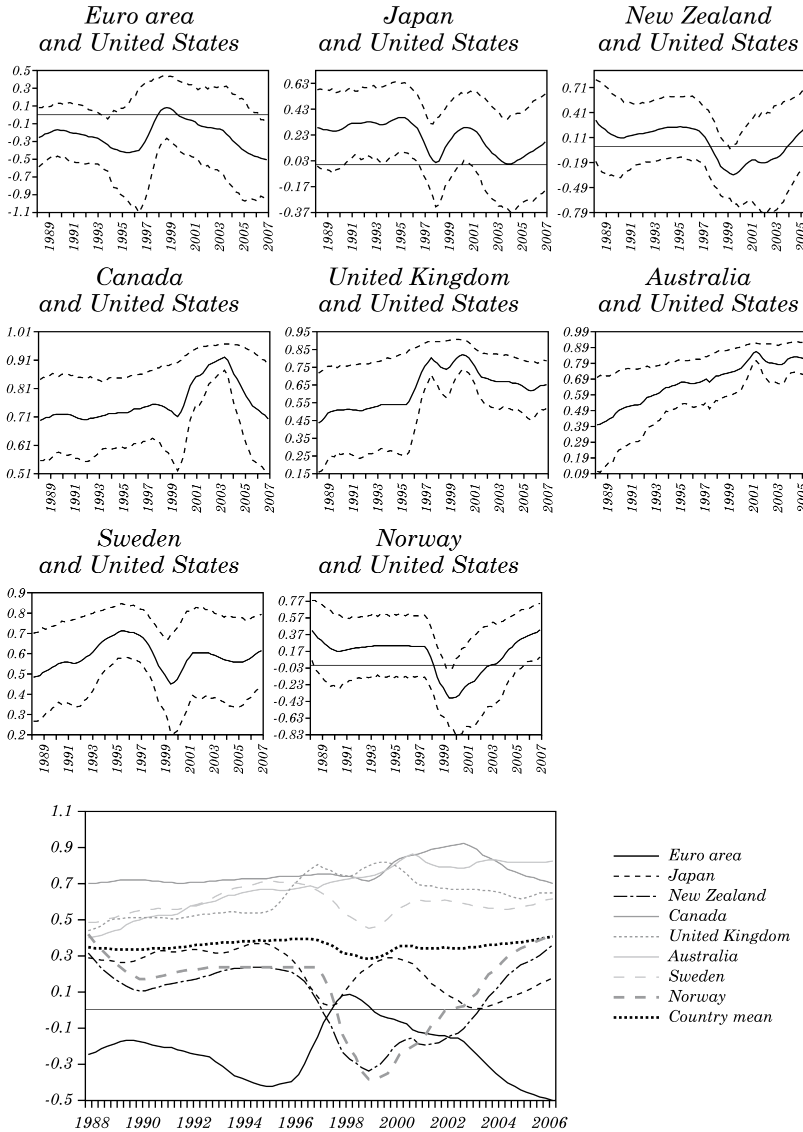
Figure 23. Potential Output Growth Correlations of Eight Industrial Economies with the United States, 1970:2–2006:4[a]



Source: Authors' calculations.

a. The sample period starts in 1974:2 for New Zealand and 1979:2 for Norway. The window size for the rolling estimations is seventy-four quarters. For instance, the first point estimate corresponds to 1988:3, which is based on the period 1970:2–1988:3.

Figure 24. Output Gap Correlations of Eight Industrial Economies with the United States, 1970:2–2006:4[a]



Source: Authors' calculations.

a. The sample period starts in 1974:2 for New Zealand and 1979:2 for Norway. The window size for the rolling estimations is seventy-four quarters. For instance, the first point estimate corresponds to 1988:3, which is based on the period 1970:2–1988:3.

Real interest rate correlations with the United States display a U-shaped pattern over the last four decades, reaching their lowest values during the 1980s and early 1990s and rising to high levels again in the late 1990s and early 2000s. This suggests rising monetary integration (or declining monetary independence) in the last decade (figure 25). Regarding neutral real interest rate correlations with the United States, the U-shaped pattern is confirmed in most economies, while in Japan and Norway correlations turn from negative to positive (figure 26). New Zealand displays the opposite pattern, from positive to negative. The country pattern of interest rate gap correlations with the United States replicates that of actual interest rate correlations, reflecting the smoothness of neutral rates (figure 27).

Summing up, country averages of the rolling correlation coefficients of country variables with those of the United States display slightly rising trends for the output gap, the actual interest rate, the neutral interest rate, and the interest rate gap (the lower panels in figures 22 through 27). The opposite is observed regarding average trends in actual and potential output growth with the United States, which decline over time.

5.4 Convergence

In this section, we test for cross-country convergence with the United States and the euro area in key variables for our full sample of eight countries. Because rising correlations over time do not imply convergence in levels, we carry out this final set of exercises on convergence to complement the previous evidence on increasing comovements.

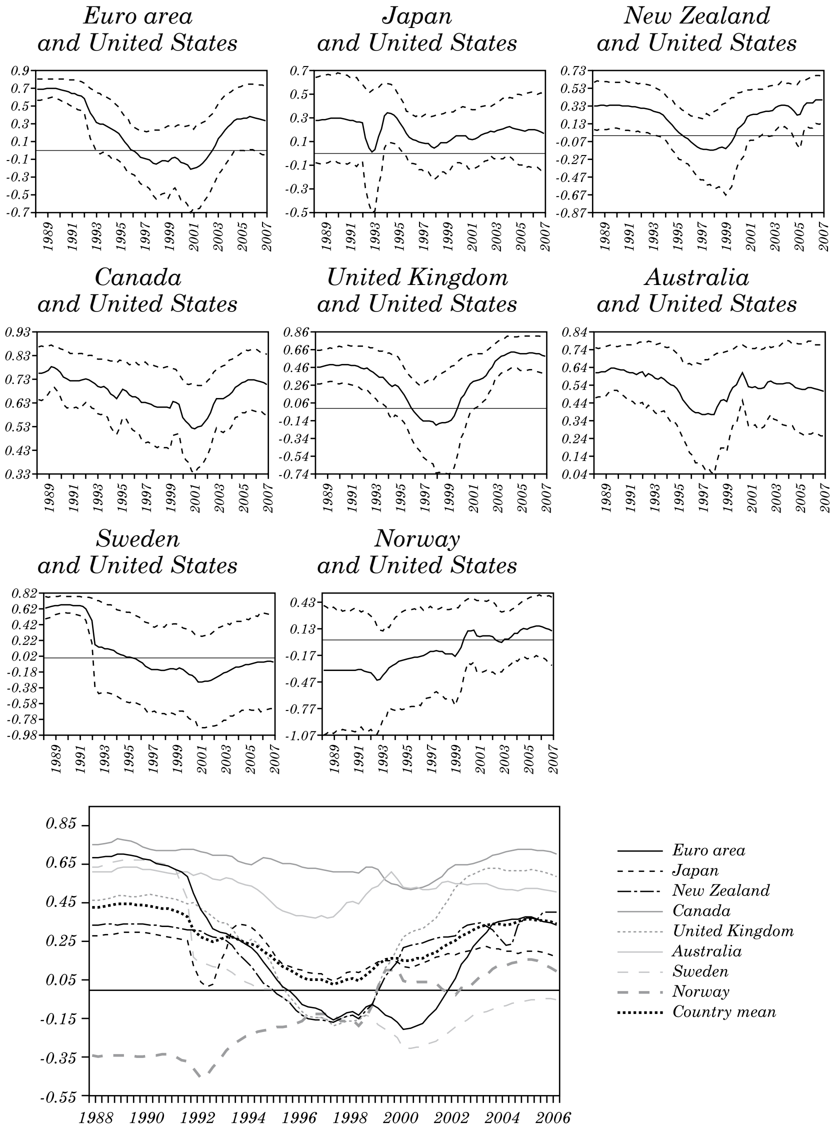
We test for convergence across countries using the following simple autoregressive models for the difference in country *j*'s variable v with respect to that of the United States or the euro area:

$$v_{j,t} - v_{us,t} = \alpha_0 + \sum_{i=1}^{p} \alpha_i \left(v_{j,t-i} - v_{us,t-i}\right) + u_{j,us,t} \tag{18}$$

or

$$v_{j,t} - v_{euroarea,t} = \alpha_0 + \sum_{i=1}^{p} \alpha_i \left(v_{j,t-i} - v_{euroarea,t-i}\right) + u_{j,euroarea,t},$$

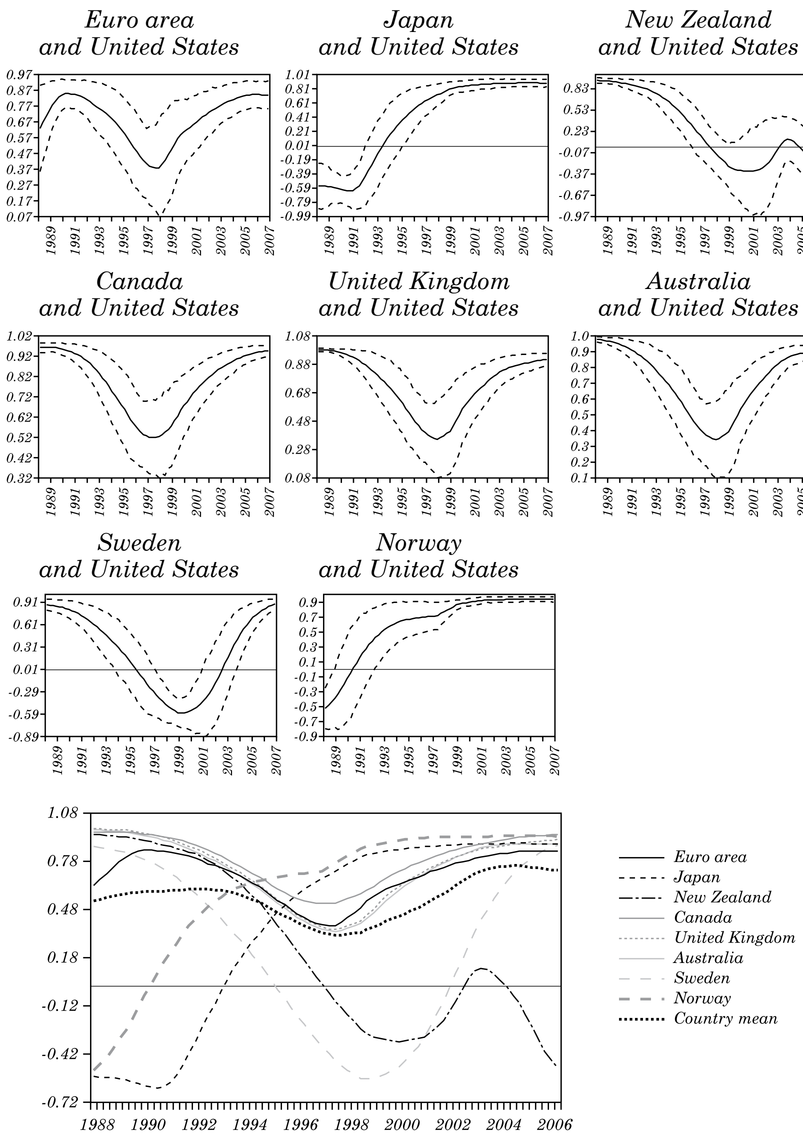
Figure 25. Actual Interest Rate Correlations of Eight Industrial Economies with the United States, 1970:2–2006:4[a]



Source: Authors' calculations.

a. The sample period starts in 1974:2 for New Zealand and 1979:2 for Norway. The window size for the rolling estimations is seventy-four quarters. For instance, the first point estimate corresponds to 1988:3, which is based on the period 1970:2–1988:3.

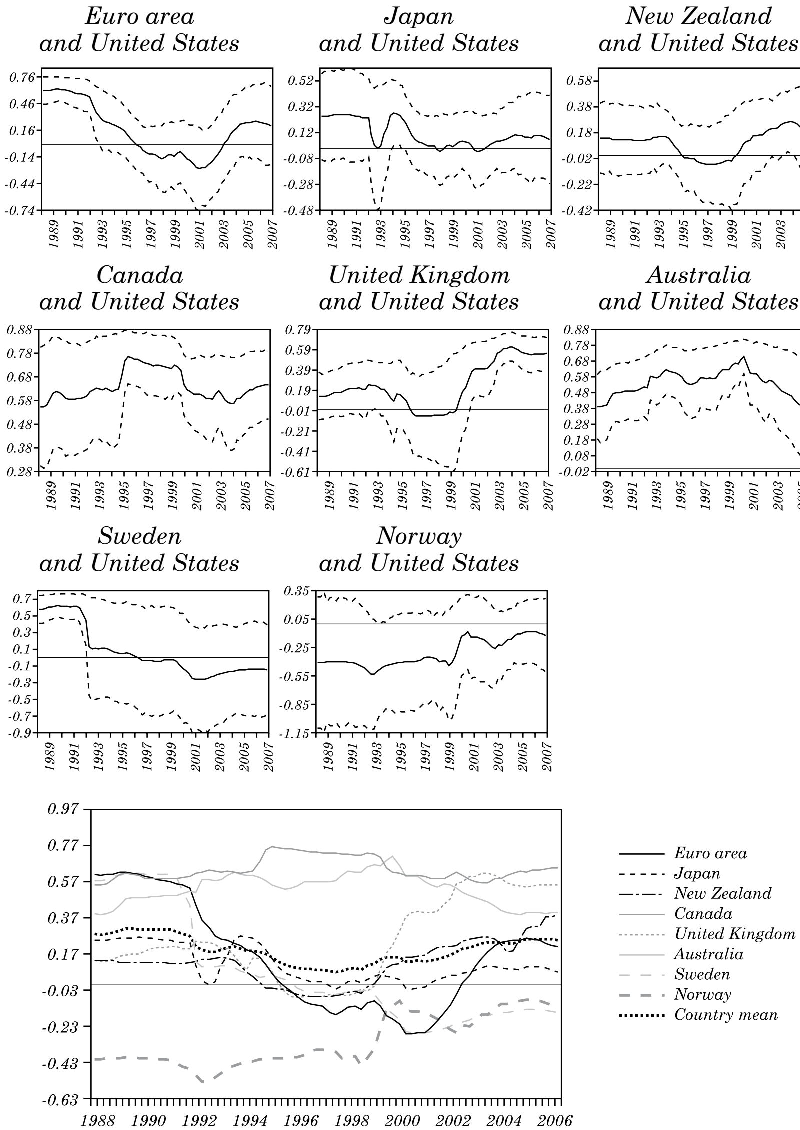
Figure 26. Neutral Interest Rate Correlations of Eight Industrial Economies with the United States, 1970:2–2006:4[a]



Source: Authors' calculations.

a. The sample period starts in 1974:2 for New Zealand and 1979:2 for Norway. The window size for the rolling estimations is seventy-four quarters. For instance, the first point estimate corresponds to 1988:3, which is based on the period 1970:2–1988:3.

Figure 27. Interest Rate Gap Correlations of Eight Industrial Economies with the United States, 1970:2–2006:4[a]



Source: Authors' calculations.

a. The sample period starts in 1974:2 for New Zealand and 1979:2 for Norway. The window size for the rolling estimations is seventy-four quarters. For instance, the first point estimate corresponds to 1988:3, which is based on the period 1970:2–1988:3.

where v_j (v_{us}, $v_{euroaera}$) is an observable variable or an unobservable estimate for country *j* (for the United States, for the euro area), u_j (u_{us}, $u_{euroarea}$) is a zero-mean stochastic error term for country *j* (for the United States, for the euro area), and α_0 and α_i ($i = 1, \ldots p$) are the autoregressive coefficients of the AR(p) process.

For the AR(p) model, we obtain convergence across countries if the AR polynomial is stationary.[22] To test for stationarity, we use a grid bootstrap method to estimate confidence intervals for the parameters of interest (Hansen, 1999).[23]

The variable v represents observable variables (output growth and the interest rate), our estimates for unobservables (potential output growth and the neutral interest rate), and our estimated unobservable gaps (the output gap and the interest rate gap). We do not test for convergence in levels of cross-country gap measures, however, as they tend to zero by construction.

The convergence tests for actual output growth (table 9) and interest rates (table 10) reveal the following results. For actual growth convergence with the United States, we find that all countries are characterized by an AR(1) model, except Sweden with an AR(2) process. We find (weak) evidence of convergence with the United States for all countries, although α_j is only significant in Chile, New Zealand, Norway, and Sweden. For the remaining countries we are not able to reject a white-noise process.[24] For all countries, we obtain small half-lives of shocks (HLS) of only 0.6 quarters, on average.

When we examine actual growth convergence with the euro area, the relationships are characterized by higher-order AR processes in Japan, Norway, Sweden, and the United Kingdom. We find evidence of convergence with the euro area for all countries. The smallest HLS is 0.2 quarters (Australia) and the largest is 2.3 quarters (United Kingdom); the average HLS is 1.1 quarters.

22. For example, convergence of an AR(1) model requires that $|\alpha_1| < 1$; convergence of an AR(2) model requires that $\alpha_1 + \alpha_2 < 1$, $\alpha_2 - \alpha_1 < 1$, and $\alpha_2 > -1$. Hamilton (1994) provides a more detailed discussion of stationarity conditions.

23. The bootstrap method works as follows. Pick a grid over the parameters of interest and calculate the confidence interval by bootstrap at each parameter value, then smooth the estimated function for the confidence interval using a kernel regression, and finally obtain the confidence interval estimated by the kernel for a given value of the parameter. Lag lengths (p lags) are determined using the Akaike information criterion (AIC), the Hannan-Quinn criterion (HQC), and the Bayesian information criterion (BIC).

24. All autocorrelations and partial correlations are not significantly different from zero.

Table 9. Convergence of Actual Output Growth in Eight Countries with the United States and the Euro Area, 1970:2–2006:4[a]

Country	*I(0)* *(1)*	*Order* *(2)*	*Drift* *(3)*	*AR coefficients* *(4)*		*HLS* *(5)*
A. Convergence with the United States						
Euro area	Yes	1	0.0000	0.1260		0.3
Japan	Yes	1	0.0000	0.1105		0.3
New Zealand	Yes	1	0.0000	**−0.1836**		0.6
Canada	Yes	1	0.0000	−0.0273		0.5
United Kingdom	Yes	1	**−0.6998**	−0.1092		0.3
Australia	Yes	1	0.0000	−0.0684		0.5
Sweden	Yes	2	0.0000	−0.0162	**0.1742**	1.6
Norway	Yes	1	0.0000	**−0.3124**		0.7
Chile	Yes	1	**2.6974**	**0.2233**		0.5
				Average HLS		0.6

Table 9. (continued)

B. Convergence with the euro area								
Japan	Yes	4	0.0000	0.0797	0.1253	**0.2402**	**-0.1871**	1.7
New Zealand	Yes	1	0.0000	**−0.2010**				0.6
Canada	Yes	1	**0.6894**	**0.2240**				0.5
United Kingdom	Yes	3	0.0000	−0.0377	**0.1720**	**0.1801**		2.3
Australia	Yes	1	0.0000	0.0244				0.2
Sweden	Yes	3	0.0000	**−0.1829**	**0.2448**	**0.1837**		1.6
Norway	Yes	3	0.0000	**−0.2135**	**0.3031**	0.1768		1.1
Chile	Yes	1	**3.3966**	**0.2991**				0.6
				Average HLS				1.1

Source: Authors' calculations.

a. The sample period for Chile is 1986–2006. Significant estimates are in bold. In column 1, we use the grid bootstrap (Hansen, 1999) for autoregressive models to compute confidence intervals for all AR coefficients. In column 2, we use AIC, BIC and HQC criteria to determine lag lengths. Column 3 reports the value of the constant in the AR model. Column 4 presents the estimated AR coefficients. Column 5 reports the half-life of a unit shock (HLS) coefficient, which is defined as HLS = abs(log(1/2)/log(α)) for AR(1) model (with $\alpha \geq 0$). The HLS for AR(p) models can be calculated directly from the impulse response functions. We did not find convergence for the unobservables (natural rate of interest and potential output growth) in either case (with the United States or the euro area), since the series are not I(0) (stationary). In these cases, the HLS coefficients are explosive (∞ or a large number).

Table 10. Convergence of the Actual Interest Rate in Eight Countries with the United States and the Euro Area, 1970:2–2006:4[a]

Country	I(0) (1)	Order (2)	Drift (3)	AR coefficients (4)		HLS (5)
A. Convergence with the United States						
Euro area	Yes	1	0.0000	**0.8650**		4.8
Japan	Yes	1	0.0000	**0.8274**		3.7
New Zealand	Yes	1	0.0000	**0.7494**		2.4
Canada	Yes	1	0.0000	**0.7571**		2.5
United Kingdom	Yes	1	0.0000	**0.7625**		2.6
Australia	Yes	1	0.0000	**0.7107**		2.0
Sweden	Yes	1	0.0000	**0.6806**		1.8
Norway	Yes	1	0.0000	**0.8826**		5.6
Chile	Yes	2	0.0000	**0.7066**	**0.2182**	7.5
				Average HLS		3.6
B. Convergence with the euro area						
Japan	Yes	1	0.0000	**0.7554**		2.7
New Zealand	Yes	1	0.0000	**0.7060**		2.0
Canada	Yes	2	0.0000	**1.0074**	**-0.2645**	3.0
United Kingdom	Yes	1	0.0000	**0.6695**		1.7
Australia	Yes	1	0.0000	**0.5953**		1.3
Sweden	Yes	1	0.0000	**0.4365**		0.8
Norway	Yes	1	0.0000	**0.8115**		3.3
Chile	Yes	1	0.0000	**0.8813**		5.5
				Average HLS		2.6

Source: Authors' calculations.

a. The sample period for Chile is 1986–2006. Significant estimates are in bold. In column 1, we use the grid bootstrap (Hansen, 1999) for autoregressive models to compute confidence intervals for all AR coefficients. In column 2, we use AIC, BIC and HQC criteria to determine lag lengths. Column 3 reports the value of the constant in the AR model. Column 4 presents the estimated AR coefficients. Column 5 reports the half-life of a unit shock (HLS) coefficient, which is defined as HLS = abs(log(1/2)/log(α)) for AR(1) model (with $\alpha \geq 0$). The HLS for AR(p) models can be calculated directly from the impulse response functions. We did not find convergence for the unobservables (natural rate of interest and potential output growth) in either case (with the United States or the euro area), since the series are not I(0) (stationary). In these cases, the HLS coefficients are explosive (∞ or a large number).

Turning to convergence of actual interest rates with U.S. interest rates, we estimate an AR(1) process for almost all countries, except Chile with an AR(2) process (table 10). We find that all countries converge to the United States (and all estimated parameters are significant). As above, we also estimate HLS coefficients, which are much larger than those obtained for growth convergence. HLS coefficients range from 1.8 quarters (Sweden) to 7.5 quarters (Chile), with an average HLS of 3.7 quarters.

For interest rate convergence with the euro area, we estimate an AR(1) process for all countries, less Canada with an AR(2). All countries' interest rates converge to the euro area's. Our HLS estimates range from 0.8 quarters (Sweden) to 5.5 quarters (Chile), with an average HLS of 2.6 quarters.

We did not find country convergence of our two key estimated unobservables (that is, the potential output growth rate and the neutral real interest rate) with either the United States or the euro area. This reflects the fact that country differentials in unobservables—with either the United States or the euro area—are not stationary in the 1970–2006 sample.

6. Conclusions and Possible Extensions

The conduct of monetary policy is crucially dependent on several key unobservables. The output gap, the neutral real interest rate, the natural rate of unemployment, and expected inflation are the most critical for central bank models, forecasts, and policy decisions. Individual central banks have developed methodologies for estimating unobservable variables. Many researchers have derived estimates for single countries (usually the United States) or for a small number of developed economies. We have extended this literature by providing new estimates of key unobservables for ten economies, including the world's three largest economies and seven inflation-targeting countries. In addition, we have exploited our time-series estimates of unobservables for ten economies to test for common trends, more macroeconomic stability, comovements, and convergence across economies.

We adopted a very parsimonious monetary policy model comprising an IS relation, a Phillips curve, a Taylor rule, and transition equations for key observables and unobservables. This model was applied to all sample countries. An extended version, including Okun's law, was also applied to the United States and Chile. Our estimation model, which closely follows Laubach and Williams' (2003) sequential-step estimation procedure, yields country estimates for model parameters and unobservable-variable time series for each country.

Structural parameter estimates vary from country to country and sometimes from sample to sample for a given country. The results conform to our prior assumptions in the case of the United States, Canada, and Chile, less so for six other economies, and the least for Japan.

We also obtain reasonable and precise estimates for unobservable variables and for all countries. The evidence points to time variation in trend output growth, the neutral real interest rate, and (for the United States and Chile) the natural rate of unemployment. This time variation has important implications for the conduct of monetary policy. For example, if trend growth of potential output were constant, then policy rules that focus on the growth rate of output relative to the growth rate of potential (such as speed limit policies of the type analyzed in Walsh, 2003) might serve to eliminate (or at least significantly reduce) measurement problems in estimating the level of potential output. But if the growth rate of potential output is also subject to stochastic variation, as we find it to be, then the problem of estimating the level of potential cannot be eliminated by simply focusing on growth rates. Similarly, time variation in the neutral real interest rate implies that simple Taylor rules for the policy interest rate, which very commonly assume that the equilibrium real interest rate is constant, may lead to policy errors.

Finally we have used our estimates of unobservables and the data on observables to test for common trends and comovements across countries, the time trend toward more macroeconomic stability, and convergence in variable levels toward those observed in the United States and the euro area.

Consistent with the notion of a Great Moderation over the 1988–2006 period, measures of inflation volatility showed a marked common decline over the past decade. Output growth also declined in volatility. However, little of this decline in output growth volatility seems due to a decline in the volatility of the growth rate of potential output. The volatility of the latter has fallen slightly over the past twenty years, but this decline is small relative to the overall reduction in output growth volatility. Given these results, it is perhaps surprising that the volatility of the output gap displays only a modest decline over the sample. This reflects, in part, a rise in the average output gap volatility among our sample countries over the past decade. This is an interesting finding since it offers evidence, consistent with standard theoretical models, that greater inflation stability should come at the cost of some increase in output gap volatility. The failure of output gap volatility to fully reflect the decline in output growth volatility suggests that there may have been an increase in the volatility of the *level* of potential output over this period.

We find evidence that the volatility of the neutral real interest rate has declined when we look at the average across the sample economies. However, this masks significant differences among the individual economies.

Interestingly, we find neutral real interest rates to be more highly correlated across countries than either actual real rates or Wicksellian interest rate gaps. The notable exception to this finding is Japan. While neutral real rates were highly correlated across countries, this did not reflect a common pattern of convergence to the level of the U.S. or euro area neutral real rates. In fact, the neutral real rate differentials were nonstationary, indicating no long-run tendency to converge.

There are several extensions of the analysis that would be interesting to pursue. We would like to extend the approach to allow for richer and potentially different dynamics across the set of countries. Undoubtedly, one reason for some of our mixed results for individual countries arises from our use of a common specification of dynamics across all countries, particularly since our parsimonious model incorporated a fairly simple dynamic structure. It would also be useful to extend the sample to include more emerging market and developing economies. Many of these economies have adopted inflation-targeting frameworks in which the output gap and the neutral real interest rate are central to the design of policy. They are generally small open economies, making them candidates for exploring issues of convergence and comovements among these countries and the large industrialized economies.

Data Appendix[a]

Variable	*Description*	*Source*	*Countries*
Inflation measure	Consumer price index	IFS	Australia, Canada, Japan, New Zealand, Norway, Sweden, and United Kingdom
	Consumer price index	INE and CBC	Chile
	Price index for personal consumption expenditures	LW	United States
	Consumption deflactor	ECB	Euro area
Inflation targets	A composite measure that joins the HP-filtered inflation rate and the observed inflation targets for inflation targeters. For nontargeters (the euro area, Japan, and the United States) we use the HP-filtered series for the inflation measure.	Authors' construction	All countries
Inflation expectation	Calculations based on four step-ahead forecasts stemming from an AR(4) for the actual inflation rate.	Authors' construction	All countries
Gross domestic product	Seasonally adjusted real gross domestic product	OECD	Australia, Canada, Japan, New Zealand, Norway, Sweden, United Kingdom, and United States
		ECB	Euro area
		CBC	Chile
		LW	United States
Unemployment rate	Seasonally adjusted unemployment rate	OECD	Australia, Canada, Japan, New Zealand, Norway, Sweden, United Kingdom, and United States
		ECB	Euro area
		INE	Chile
Interest rate	Short-term interest rate. The real interest rate is calculated as the difference of the nominal interest rate and our estimation of the inflation expectations.	OECD	Australia, Canada, Japan, New Zealand, Norway, Sweden, and United Kingdom
		ECB	Euro area
	Real monetary policy rate. Previous to 1994 indexed CBC's 90-day bond rate. Since 2001, official nominal MPR less expected inflation from inflation reports.	CBC	Chile
	Monetary policy rate	LW	United States

a. IFS: International Financial Statistics; INE: National Statistics Institute of Chile; CBC: Central Bank of Chile; LW: Laubach and Williams (2003); ECB: European Central Bank; OECD: Organization for Economic Cooperation and Development.

References

Backus, D.K., P.J. Kehoe and F.E. Kydland. 1992. "International Real Business Cycles." *Journal of Political Economy* 100(4): 745−75.

Basdevant, O., N. Björksten and Ö. Karagedikli. 2004. "Estimating a Time Varying Neutral Real Interest Rate for New Zealand." Discussion paper 2004/01. Wellington: Reserve Bank of New Zealand.

Basistha, A. and C.R. Nelson. 2007. "New Measures of the Output Gap Based on the Forward-Looking New Keynesian Phillips Curve." *Journal of Monetary Economics* 54(2): 498–511.

Benati, L. and G. Vitale. 2007. "Joint Estimation of the Natural Rate of Interest, the Natural Rate of Unemployment, Expected Inflation, and Potential Output." Working paper 797. Frankfurt: European Central Bank.

Björksten, N, and Ö. Karagedikli. 2003. "Neutral Real Interest Rates Revisited." *Reserve Bank of New Zealand Bulletin* 66(3): 18−27.

Blanchard, O. and J. Galí. 2007. "Real Wage Rigidities and the New Keynesian Model." *Journal of Money, Credit, and Banking* 39(1): 35–65.

Clarida, R., J. Galí and M. Gertler. 1999. "The Science of Monetary Policy: A New Keynesian Perspective." *Journal of Economic Literature* 37(4): 1661–707.

Erceg, C. and A. Levin. 2003. "Imperfect Credibility and Inflation Persistence." *Journal of Monetary Economics* 50(4): 915−44.

Fuentes, R. and F. Gredig. 2008. "La tasa de interés neutral: estimaciones para Chile." *Economía Chilena* 11(2): 47–58.

Fuentes, R., F. Gredig and M. Larraín. 2008. "La brecha de producto en Chile: medición y evaluación." *Economía Chilena* 11(2): 7−30.

Garnier, J. and B.R. Wilhelmsen. 2009. "The Natural Rate of Interest and the Output Gap in the Euro Area: A Joint Estimation." *Empirical Economics* 36(2): 297-319.

Hamilton, J.D. 1986. "A Standard Error for the Estimated State Vector of a State-Space Model." *Journal of Econometrics* 33(3): 387−97.

————. 1994, *Time Series Analysis*. Princeton University Press.

Hansen, B. 1999. "The Grid Bootstrap and the Autoregressive Model." *Review of Economics and Statistics* 81(4): 594−607.

King, T.B. and J. Morley. 2007. "In Search of the Natural Rate of Unemployment." *Journal of Monetary Economics* 54(2): 550−64.

Kuttner, K.N. 1994. "Estimating Potential Output as a Latent Variable." *Journal of Business and Economic Statistics* 12(3): 361−68.

Laubach, T. and J.C. Williams. 2003. "Measuring the Natural Rate of Interest." *Review of Economics and Statistics* 85(4): 1063−70.

Levin, A., A. Onatski, J.C. Williams and N. Williams. 2006. "Monetary Policy under Uncertainty in Micro-Founded Macroeconomic Models." In *NBER Macroeconomics Annual 2005*, edited by M. Gertler and K. Rogoff, pp. 229–87. Cambridge, Mass.: MIT Press.

McCallum, B.T. 2001. "Should Monetary Policy Respond Strongly to Output Gaps?" *American Economic Review Papers and Proceedings* 91(2): 258−62.

Murchison, S. and A. Rennison. 2006. "ToTEM: The Bank of Canada's New Quarterly Projections Model." Technical report 97. Ottawa: Bank of Canada.

Orphanides, A. 2002. "Monetary Policy Rules and the Great Inflation." *American Economic Review* 92(2): 115–20.

————. 2003. "The Quest for Prosperity without Inflation." *Journal of Monetary Economics* 50(3): 633−63.

Orphanides, A. and S. Van Norden. 2002. "The Unreliability of Output-Gap Estimates in Real Time." *Review of Economics and Statistics* 84(4): 569−83.

Orphanides, A. and J.C. Willliams. 2002. "Robust Monetary Policy Rules with Unknown Natural Rates." *Brookings Papers on Economic Activity* 2002(2): 63–145.

Politis, D.N. and J.P. Romano. 1994. "The Stationary Bootstrap." *Journal of the American Statistical Association* 89(428): 1303–13.

Primiceri, G. E. 2006. "Why Inflation Rose and Fell: Policy-Makers' Beliefs and U.S. Postwar Stabilization Policy." *Quarterly Journal of Economics* 121(3): 867–901.

Reis, R. 2003. "Where Is the Natural Rate? Rational Policy Mistakes and Persistent Deviations of Inflation from Target." *Advances in Macroeconomics* 3(1).

Reserve Bank of New Zealand. 2004. *The Reserve Bank's Forecasting and Policy System*. Wellington: Reserve Bank of New Zealand, Economics Department.

Restrepo, J.E. 2008. "Estimaciones de la NAIRU para Chile." *Economía Chilena* 11(2): 31−46.

Rudebusch, G.D. and L.E.O. Svensson. 1999. "Policy Rules for Inflation Targeting." In *Monetary Policy Rules*, edited by J. Taylor, pp. 203–46, University of Chicago Press.

Sargent, T.J. 2008. "Evolution and Intelligent Design." *American Economic Review* 98(1): 5−37.

Stock, J.H. 1994. “Unit Root, Structural Breaks, and Trends,” In *Handbook of Econometrics*, vol. 5, edited by R.F. Engle and D.L. McFadden, chap. 6. Amsterdam: North-Holland.

Stock, J.H. and M. W. Watson. 1998. “Median Unbiased Estimation of Coefficient Variance in a Time-Varying Parameter Model.” *Journal of the American Statistical Association* 93(441): 349−58.

Svensson, L.E.O. and M. Woodford. 2005. “Implementing Optimal Policy through Inflation-Forecast Targeting.” In *The Inflation-Targeting Debate*, edited by B.S. Bernanke and M. Woodford, pp. 19–83. University of Chicago Press.

Walsh, C.E. 2003. “Speed Limit Policies: The Output Gap and Optimal Monetary Policy.” *American Economic Review* 93(1): 265−78.

Woodford, M. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Inflation Target Transparency and the Macroeconomy

Martin Melecký
World Bank

Diego Rodríguez Palenzuela
European Central Bank

Ulf Söderström
Sveriges Riksbank

Over the last twenty years, many central banks have adopted increasing standards of transparency in communicating their monetary policy objectives, in particular regarding the explicit definition and quantification of their price stability objective or inflation target. One important benefit of increased transparency is that it prepares the ground for central banks to improve their credibility and facilites the anchoring of private sector inflation expectations to stated objectives (see, for instance, Leiderman and Svensson, 1995; Bernanke and others, 1999). Economic theory suggests that private decisions are partly determined by agents' expectations concerning the future. Inflation targeting, by anchoring inflation expectations, can thus be expected to simplify private agents' decisions, thereby reducing macroeconomic volatility and increasing overall welfare.

Several authors present empirical evidence that inflation targeting coupled with central bank independence has had the effect of anchoring

At the time of the conference, Ulf Söderström was affiliated with Bocconi University.

This paper circulated previously under the title "Monetary Policy Credibility and the Macroeconomy." We are grateful for comments from Alex Bowen, Jim Bullard, Andy Levin, John McDermott, Luca Sala, Paul Söderlind, Antonella Trigari, John Williams, Tony Yates, and the editors, Klaus Schmidt-Hebbel and Carl E. Walsh. We also thank Eric Swanson for sharing his Matlab code. All views, conclusions, and opinions expressed in this paper reflect solely those of the authors and do not represent those of the World Bank or the European Central Bank.

Monetary Policy under Uncertainty and Learning, edited by Klaus Schmidt-Hebbel and Carl E. Walsh, Santiago, Chile.

inflation expectations. For instance, Levin, Natalucci, and Piger (2004) find that private sector inflation forecasts in the United States (where monetary policy is not guided by an inflation target) are highly correlated with a moving average of lagged inflation, while this correlation is essentially zero in a number of countries with formal inflation targets. Gürkaynak, Levin, and Swanson (2006) and Gürkaynak and others (2007) show that long-term inflation expectations tend to be less responsive to macroeconomic announcements in countries with independent inflation-targeting central banks, such as Canada, Sweden, or the United Kingdom after 1997, than in countries where the central bank is either not independent or does not have an explicit inflation target, such as the United States or the United Kingdom before formal independence in 1997.

There is no strong evidence, however, that this effect on inflation expectations has reduced macroeconomic volatility in general. While many economies, including the United Kingdom and Sweden, have performed well since the introduction of inflation targets, other economies without formal inflation targets, in particular the United States, have posted a similar, or even more impressive, performance.[1]

This paper aims at better understanding the links between monetary policy credibility and communication, on the one hand, and private sector expectations and macroeconomic volatility, on the other. We study an empirical dynamic stochastic general equilibrium (DSGE) model of the euro area, estimated by Smets and Wouters (2003). In our specification of the model, private agents observe changes in the monetary policy stance (the central bank's interest rate instrument), but they are unable to distinguish between temporary deviations from the central bank's monetary policy rule and permanent shifts in the inflation target. Agents therefore use the Kalman filter to construct optimal estimates of the current inflation objective and the temporary monetary policy shock and to make forecasts of the future path of

1. Cecchetti and Ehrmann (1999) and Levin, Natalucci, and Piger (2004) instead suggest that the introduction of a formal inflation target may lead to higher volatility in output, as the central bank shifts its preference toward stabilizing inflation and the economy moves along a fixed inflation/output volatility frontier. However, they do not find strong empirical support for this hypothesis. Benati (2006) finds that explicit inflation targeting (as in the United Kingdom, Sweden, Canada, and New Zealand) or the adoption of a quantitative definition of price stability (as in Switzerland and the euro area) has led to a significantly lower degree of inflation persistence. At the same time, he also finds that the United States has been able to achieve a low degree of inflation persistence since former Federal Reserve Chairman Paul Volcker's mandate, even without announcing an explicit inflation target.

monetary policy, and they update these estimates and forecasts as more information arrives. This learning behavior affects private agents' decisions and therefore all endogenous variables in the economy, with consequences for macroeconomic volatility in general.

Within this model, we first quantify the macroeconomic benefits of credibly announcing the (time-varying) level of the central bank's inflation objective. Such an announcement enables private agents to directly observe movements in the central bank's inflation objective and temporary deviations from the monetary policy rule. We then study the design of optimized rules for monetary policy within our framework, assuming a standard objective function for the central bank. In particular, we analyze whether rules optimized for the full information specification of the model need to be altered if agents do not observe the central bank's inflation objective.

Our results suggest that the macroeconomic benefits of credibly announcing the current level of the time-varying inflation target may be reasonably small as long as private agents correctly understand the stochastic processes governing the unobservable inflation target and the temporary policy shock and as long as the standard deviation of these shocks remains relatively small. We find that economic volatility decreases substantially after shocks to monetary policy. The overall gains from announcing the inflation target are fairly small, however, since these shocks account for a small fraction of overall volatility in our economy.[2] On the other hand, if private agents overestimate the volatility of the inflation target, the overall gains of credibly announcing the target can be large.

We also find that optimized monetary policy rules tend to respond more aggressively to inflation when private agents have imperfect information. By responding more aggressively to inflation, the central bank helps private agents in their learning process, thus reducing the deviation of inflation from the target with small consequences for volatility in the remaining macroeconomic variables.

Our model setup is closely related to those of Erceg and Levin (2003), Andolfatto, Hendry, and Moran (2005), and Kozicki and Tinsley (2005). Erceg and Levin (2003) study inflation persistence and the cost of disinflation in a model in which private agents cannot distinguish between temporary and permanent monetary policy shocks that follow

2. Our model is estimated over a period that does not include the great inflation of the 1970s, so monetary policy shocks are not very volatile and account for a small fraction of overall volatility. The effects of announcing the inflation target might be larger if monetary policy shocks were more volatile, but we do not explore this issue here.

stationary autoregressive processes, as in our setup. Their model is able to generate substantial inflation persistence and large disinflation costs as a consequence of the learning behavior of private agents, properties that are also present in our model. Andolfatto, Hendry, and Moran (2005) study the properties of inflation expectations in a model in which the temporary shock follows an autoregressive process, but the permanent shock follows a Bernoulli process. They show that common econometric tests tend to reject the rationality of inflation expectations when private agents learn about the properties of monetary policy shocks over time. Relative to these contributions, our purpose is somewhat broader, in that we try to understand the overall costs of imperfect information about monetary policy in terms of macroeconomic volatility, and we also study the appropriate design of monetary policy.

Moran (2005) uses a similar model to study the welfare effects of reducing the inflation target when agents learn about the shift in the inflation target using Bayesian updating. The welfare benefits are significant when comparing steady states, but much smaller if the transitional period of learning is also taken into account.

Kozicki and Tinsley (2005) use a reduced-form model of the U.S. economy to analyze the role of imperfect central bank credibility in the economy's transition to a new level of the inflation objective. Their model generates a rather large contribution of monetary policy to the volatility of inflation and other nominal variables after permanent shifts in the inflation target.

A number of other recent contributions study the consequences for monetary policy of private sector learning about the general structure of the economy in the stylized "New-Keynesian" model framework developed by Clarida, Galí, and Gertler (1999), Woodford (2003), and others. For instance, Nunes (2005) uses a model in which a proportion of private agents learn about the economic structure; he finds that his model explains well the transitional dynamics of the economy after a disinflationary shock. Gaspar, Smets, and Vestin (2006a, 2006b, 2006c) show that in order to reduce the persistence and volatility of inflation, optimal monetary policy responds more persistently to shocks when private agents learn about the structure of the economy than when they operate under rational expectations. Similarly, Molnár and Santoro (2006) show that optimal monetary policy responds more aggressively to shocks under private sector learning than when private agents have rational expectations. We present similar results in our framework.

Also in a New-Keynesian framework, Orphanides and Williams (2007) study monetary policy in a small estimated model in which the central bank learns about the natural rates of unemployment and interest and private agents learn about the structure of the economy. They show that the explicit communication of the central bank's inflation objective substantially improves macroeconomic performance under a suboptimal policy, while the gains are fairly modest under the optimal policy. Rudebusch and Williams (2008) instead study how the publication of the central bank's interest rate projections can better align private sector expectations when private agents do not observe either the coefficients in the monetary policy rule or the central bank's target level for inflation. Aoki and Kimura (2007) show that the learning processes of the central bank and the private sector imply that higher-order beliefs become relevant, leading to an increase in macroeconomic persistence and volatility. They also show that private sector learning can reduce macroeconomic volatility over time, and announcing the inflation objective can help the central bank to estimate the natural rate of interest.

A different but related strand of the literature explores the implications of variability in the central bank's preferences or in the inflation objective for the dynamic properties of the economy, under the assumption that central bank preferences and objectives are perfectly observable and credible. Cogley, Primiceri, and Sargent (2008) attribute the decline in the persistence of the inflation gap (defined as the deviation of inflation from the measured time-varying inflation objective) to the decline in the variance of permanent shocks to a time-varying but observable inflation target. Ireland (2007) argues that monetary policy has increased the degree of inflation persistence by shifting the inflation objective in accordance with realized supply-side shocks, to effectively accommodate them. Finally, Dennis (2006) and Beechey and Österholm (2007) argue that shifts in the central bank's preferences, toward a sharper focus on inflation stabilization at the expense of output stabilization, are behind the lower degrees of macroeconomic persistence and in particular inflation persistence in the U.S. economy since the time of Paul Volcker's chairmanship of the Federal Reserve.

In contrast to these papers, we study an estimated medium-sized DSGE model often used for quantitative analysis. We show that while announcing the inflation target reduces the volatility originating in shocks to monetary policy, this volatility is small relative to that from the remaining shocks in the model. This result partly reflects the fact that the standard deviation of monetary policy shocks in our model, which is calibrated for a period with broadly anchored inflation trends, is relatively small compared, for instance, with the great inflation period of the 1970s.

Finally, Beechey (2004) and Gürkaynak, Sack, and Swanson (2005) use similar models to explore the relationship between monetary policy and the yield curve. Beechey uses a stylized model with optimizing agents to study the effects on the yield curve of central bank private information concerning macroeconomic shocks and the central bank's preferences, following Ellingsen and Söderström (2001, 2005). In her model, the central bank sets monetary policy optimally given a quadratic loss function, and private agents use a Kalman filter to construct estimates of the unobservable shocks. Gürkaynak, Sack, and Swanson (2005) use a small macroeconometric model (without complete microfoundations) to study the effects of macroeconomic announcements on the yield curve. They rationalize the large response of long-term forward rates found in case studies through a model in which the central bank's inflation target moves with actual inflation, but the target is unobservable to the private sector, and private agents use a signal extraction methodology to estimate the current inflation target from observed movements in the short-term interest rate.[3] We deviate from these authors by studying an estimated medium-scale DSGE model. While our model is also suited to studying the behavior of the yield curve, we focus here on macroeconomic volatility in general.

Our paper is organized as follows. We present the structure of the model economy, following Smets and Wouters (2003), and discuss the restrictions on the private sector's information set and the Kalman filter used to construct estimates of the two monetary policy shocks in section 1. We then present the results concerning volatility in private expectations and the macroeconomy in section 2, and we study the design of optimized rules for monetary policy in section 3. Finally, we summarize our findings and conclude in section 4.

1. The Model

We use the dynamic stochastic general equilibrium model developed and estimated on quarterly euro area data by Smets and

3. Gürkaynak, Levin, and Swanson (2006) use a similar model.

Wouters (2003).[4] Here we briefly present the log-linearized version of the model; we refer to Smets and Wouters (2003) for a more extensive discussion.

1.1 The Structural Model

Households choose consumption, labor supply, and holdings of a one-period bond to maximize lifetime utility, which depends on consumption relative to an external habit level and leisure. Utility maximization subject to a standard budget constraint gives the log-linearized consumption Euler equation

$$C_t = \frac{h}{1+h}C_{t-1} + \frac{1}{1+h}E_t C_{t+1} - \frac{1-h}{\sigma_c(1+h)}(R_t - E_t\pi_{t+1} - \varepsilon_t^b), \tag{1}$$

where C_t is aggregate consumption, R_t is the nominal one-period interest rate (measured at a quarterly rate), π_t is the one-period rate of inflation, $h \in [0, 1)$ determines the importance of habits, $\sigma_c > 0$ is related to the intertemporal elasticity of substitution, and ε_t^b is a shock to household preferences.

Households act as price setters in the labor market, but wages are set in a staggered fashion: a fraction $1 - \xi_w$ of wages are reset in a given period, and the remaining fraction is partially indexed to past inflation. This gives the log-linearized real wage equation

$$\begin{aligned} W_t = {} & \frac{\beta}{1+\beta}E_t W_{t+1} + \frac{1}{1+\beta}W_{t-1} + \frac{\beta}{1+\beta}E_t\pi_{t+1} - \frac{1+\beta\gamma_w}{1+\beta}\pi_t + \frac{\gamma_w}{1+\beta}\pi_{t-1} \\ & - \frac{(1-\beta\xi_w)(1-\xi_w)\lambda_w}{\left[\lambda_w + (1+\lambda_w)\sigma_l\right](1+\beta)\xi_w}\left[W_t - \sigma_l L_t - \frac{\sigma_c}{1-h}(C_t - hC_{t-1}) + \varepsilon_t^l\right] + \eta_t^w, \end{aligned} \tag{2}$$

where W_t is the real wage, L_t is aggregate labor demand, $\beta \in [0, 1]$ is a discount factor, γ_w is the degree of wage indexation, σ_l measures the elasticity of labor supply, λ_w is the steady-state wage markup, ε_t^l is a labor supply shock, and η_t^w is a wage markup shock.

4. This model is based on Christiano, Eichenbaum, and Evans (2005). Other versions of the model include Smets and Wouters (2005, 2007), Levin and others (2005), and Del Negro and others (2005). The model specification used here corresponds to that estimated by Smets and Wouters (2003), and it differs slightly from the specification presented in their paper. Frank Smets and Raf Wouters kindly provided the specification of the estimated model.

Households also own the capital stock, which is rented to firms producing intermediate goods at the rental rate r_t^k. They can increase the supply of capital by either investing in new capital or changing the utilization rate of installed capital, and both actions are costly in terms of foregone consumption. The optimal choice of the capital stock, investment, and the utilization rate give the log-linearized conditions

$$I_t = \frac{1}{1+\beta} I_{t-1} + \frac{\beta}{1+\beta} E_t I_{t+1} + \frac{1}{\varphi_i (1+\beta)} Q_t + \frac{1}{\varphi_i (1+\beta)\left[1-\beta\rho_i (1-\tau)\right]} \varepsilon_t^i, \tag{3}$$

$$Q_t = -(R_t - E_t \pi_{t+1}) + \beta(1-\tau) E_t Q_{t+1} + \left[1-\beta(1-\tau)\right] \frac{1+\psi}{\psi} E_t r_{t+1}^k + (1+\beta)\varphi_i \eta_t^q, \tag{4}$$

and

$$K_t = (1-\tau)\, K_t-1 + \tau I_{t-1}, \tag{5}$$

where I_t is investment, Q_t is Tobin's Q, K_t is the total capital stock, φ_i is the second derivative of the investment adjustment cost function, τ is the depreciation rate of capital, ψ is the elasticity of the capital utilization cost function, ε_t^i is a shock to the investment cost function, and η_t^q is a shock that captures variations in the external finance premium.

There is a single final good that is produced under perfect competition using a continuum of intermediate goods. These intermediate goods, in turn, are produced under monopolistic competition using capital and labor inputs with a Cobb-Douglas technology. Prices on intermediate goods are staggered as in Calvo (1983), so a fraction $1-\xi_p$ of prices are reset in a given period. The remaining prices are partially indexed to past inflation.[5] The optimal price-setting behavior then implies that

5. More recent models instead assume that the prices that are not reoptimized are indexed partly to past inflation and partly to the (nonzero) inflation target or steady-state inflation (see, for instance, Smets and Wouters, 2007). This assumption would imply that changes in the perceived inflation target have a direct effect on price setting and therefore on welfare (see below).

aggregate inflation is determined by the New-Keynesian Phillips curve:

$$\pi_t = \frac{\beta}{1+\beta\gamma_p} E_t \pi_{t+1} + \frac{\gamma_p}{1+\beta\gamma_p} \pi_{t-1} + \frac{(1-\beta\xi_p)(1-\xi_p)}{\xi_p(1+\beta\gamma_p)} \left[\alpha r_t^k + (1-\alpha) W_t - \varepsilon_t^a \right] + \eta_t^p, \tag{6}$$

where γ_p is the degree of indexation to past inflation, α is the Cobb-Douglas parameter on capital, ε_t^a is a technology shock, and η_t^p is a price markup shock. Profit optimization also gives the labor demand function,

$$L_t = -W_t + \frac{1+\psi}{\psi} r_t^k + K_{t-1}. \tag{7}$$

Finally, market clearing implies that

$$Y_t = \frac{\alpha\varphi_y}{\psi} r_t^k + \alpha\varphi_y K_{t-1} + (1-\alpha)\varphi_y L_t + \varphi_y \varepsilon_t^a, \tag{8}$$

where Y_t is the aggregate level of output, and φ_y is equal to one plus the share of the fixed cost in production. The resource constraint gives

$$Y_t = c_y C_t + \tau k_y I_t + \varepsilon_t^g, \tag{9}$$

where c_y and k_y are the steady-state ratios of consumption and capital to output, and ε_t^g is government spending.[6]

The model contains eight structural shocks. Three of these—the price and wage markup shocks, η_t^p and η_t^w, and the equity premium shock, η_t^q—are assumed to be white noise with variances σ_p^2, σ_w^2, and σ_q^2. The remaining five shocks—to preferences, the investment adjustment cost, technology, labor supply, and government spending—are assumed to follow the stationary autoregressive processes:

6. Onatski and Williams (2004) add a term on the right-hand side of equation (9) to include capital utilization costs, which was omitted in the original Smets and Wouters (2003) model. We choose to use the latter specification, which was estimated on euro area data.

$$\varepsilon_t^j = \rho_j \varepsilon_{t-1}^j + \eta_t^j, \qquad j = b,i,a,l,g, \tag{10}$$

where $\rho_j \in [0, 1)$, and the innovations η_t^j are white noise with variance σ_j^2.

1.2 Monetary Policy

For the specification of monetary policy, we depart slightly from Smets and Wouters (2003) by assuming that monetary policy is set according to the following interest rate rule:[7]

$$R_t = (1 - g_r)\left[\pi_t^* + g_\pi\left(\pi_{t-1} - \pi_t^*\right) + g_y\left(Y_{t-1} - Y_{t-1}^n\right)\right] + g_r R_{t-1} + \varepsilon_t^r. \tag{11}$$

Thus, the nominal one-period interest rate, R_t, is a linear combination of the deviation of the previous period's rate of inflation, π_{t-1}, from the central bank's current inflation objective, π_t^*, the previous period's output gap (the log deviation of real output, Y_t, from its natural level, Y_t^n), and the previous period's interest rate.[8] There are two exogenous elements in the policy rule: the inflation objective, π_t^*, and the monetary policy shock, ε_t^r. In general, these are assumed to follow stationary first-order autoregressive processes:

$$\pi_t^* = \rho_* \pi_{t-1}^* + \eta_t^* \tag{12}$$

7. Smets and Wouters (2003) instead specify their monetary policy rule as follows:

$$R_t = (1 - g_r)\left[\pi_t^* + g_\pi(\pi_{t-1} - \pi_t^*) + g_y(Y_t - Y_t^n)\right] + g_{\Delta\pi}(\pi_t - \pi_{t-1})$$
$$+ g_{\Delta y}\left[(Y_t - Y_t^n) - (Y_{t-1} - Y_{t-1}^n)\right] + g_r R_{t-1} + \varepsilon_t^r,$$

and obtain the estimates $g_\pi = 1.684$, $g_y = 0.099$, $g_{\Delta\pi} = 0.140$, and $g_{\Delta y} = 0.159$, and $g_r = 0.961$. Also, they estimate the autoregressive coefficient of the inflation target to $\rho_* = 0.924$. Using this rule instead of our rule gives very similar qualitative results. We also experimented with rules including the current rate of inflation and output gap, and rules with persistent monetary policy shocks rather than gradual behavior, as advocated by Rudebusch (2002). Again, the results with these rules are similar to those presented here.

8. The natural output level is defined as the output level in the equilibrium with flexible wages and prices and without the shocks to the wage and price markups and the external finance premium. The presence of the past inflation rate and output gap in the policy rule implies that monetary policy only responds to predetermined variables. In the terminology of Svensson and Woodford (2004), the policy rule is an operational or explicit instrument rule, as opposed to an implicit instrument rule that includes variables that are not predetermined. Such rules are also recommended by McCallum (1997).

and

$$\varepsilon_t^r = \rho_r \varepsilon_{t-1}^r + \eta_t^r, \tag{13}$$

where $\rho_*, \rho_r \in [0,1)$ and η_t^* and η_t^r are white noise processes with variances σ_*^2 and σ_r^2. However, we assume that the inflation target is very persistent (close to a random walk), while the monetary policy shock is almost white noise.[9]

1.3 Parameterization

For the structural parameters, we use the calibrated or estimated values from Smets and Wouters (2003), summarized in table 1. These estimates were obtained using quarterly data from the euro area from 1980:2 to 1999:4. For the monetary policy parameters, we start with a fairly standard calibration of the policy rule (11), setting $g_\pi = 2.0$, $g_y = 0.2$, and $g_r = 0.9$ (also reported in table 1), while in section 3 we choose the policy rule parameters to minimize a standard objective function for the central bank. The inflation objective, π_t^*, is assumed to be a near-random walk, with $\rho_* = 0.99$, while the temporary monetary policy shock, ε_t^r, is essentially white noise, with $\rho_r = 0.01$. Changes in the inflation objective are thus highly persistent (the half-life of a shock is close to 70 quarters), while other deviations from the policy rule are entirely temporary. The standard deviations of the two monetary policy shocks are set to the Smets and Wouters (2003) estimates: $\sigma_* = 0.017$ percentage point and $\sigma_r = 0.081$ percentage point, respectively. Innovations to the temporary shock are thus almost five times as volatile as those to the inflation target.[10] However, since the model is estimated on a sample with changing monetary regimes and high inflation in Europe, the estimated volatility of the inflation target is likely an upper bound on the true volatility.

9. Time variation in the inflation target could be due to true time variation in the preferred inflation rate of an individual central banker, time variation in the composition of the monetary policy committee (and thus in the average preferred inflation rate of the committee), or time variation in the committee's concerns for the zero lower bound of interest rates. We assume that the inflation target is close to a random walk, so changes in the inflation target are not expected to be reversed immediately, but are seen as close to permanent.

10. Andolfatto, Hendry, and Moran (2005) model the inflation target as a Bernoulli process, so occasional shifts in the inflation target are followed by long periods of a constant target. Our specification implies that the inflation target changes in every period, but with a very low variance. One advantage of this specification is that the Kalman filter produces optimal forecasts of the future temporary shock and inflation target.

Table 1. Parameter Values[a]

Parameter	*Value*	*Description*
Calibrated parameters		
β	0.99	Discount factor
τ	0.025	Depreciation rate of capital
α	0.30	Capital share in production
k_y	8.8	Capital/output ratio
c_y	0.60	Consumption/output ratio
λ_w	0.5	Average wage markup
Estimated structural parameters		
φ_i	6.771	Investment adjustment cost parameter
σ_c	1.353	Coefficient of relative risk aversion
h	0.573	Consumption habit parameter
σ_l	2.400	Elasticity of labor supply
φ_y	1.408	Fixed cost in production
ψ	0.169	Elasticity of capital utilization cost function
ξ_w	0.737	Calvo wage parameter
ξ_p	0.908	Calvo price parameter
γ_w	0.763	Rate of wage indexation
γ_p	0.469	Rate of price indexation
Estimated autoregressive parameters		
ρ_b	0.855	Preference shock
ρ_i	0.927	Investment cost shock
ρ_a	0.823	Productivity shock
ρ_l	0.889	Labor supply shock
ρ_g	0.949	Government spending shock
Estimated standard deviations		
σ_b	0.336	Preference shock
σ_i	0.085	Investment cost shock
σ_q	0.604	Equity premium shock
σ_a	0.598	Productivity shock
σ_p	0.160	Price markup shock
σ_w	0.289	Wage markup shock
σ_l	3.520	Labor supply shock
σ_g	0.325	Government spending shock
σ_*	0.017	Inflation objective
σ_r	0.081	Temporary monetary policy shock

Table 1. (continued)

Parameter	*Value*	*Description*
Calibrated monetary policy parameters		
g_π	2.0	Coefficient on inflation
g_y	0.2	Coefficient on output gap
g_r	0.9	Coefficient on lagged interest rate
ρ_*	0.99	Persistence in inflation objective
ρ_r	0.01	Persistence in temporary monetary policy shock

Source: Smets and Wouters (2003).
a. The estimated parameter values are taken from Smets and Wouters (2003) (the mode of their estimated posterior distribution), using euro area data from 1980:2 to 1999:4.

1.4 Private Sector Information

Our key assumption is that private agents are unable to distinguish between the two exogenous shocks to the monetary policy rule—namely, the inflation objective, π_t^*, and the temporary monetary policy, shock ε_t^r. However, they are perfectly informed about all other aspects of the economy. Since they can observe the interest rate, R_t, private agents can use the policy rule (11) to back out the combination

$$\hat{\varepsilon}_t = (1-g_r)(1-g_\pi)\pi_t^* + \varepsilon_t^r, \tag{14}$$

and then use the Kalman filter to calculate optimal estimates of the inflation target, π_t^*, and the policy shock, ε_t^r.[11] The Kalman filter is thus characterized by the state equation

$$\begin{aligned}\begin{bmatrix}\pi_{t+1}^*\\ \varepsilon_{t+1}^r\end{bmatrix} &= \begin{bmatrix}\rho_* & 0\\ 0 & \rho_r\end{bmatrix}\begin{bmatrix}\pi_t^*\\ \varepsilon_t^r\end{bmatrix} + \begin{bmatrix}\eta_{t+1}^*\\ \eta_{t+1}^r\end{bmatrix}\\ &\equiv \mathbf{F}\begin{bmatrix}\pi_t^*\\ \varepsilon_t^r\end{bmatrix} + \begin{bmatrix}\eta_{t+1}^*\\ \eta_{t+1}^r\end{bmatrix}\end{aligned} \tag{15}$$

11. As mentioned earlier, this specification is similar to those of Erceg and Levin (2003) and Andolfatto, Hendry, and Moran (2005).

and the observation equation

$$\begin{aligned}\hat{\varepsilon}_t &= \left[(1-g_r)(1-g_\pi) \quad 1\right]\begin{bmatrix}\pi_t^*\\ \varepsilon_t^r\end{bmatrix}\\ &\equiv \mathbf{H}'\begin{bmatrix}\pi_t^*\\ \varepsilon_t^r\end{bmatrix}.\end{aligned} \tag{16}$$

Optimal forecasts of the future inflation target and policy shock are then calculated as

$$\begin{bmatrix}\hat{E}_t\pi_{t+1}^*\\ \hat{E}_t\varepsilon_{t+1}^r\end{bmatrix} = (\mathbf{F}-\boldsymbol{\kappa}\mathbf{H}')\begin{bmatrix}\hat{E}_{t-1}\pi_t^*\\ \hat{E}_{t-1}\varepsilon_t^r\end{bmatrix} + \boldsymbol{\kappa}\mathbf{H}'\begin{bmatrix}\pi_t^*\\ \varepsilon_t^r\end{bmatrix}, \tag{17}$$

where $\boldsymbol{\kappa}$ is the Kalman gain.[12] The optimal estimates of the current target and policy shock are given by

$$\begin{bmatrix}\hat{E}_t\pi_t^*\\ \hat{E}_t\varepsilon_t^r\end{bmatrix} = \mathbf{F}^{-1}\begin{bmatrix}\hat{E}_t\pi_{t+1}^*\\ \hat{E}_t\varepsilon_{t+1}^r\end{bmatrix}. \tag{18}$$

Although private agents' estimates of π_t^* and ε_t^r do not enter the model explicitly, these estimates affect private expectations of future monetary policy and therefore indirectly affect all other endogenous variables. Since agents learn over time, private expectations are generally biased predictors of future outcomes. This bias may

12. To determine the Kalman gain $\boldsymbol{\kappa}$, let $\boldsymbol{\Sigma}$ be the variance-covariance matrix of $[\eta_{t+1}^* \ \eta_{t+1}^r]'$ and let $\mathbf{P}_{t+1|t}$ denote the mean-squared error of the forecast of $\mathbf{x}_{t+1} \equiv [\pi_{t+1}^* \ \varepsilon_{t+1}^r]'$, that is,

$$\mathbf{P}_{t+1|t} = E\left[(\boldsymbol{\xi}_{t+1} - \hat{E}_t\boldsymbol{\xi}_{t+1})(\boldsymbol{\xi}_{t+1} - \hat{E}_t\boldsymbol{\xi}_{t+1})'\right].$$

Starting from the unconditional mean-squared error, given by $\mathrm{vec}(\mathbf{P}_{1|0}) = (\mathbf{I} - \mathbf{F}\otimes\mathbf{F})^{-1}\mathrm{vec}(\boldsymbol{\Sigma})$, the Kalman gain matrix and the mean-squared error are found by iterating on

$$\boldsymbol{\kappa}_t = \mathbf{F}\mathbf{P}_{t|t-1}\mathbf{H}(\mathbf{H}'\mathbf{P}_{t|t-1}\mathbf{H})^{-1}$$

and

$$\mathbf{P}_{t+1|t} = (\mathbf{F} - \boldsymbol{\kappa}_t\mathbf{H}')\mathbf{P}_{t|t-1}(\mathbf{F} - \boldsymbol{\kappa}_t\mathbf{H}')' + \boldsymbol{\Sigma}.$$

See Hamilton (1994, chap. 13) for details. Thus, the Kalman gain depends on all elements of **F, H**, and $\boldsymbol{\Sigma}$, that is, on g_π, g_r, ρ_*, ρ_r, σ_*, and σ_r.

lead private agents to make inefficient decisions, so the economy may experience inefficient volatility relative to the case of perfect information. If the central bank were instead to announce the current level of the inflation target, π_t^*, private agents would be able to perfectly infer the realization of the shock ε_t^r, and the perfect-information equilibrium would be attainable. We next study the effects on macroeconomic volatility of announcing the inflation target—that is, we move from the equilibrium with imperfect information to that with perfect information.

2. Macroeconomic Dynamics and Volatility

This section explores the dynamics of our model economy, first in terms of impulse responses to the two monetary policy shocks and then in terms of the volatility of simulated time series.

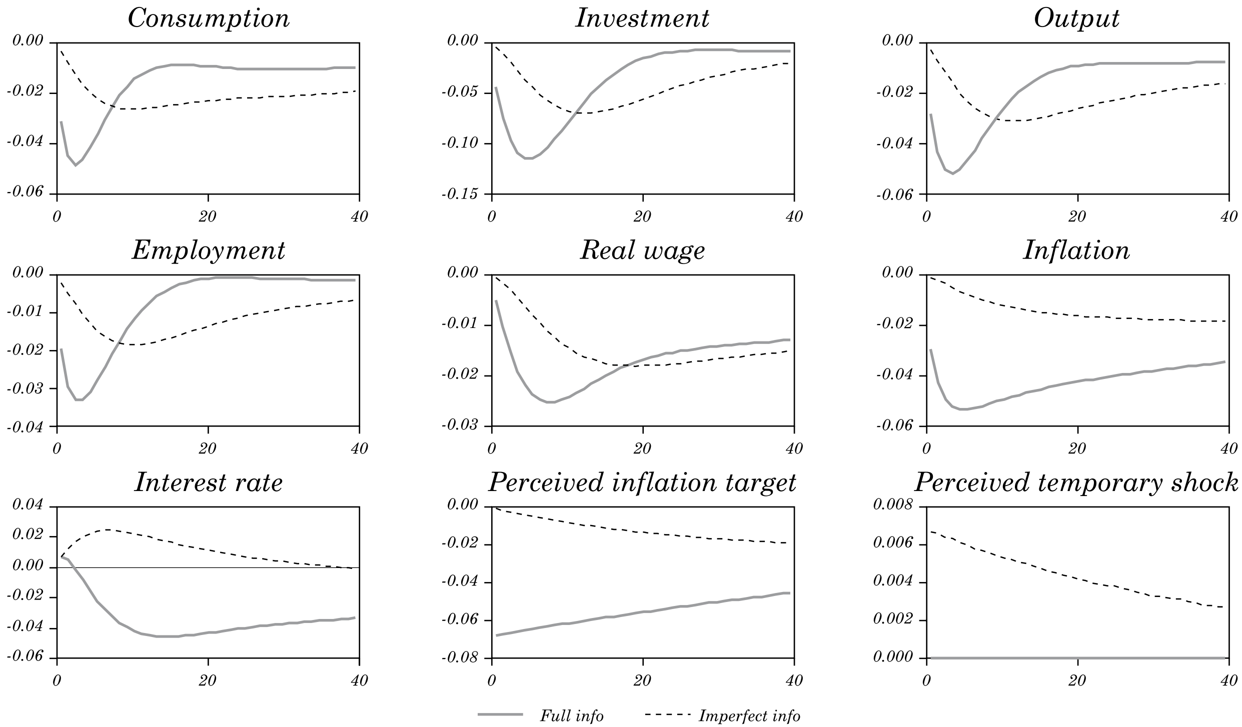
2.1 The Effects of Monetary Policy Shocks

Figures 1 and 2 show impulse responses to one-standard-deviation innovations to the inflation objective and the temporary monetary policy shock, respectively. The solid lines represent the impulse responses (and forecasts) in the benchmark case of full information (when all shocks are observable), the dash-dotted lines represent optimal forecasts with imperfect information, and the dashed lines show the effects of shocks on the economy when there is imperfect information and agents learn over time.[13]

Consider first the case of full information, represented by the solid lines in figures 1 and 2. Figure 1 shows impulse responses and forecasts after a negative shock to the inflation target, π_t^*. With full information, private agents immediately notice that the inflation target has decreased, so the perceived target jumps down to its new level and agents adjust their expectations accordingly. As a consequence, inflation falls in the initial period, and the central bank is able to increase the real interest rate with only a slight increase in the nominal interest rate, which is soon reversed. This leads to a decrease in consumption, investment, output, employment, and the real wage and, therefore, a drop in inflation. When inflation and the

13. In all figures and tables, the inflation and interest rates are measured on an annualized basis. The appendix outlines how we simulate the model and construct impulse responses with imperfect information.

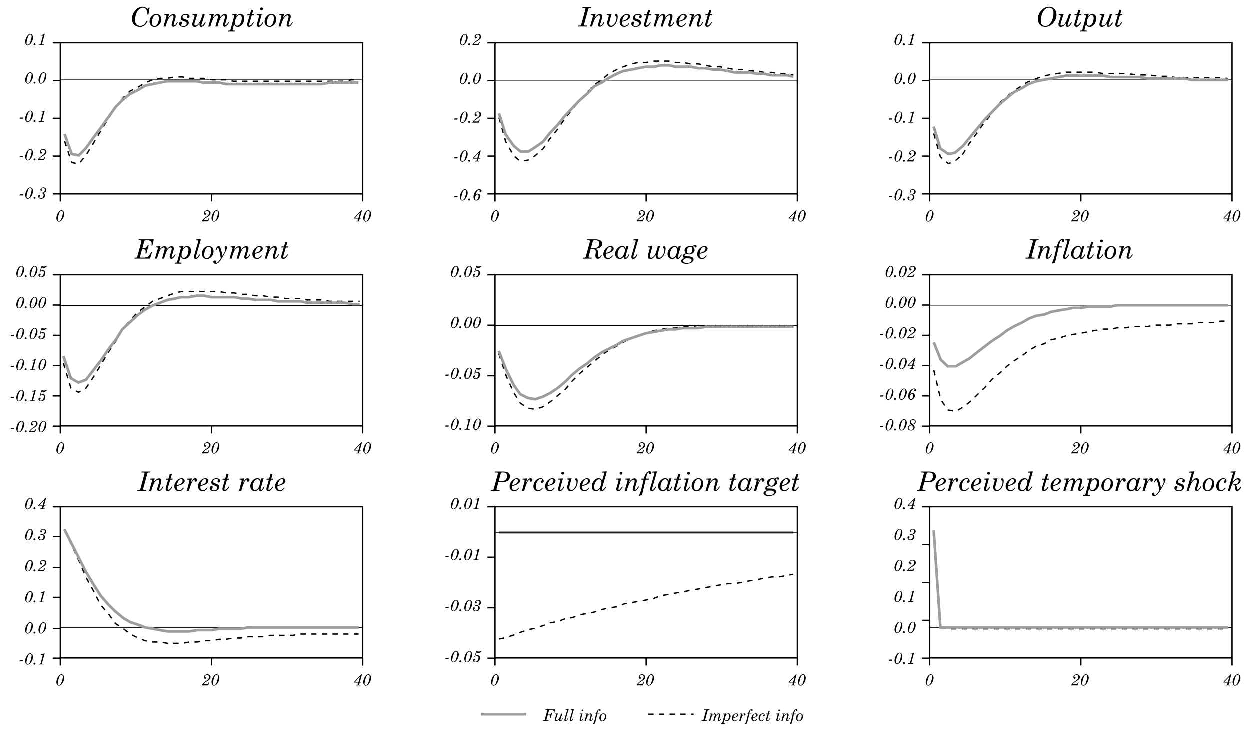
Figure 1. Impulse Responses to an Inflation Target Shock[a]



Source: Authors' calculations.

a. This figure shows impulse responses to a one-standard-deviation negative innovation to the inflation target, π_t^*.

Figure 2. Impulse Responses to a Temporary Monetary Policy Shock[a]



Source: Authors' calculations.

a. This figure shows impulse responses to a one-standard-deviation innovation to the temporary monetary policy shock, ε_t^r.

time-varying inflation target are close, they move back together to the initial level, and the nominal interest rate follows them back. The real interest rate is therefore close to its neutral level, and all real variables return toward steady state. There is thus a hump-shaped response of all variables, with the maximum effect on output (around 5 basis points) after four to six quarters.

After a positive innovation to the temporary monetary policy shock, ε_t^r, in figure 2, the interest rate increases by the full amount of the shock (32 basis points), and the real interest rate increases even more as expected inflation falls. This leads to a reduction in all real variables, which motivates the decrease in inflation. Again, all responses are hump-shaped, and the maximum effects on output (–20 basis points) and inflation (–4 basis points) occur after three quarters.

Under imperfect information, private agents use the Kalman filter to make optimal estimates of the current and future inflation target and policy shock, and they adjust their expectations accordingly. Figure 1 shows that after a negative inflation target shock, a persistent increase in the interest rate is necessary to reduce inflation expectations. Private agents observe the small increase in the nominal interest rate, and they attribute this partly to a negative inflation target shock and partly to a positive temporary policy shock. As they know that the inflation target is much less volatile than the temporary shock, their optimal estimate of the inflation target initially falls very little (by 0.09 basis point), while the estimate of the temporary shock increases more (by 0.67 basis point).

As time passes, the central bank increases the interest rate further, and when agents update their information set, they find it increasingly likely that the inflation target has in fact decreased. Inflation therefore falls further, and all real variables continue to drop as the real interest rate increases. As agents learn, the perceived and actual inflation target slowly converge, and the perceived temporary monetary policy shock approaches zero. This slow learning process implies that all variables respond more gradually and persistently to the inflation target shock than in the case of full information, and the maximum effects on output now occur after twelve quarters. As in Erceg and Levin (2003) and Nunes (2005), the presence of imperfect information substantially increases the real cost of disinflation.

After a temporary policy shock in figure 2, private agents again observe an increase in the nominal interest rate and attribute almost all of this (32 basis points) to a positive temporary shock and very little (4 basis points) to a negative inflation target shock. In the initial period, the main difference compared with the full information case is a larger fall in inflation, as private agents believe that the inflation objective is lower. Thus, the same increase in the interest rate leads to a larger increase in the real interest rate under imperfect information, with a larger effect on real variables.

As agents learn over time, the monetary policy tightening leads to a slightly deeper recession than under full information, and the central bank needs to lower the interest rate below the initial level to stimulate the economy. The real variables then return toward steady state, often with some overshooting, while inflation and the interest rate return very slowly to their initial levels, together with the perceived inflation target.

To summarize, imperfect information about the two policy shocks implies that agents optimally attribute almost all unexpected movements in the nominal interest rate to the more volatile temporary shock and very little to the persistent inflation target shock, which is less volatile. To persuade private agents that the inflation target is lower, the central bank needs to tighten policy more, resulting in a deeper recession. The learning process implies that all variables respond more gradually to an inflation target shock with imperfect than with full information. The temporary policy shock, on the other hand, has very similar effects under imperfect and full information, as agents attribute most of the unexpected interest rate movement to the temporary shock.

2.2 Imperfect Information and Macroeconomic Volatility

It is clear from the impulse responses and forecasts in figures 1 and 2 that imperfect information about the two monetary policy shocks has large effects on the dynamic behavior of the economy and private sector forecasts, particularly after shocks to the inflation target. This impression is confirmed by panel A of table 2, which shows the variance in some key macroeconomic variables in the model that is due to the two monetary policy shocks.[14]

14. The reported variances are averages across 1,000 simulated samples of 10,000 observations (after discarding the initial 500 observations). Inflation and the interest rate are in annualized terms, so $\bar{\pi}_t = 4\pi_t$ and $\bar{R}_t = 4R_t$.

Table 2. Variances of Simulated Data under Full and Imperfect Information[a]

Type of information	C_t	Y_t	I_t	L_t	W_t	$\overline{\pi}_t$	$Y_t - Y_t^n$	$\overline{R}_t$	$\overline{\pi}_t - \overline{\pi}_t^*$
A. Monetary policy shocks only									
Full information	0.21	0.24	1.15	0.094	0.068	0.140	0.24	0.42	0.025
Imperfect information	0.26	0.30	1.44	0.120	0.079	0.089	0.30	0.35	0.150
B. All shocks									
Full information	6.89	7.12	77.23	3.54	1.60	1.34	3.76	1.29	1.22
Imperfect information	6.94	7.18	77.51	3.57	1.61	1.29	3.82	1.22	1.34

Source: Authors' calculations.

a. This table reports simulated variances (averages over 1,000 simulated series of 10,000 observations) in the models with full information and with imperfect information. Inflation and the interest rate are in annualized terms: $\overline{\pi}_t = 4\pi_t$ and $\overline{R}_t = 4R_t$.

Conditional on the two monetary policy shocks, most variables are considerably more volatile under imperfect information than under full information, with the exception of inflation and the interest rate. The variance of the real variables resulting from monetary policy shocks is 20 to 25 percent larger with imperfect information than with full information, while inflation and the nominal interest rate are considerably less volatile with imperfect information. A review of figures 1 and 2 reveals that this effect on volatility is mainly due to the effect of shocks to the inflation target, where the response of all real variables is more gradual with imperfect information an leads to larger volatility. Since inflation target shocks have a smaller impact on inflation and the interest rate with imperfect information than with full information, these variables are also less volatile. Thus, imperfect information about the monetary policy shocks has an important impact on macroeconomic volatility, conditional on the two monetary policy shocks.

However, the remaining eight shocks are observable to the private sector and therefore are not affected by the information restrictions, so the total effect of imperfect information on macroeconomic volatility depends on the overall contribution of the monetary policy shocks to volatility. Panel B of table 2 reports the effects of imperfect information on aggregate volatility. This panel reveals that imperfect information has very small effects on the volatility of macroeconomic variables once we take into account all structural shocks: the variance of most real variables increases by less than one percent. The largest effects are on inflation and interest rate volatility, which is lower with imperfect information, and on the volatility of inflation around the target, which is substantially higher. This is because actual inflation adjusts slowly to changes in the inflation target when private agents cannot directly observe the target (see figure 1). Nevertheless, the overall effects of imperfect information on macroeconomic volatility—and thus the potential benefits of credibly announcing the central bank's target for inflation—seem modest.[15]

15. In the case of full information, the inflation target is not constant but varies over time. Since the volatility of the inflation target is very low, however, the outcome with a known constant inflation target is very similar to the full information case reported here.

2.3 The Role of Private Sector Information about Monetary Policy Shock Processes

The above results suggest that the presence of imperfect information has small effects on macroeconomic volatility, so the gains of announcing the exact inflation target are small. As discussed earlier, however, the response of private expectations to the unobservable shocks depends crucially on the perceived volatility of the shocks. In the benchmark calibration, the temporary shock is considerably more volatile than the inflation target shock. Private agents therefore attribute a small fraction of the unexpected movement in the interest rate to the inflation target and a large fraction to the temporary shock, with a small effect on overall volatility as a result.

If the central bank is unwilling to announce its inflation target, it may be difficult for private agents to estimate the variance of the target. In this section, we therefore analyze an alternative scenario in which private agents overestimate the variance of the inflation target. Specifically, we set the perceived standard deviation of the inflation target five times larger than the actual standard deviation, so the perceived standard deviation is $\hat{\sigma}_* = 0.085$, which is of similar magnitude to the standard deviation of the temporary policy shock. In this situation, private agents will attribute a greater part of the unexpected movements in the interest rate to inflation target shocks than when they know the true variance of the inflation target.

To illustrate how private agents' perceptions affect the speed with which they update their forecasts as new information arrives, figures 3 and 4 show how the sensitivity of the optimal forecasts for the inflation target and the temporary policy shock to the observed interest rate depends on the perceived coefficients in the monetary policy rule and the persistence and volatility of the two monetary policy shocks.[16] Figure 3 reveals that private agents' inflation target forecast is more sensitive to unexpected changes in the observed interest rate either when the central bank is more responsive to inflation deviations from target (that is, when g_π is large) or when the inflation target process is seen to be more persistent or volatile (that is, when ρ_* or

16. The figures thus plot the two updating coefficients in the Kalman gain, κ, in equation (17) as a function of g_π, g_r, ρ_*, ρ_r, σ_*, and σ_r. Rudebusch and Williams (2008) also discuss how the private sector's information set affects the optimal updating scheme in a model in which private agents are unable to observe the inflation target and the central bank helps private agents by publishing its forecast for the interest rate.

σ_* is large).[17] A larger central bank response to the lagged interest rate or more persistence or volatility in the temporary policy shock instead reduce the effect of new information on the inflation target forecast. Figure 4 shows the opposite pattern for the sensitivity of the temporary shock forecast. In our benchmark calibration (marked by vertical lines in the figures), private agents' forecasts are particularly sensitive to the perceived volatility of the inflation target: an increase in the perceived volatility leads to much larger effects of unexpected interest rate movements on the optimal inflation target forecast, but smaller effects on the forecast of the temporary shock.

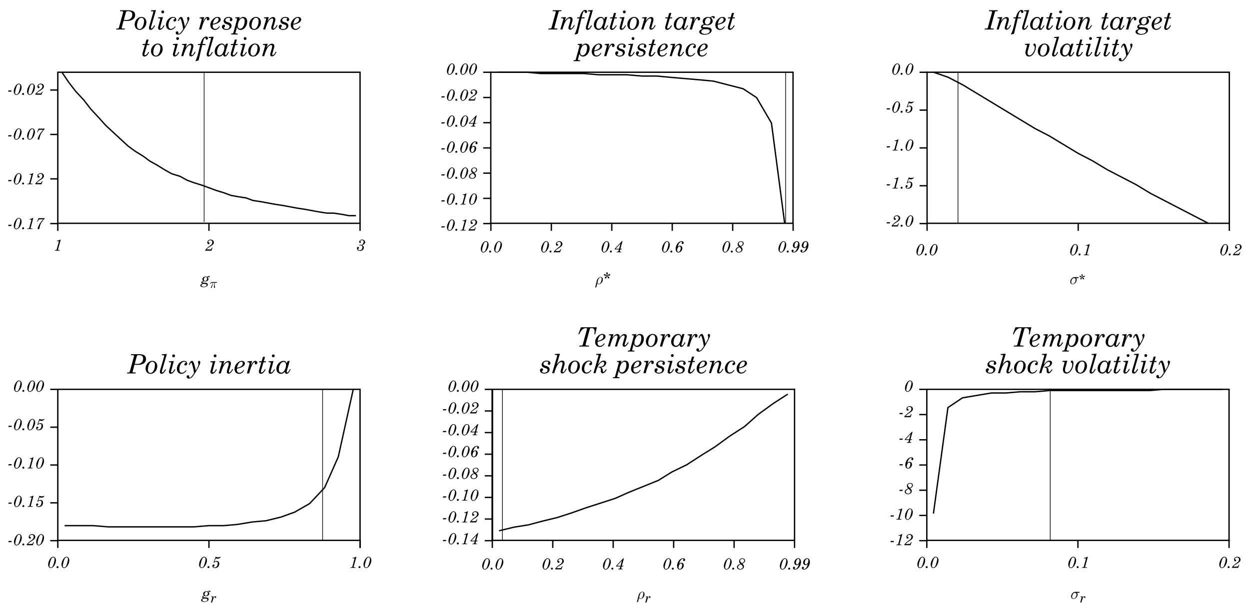
Figures 5 and 6 show impulse responses to innovations to the two monetary policy shocks when private agents overestimate the variance of the inflation target. (The responses under full information are the same as in figures 1 and 2.) After an inflation target shock in figure 5, the larger movements in the perceived inflation target imply that inflation falls faster than when private agents know the variance of the inflation target. The increase in the nominal interest rate now translates into a larger increase in the real interest rate than when private agents know the true variance of the inflation target, with a deeper and less gradual recession as a result. The central bank reduces the nominal interest rate toward the new target level more quickly, and as the perceived inflation target approaches the true target, all real variables and inflation return to their steady-state levels earlier than before. The negative humps in the impulse responses are thus deeper but less persistent than before.

After a temporary policy shock in figure 6, the differences between the cases of imperfect and full information are larger than in figure 2. The initial interest rate increase translates into a much larger fall in the perceived inflation target, which leads to lower inflation, a higher real interest rate, and a deeper initial recession. The central bank then quickly reduces the interest rate, and all variables return toward steady state with some overshooting.

In general, when private agents overestimate the volatility of the inflation target, both shocks have larger but less persistent effects on all variables. As private agents' estimate of the inflation target is more sensitive to shocks, actual inflation also responds more to these shocks, translating into larger movements in the real interest rate and the other real variables.

17. The inflation target forecast responds negatively to the observed interest rate, as an interest rate increase signals a decrease in the target.

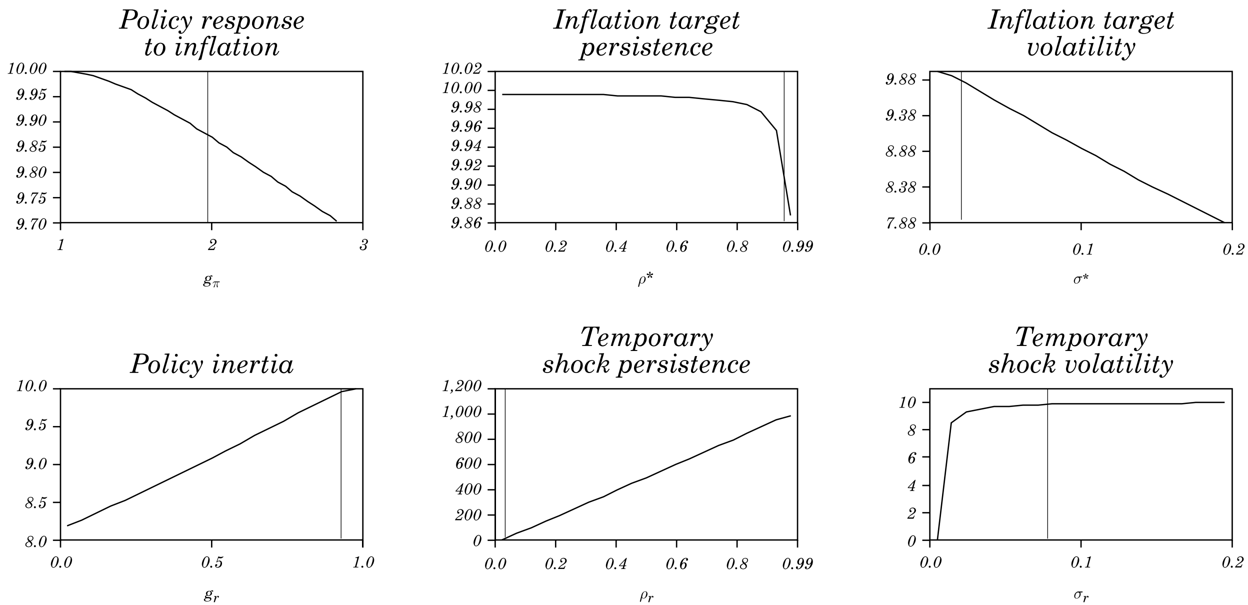
Figure 3. Sensitivity of the Inflation Target Forecast to New Information[a]



Source: Authors' calculations.

a. This figure shows the optimal updating coefficient (namely, the Kalman gain) for the inflation target forecast as key parameters vary from the benchmark calibration. Vertical lines denote benchmark values.

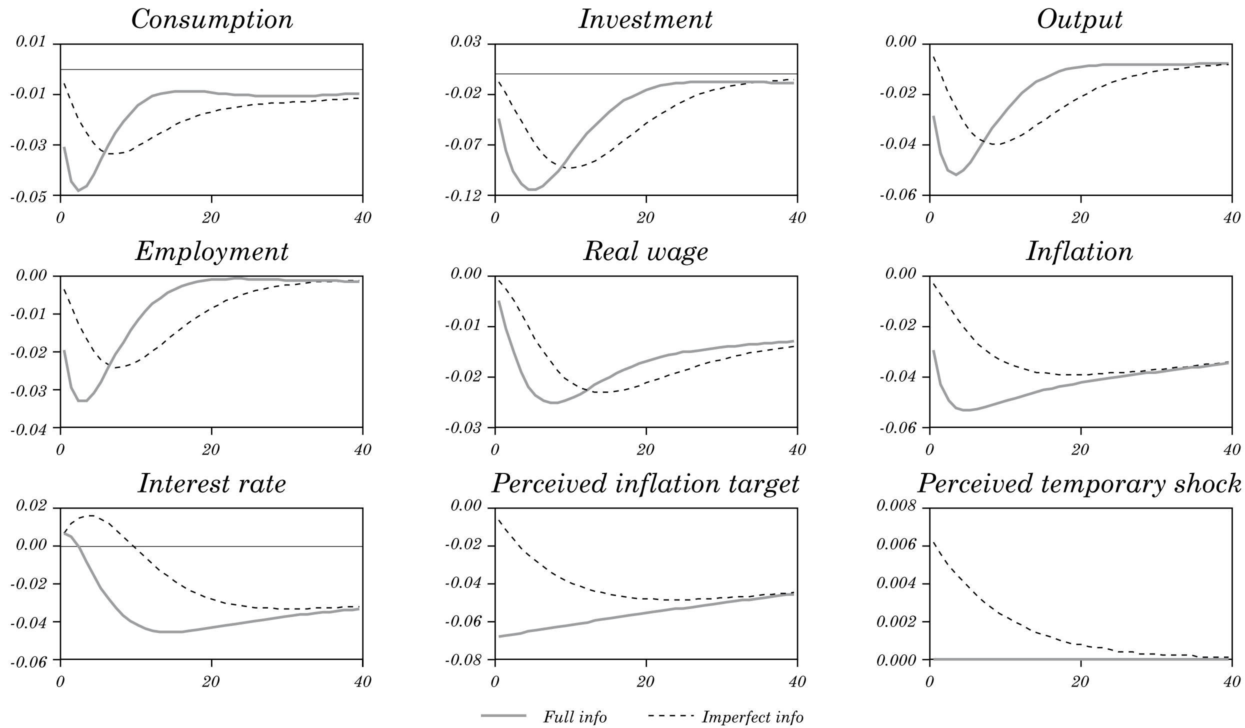
Figure 4. Sensitivity of the Temporary Policy Shock Forecast to New Information[a]



Source: Authors' calculations.

a. This figure shows the optimal updating coefficient (namely, the Kalman gain, multiplied by 1,000) for the temporary policy shock forecast as key parameters vary from the benchmark calibration. Vertical lines denote benchmark values.

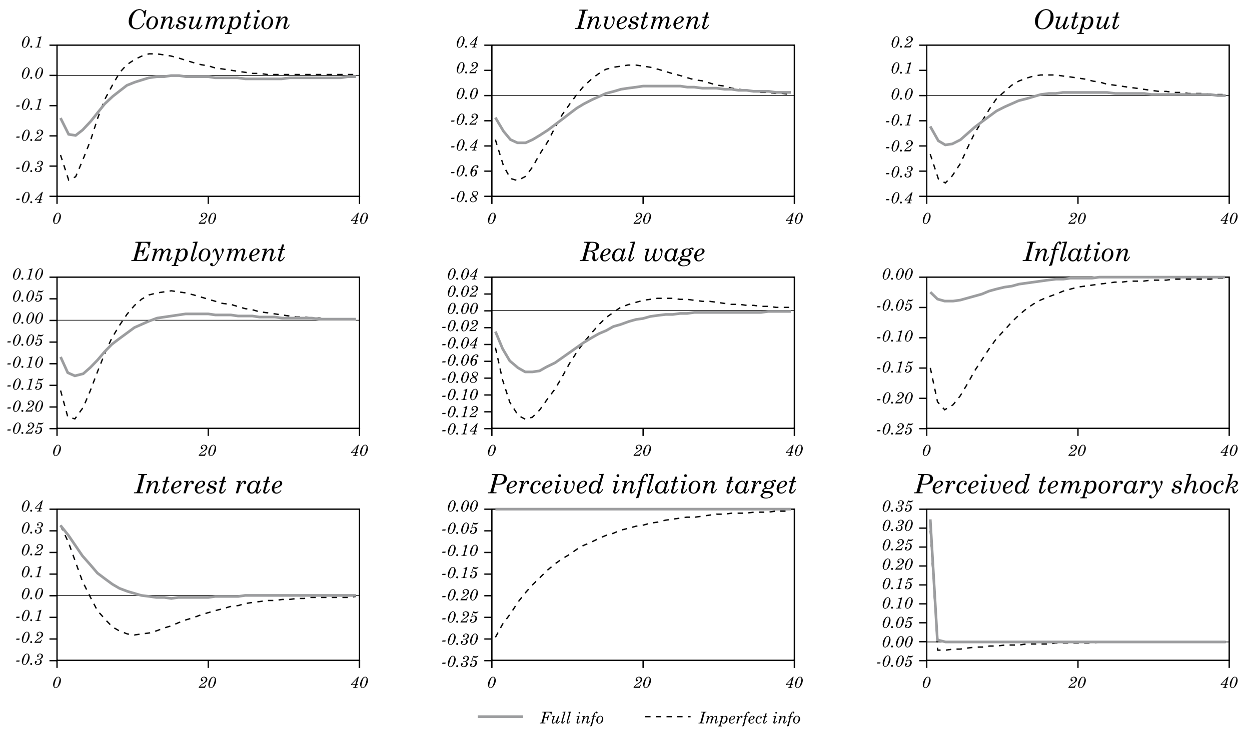
Figure 5. Impulse Responses to an Inflation Target Shock When Private Agents Overestimate the Volatility of the Inflation Target[a]



Source: Authors' calculations.

a. This figure shows impulse responses to a one-standard-deviation negative innovation to the inflation target, π_t^*, when private agents overestimate the volatility of the inflation target: $\hat{\sigma}_* = 5\,\sigma_*$.

Figure 6. Impulse Responses to a Temporary Monetary Policy Shock When Private Agents Overestimate the Volatility of the Inflation Target[a]



Source: Authors' calculations.

a. This figure shows impulse responses to a one-standard-deviation innovation to the temporary monetary policy shock, ε_t^r, when private agents overestimate the volatility of the inflation target: $\hat{\sigma}_* = 5\,\sigma_*$.

Table 3 shows that all variables are now considerably more volatile than with full information. This is particularly the case for inflation, the output gap, and the interest rate, but the variances of the real variables also increase by around five percent relative to the full information case. Thus, when we allow for imperfect information not only on the shocks to the monetary policy rule but also on the variance of these shocks, our model is able to generate fairly large effects of imperfect information on macroeconomic volatility. As a consequence, the gains in terms of macroeconomic stability from announcing the central bank's inflation target are reasonably large.

3. Optimized Monetary Policy Rules and Imperfect Credibility

We now study the properties of optimized rules for monetary policy within our framework. We assume that the central bank aims to stabilize inflation around the inflation target, the output gap, and the interest rate by minimizing the following loss function:

$$L_t = \mathrm{var}\,(\bar{\pi}_t - \bar{\pi}_t^*) + \lambda_y\,\mathrm{var}\,(Y_t - Y_t^n) + \lambda_r\,\mathrm{var}\,(\bar{R}_t), \tag{19}$$

where $\bar{\pi}_t$, $\bar{\pi}_t^*$, and $\bar{R}_t$ measure inflation, the inflation target, and the nominal interest rate in annualized terms, so, for example, $\bar{\pi}_t \equiv 4\pi_t$. While this objective function does not represent the welfare of a representative household in our economy, it is consistent with the mandates of most central banks.[18] We assume that the central bank preference parameters are given by $\lambda_y = 0.5$ and $\lambda_r = 0.1$, so the central bank attaches a larger weight to inflation stability than to output gap stability, and a small weight to interest rate stability.[19]

18. A proper welfare analysis would use an approximation of the representative household's utility as the central bank loss function (see, for instance, Woodford, 2003). In this case, the assumptions concerning firms' price setting would have a direct impact on the welfare criterion. If, as in our model, prices are indexed only to past inflation, the inflation target does not directly affect private sector behavior, and the utility-based loss function would not depend on the volatility of the inflation target. If prices were indexed to the (perceived) inflation target, changes in the target would have direct welfare effects.

19. The interest rate stabilization objective can be seen as a proxy for stability in financial markets. For instance, Tinsley (1999) argues that interest rate volatility may increase term premiums and therefore lead to higher long-term interest rates. From a theoretical perspective, Woodford (2003) shows that the welfare-maximizing policy should aim at reducing interest rate volatility when there are money transaction frictions or when the central bank wants to avoid the zero lower bound of nominal interest rates.

Table 3. Variances of Simulated Data When Private Agents Overestimate the Volatility of the Inflation Target[a]

Type of information	C_t	Y_t	I_t	L_t	W_t	$\overline{\pi}_t$	$Y_t - Y_t^n$	$\overline{R}_t$	$\overline{\pi}_t - \overline{\pi}_t^*$
A. Monetary policy shocks only									
Full information	0.21	0.24	1.15	0.094	0.068	0.14	0.24	0.42	0.025
Imperfect information	0.52	0.64	3.26	0.270	0.140	0.43	0.64	0.61	0.360
B. All shocks									
Full information	6.89	7.12	77.23	3.54	1.60	1.34	3.76	1.29	1.22
Imperfect information	7.19	7.51	79.34	3.72	1.68	1.62	4.15	1.48	1.55

Source: Authors' calculations.

a. This table reports simulated variances (averages over 1,000 simulated series of 10,000 observations) in the models with full information and with imperfect information when private agents overestimate the volatility of the inflation target: $\hat{\sigma}_* = 5\sigma_*$. Inflation and the interest rate are in annualized terms: $\overline{\pi}_t = 4\pi_t$ and $\overline{R}_t = 4R_t$.

We first choose the coefficients in the central bank's policy rule (11) to minimize the central bank loss function when private agents have perfect information about the inflation target and the temporary monetary policy shock.[20] We then evaluate this optimized rule in the case of imperfect information concerning the inflation target. Finally, we discuss whether deviating from the optimized rule may improve the outcome of monetary policy when private agents do not have full information about the inflation target.

The coefficients that minimize the value of the loss function (19) in the case of full information are given by $g_\pi = 10.740$, $g_y = 2.159$, and $g_r = 0.958$. Panel A of table 4 reports the outcome for the three alternative models under this rule, along with the value of the loss function (19). For comparison, panel B reports the corresponding results for the calibrated rule analyzed in section 2. Relative to typical parameterizations of monetary policy rules (and the calibrated rule used earlier), the optimized rule responds more aggressively to both inflation and the output gap and is also slightly more inertial.[21] Comparing the first rows of panels A and B shows that this more aggressive rule is considerably more efficient than the calibrated rule in stabilizing the output gap, at the cost of higher volatility in inflation around the target and the interest rate.

We then implement the rule optimized for the full information model in the models with imperfect information. Panel A of table 4 shows that the presence of imperfect information (when agents know the true variance of the inflation target) leads to modest increases in the volatility of the real variables, as well as the output gap and inflation around the target. Thus, the value of the loss function is only slightly higher than with full information: the increase in loss when moving from full information to imperfect information is equivalent

20. When optimizing the policy rule coefficients, we retain the temporary shocks to the policy rule, even if they are suboptimal. This allows us to compare with the case of imperfect information, where the temporary shocks are necessary to generate a nontrivial learning problem.

21. It is not uncommon for optimized policy rules to be more aggressive than estimated rules. This result is often attributed to the fact that the optimized rules do not take into account different sources of uncertainty that may make policy more cautions. See, for instance, Rudebusch (2001) or Cateau (2007).

Table 4. Performance of Optimized and Calibrated Monetary Policy Rules[a]

	Simulated variances									
Type of policy rule and information	C_t	Y_t	I_t	L_t	W_t	$\bar{\pi}_t$	$Y_t - Y_t^n$	$\bar{R}_t$	$\bar{\pi}_t - \bar{\pi}_t^*$	*Loss*
A. Optimized rule										
Full information	7.86	9.17	92.93	3.95	1.62	1.56	1.67	3.15	1.43	2.580
Imperfect information, $\hat{\sigma}_* = \sigma_*$	7.89	9.20	93.05	3.97	1.63	1.54	1.70	3.14	1.47	2.639
Imperfect information, $\hat{\sigma}_* = 5\sigma_*$	7.94	9.23	93.13	3.98	1.63	1.61	1.73	3.15	1.49	2.677
B. Calibrated rule										
Full information	6.89	7.12	77.23	3.54	1.60	1.34	3.76	1.29	1.22	3.238
Imperfect information, $\hat{\sigma}_* = \sigma_*$	6.94	7.18	77.51	3.57	1.61	1.29	3.82	1.22	1.34	3.380
Imperfect information, $\hat{\sigma}_* = 5\sigma_*$	7.19	7.51	79.34	3.72	1.68	1.62	4.15	1.48	1.55	3.785

Source: Authors' calculations.

a. This table reports simulated variances (averages over 1,000 simulated series of 10,000 observations) in the models with full information and with imperfect information. The optimized rule is the parameterization of the policy rule (11) that minimizes the loss function (19) with $\lambda_y = 0.5$ and $\lambda_r = 0.1$ under full information, and is given by $g_\pi = 10.740$, $g_y = 2.159$, and $g_r = 0.958$. The calibrated rule is given by $g_\pi = 2.0$, $g_y = 0.2$, and $g_r = 0.9$.

to a permanent deviation of inflation from the target of 0.02 percent.[22] Assuming that private agents also overestimate the variance of the inflation target leads to a further increase in volatility and loss, but again the effects are modest: the difference relative to the full information case is now equivalent to a permanent inflation gap of 0.03 percent. A comparison with the calibrated rule in panel B reveals, however, that the central bank is able to substantially reduce the effects of imperfect information by optimizing the policy rule. Under the calibrated rule, the presence of imperfect information is equivalent to a permanent inflation gap of 0.34 and 0.45 percent, respectively, for the two specifications of imperfect information.[23]

To analyze the effects of imperfect information on the optimized policy rule, we study the performance of six alternative rules, where we let one policy rule coefficient at a time deviate from the optimized rule by 10 percent while keeping the remaining coefficients at their optimized levels.[24] The results are reported in table 5. By construction, any deviations from the optimized rule will increase loss in the full information model, but panel A of the table shows that the effects of deviating from the optimized coefficients on inflation or the output gap are very small. It is more costly to deviate from the optimized coefficient on the lagged interest rate: reducing the interest rate coefficient by 10 percent increases loss substantially, and increasing the coefficient to 0.99 has an even stronger effect.[25]

Panel B shows the results for the model in which private agents have imperfect information, but know the true variance of the inflation target. Now, deviations from the optimized rule do not necessarily increase loss, as the rule is optimized for the full information model.

22. To see this, consider the quadratic version of the loss function (19) given by

$$L_t = (1-\hat{\beta})E_t\sum_{j=0}^{\infty}\hat{\beta}^j\left[(\overline{\pi}_{t+j}-\overline{\pi}^*_{t+j})^2+\lambda_y\,(Y_{t+j}-Y^n_{t+j})^2+\lambda_r\overline{R}^2_{t+j}\right],$$

which approaches the specification in equation (19) as the central bank discount factor $\hat{\beta}$ approaches one. A permanent inflation gap of x percent then implies a value of the loss function of $(1-\hat{\beta})\sum_{j=0}^{\infty}\hat{\beta}^j x^2 = x^2$. If we denote the loss under full information as L_0 and the loss under imperfect information as L_1, the permanent inflation gap that would be equivalent to moving from full information to imperfect information is given by $x = \sqrt{L_1} - \sqrt{L_0}$.

23. A similar result is obtained by Orphanides and Williams (2007).

24. The coefficient of the lagged interest rate is not allowed to be larger than 0.99.

25. One reason for the large costs of deviating from the optimized degree of policy inertia is that the long-term responses to inflation and the output gap (given by g_π and g_y) are kept unchanged in this exercise. Therefore, adjusting the coefficient on the lagged interest rate also affects the short-term responses to inflation and output, given by $(1-g_r)g_\pi$ and $(1-g_r)g_y$.

Table 5. Performance of Alternative Monetary Policy Rules[a]

Type of policy rule	*Simulated variances* $\bar{\pi}_t$	$Y_t - Y_t^n$	$\bar{R}_t$	$\bar{\pi}_t - \bar{\pi}_t^*$	*Loss*
A. Full information					
Optimized rule	1.56	1.67	3.15	1.43	2.580
Large g_π	1.51	1.76	3.32	1.37	2.586
Small g_π	1.62	1.57	2.98	1.50	2.588
Large g_y	1.61	1.54	3.26	1.48	2.585
Small g_y	1.51	1.82	3.04	1.37	2.586
Large g_r	1.66	3.10	1.09	1.53	3.196
Small g_r	1.55	1.32	8.86	1.42	2.966
B. Imperfect information, $\hat{\sigma}_* = \sigma_*$					
Optimized rule	1.54	1.70	3.14	1.47	2.639
Large g_π	1.49	1.80	3.32	1.41	2.642
Small g_π	1.60	1.61	2.98	1.54	2.648
Large g_y	1.59	1.57	3.25	1.52	2.640
Small g_y	1.49	1.86	3.03	1.41	2.647
Large g_r	1.63	3.26	1.02	1.65	3.389
Small g_r	1.54	1.33	8.91	1.43	2.988
C. Imperfect information, $\hat{\sigma}_* = 5\sigma_*$					
Optimized rule	1.61	1.73	3.15	1.49	2.677
Large g_π	1.56	1.83	3.31	1.43	2.673
Small g_π	1.68	1.64	3.00	1.57	2.694
Large g_y	1.66	1.60	3.26	1.54	2.675
Small g_y	1.56	1.89	3.04	1.43	2.689
Large g_r	2.06	3.99	1.27	1.98	4.099
Small g_r	1.56	1.33	8.85	1.43	2.980

Source: Authors' calculations.

a. This table reports simulated variances (averages over 1,000 simulated series of 10,000 observations) in the models with full information and with imperfect information for different parameterizations of the monetary policy rule (11). The optimized rule is the parameterization that minimizes the loss function (19) with $\lambda_y = 0.5$ and $\lambda_r = 0.1$ under full information, and it is given by $g_\pi = 10.740$, $g_y = 2.159$, and $g_r = 0.958$. Large and small coefficients are 10 percent larger or smaller than the optimized coefficients, with the exception of the large g_r, which equals 0.99.

Nevertheless, all deviations from the optimized rule increase loss, and the results are similar to the case of full information.

Finally, panel C shows the results when agents have imperfect information about the monetary policy shocks and overestimate the variance of the inflation target. In this case, the central bank is better off responding more aggressively to inflation or the output gap than under full information (although the gains are very small). As before, a large coefficient on the lagged interest rate is detrimental to central bank loss, even more so than in the other two cases. The reported variances show that responding more aggressively to inflation implies that inflation follows the inflation target more closely, at the cost of small increases in output and interest rate volatility. When private agents overestimate the volatility of the inflation target under imperfect information, the inflation gap is more volatile than under full information. By responding more aggressively to the inflation deviation from target, the central bank helps private agents learn the inflation target more quickly (see figure 3), which tends to reduce overall volatility.[26] The aggressive policy rule is not a perfect substitute for announcing the inflation target, however: moving from imperfect information to full information would reduce the value of the loss function considerably more than responding more aggressively to inflation.

4. Concluding Remarks

The aim of this paper was to measure the effects of monetary policy transparency and credibility on macroeconomic volatility and welfare. To this end, we use an estimated DSGE model of the euro area economy in which private agents are unable to distinguish between persistent movements in the central bank's inflation target and temporary deviations from the monetary policy rule.

Our model implies that the macroeconomic benefits of credibly announcing the current level of the time-varying inflation target are reasonably small as long as private agents correctly understand the stochastic processes governing the inflation target and the temporary policy shock. While economic volatility decreases substantially after shocks to monetary policy, these shocks account for a small fraction of

26. Similar results are obtained by Molnár and Santoro (2006) and Orphanides and Williams (2007) in models in which private agents learn about the processes for inflation, output (or unemployment), and the interest rate.

overall volatility in the economy. The overall gains from announcing the time-varying inflation target are therefore fairly small. However, if private agents overestimate the volatility of the inflation target, the overall gains of announcing the target can be substantial.

We have also demonstrated that the central bank to some extent can help private agents in their learning process by responding more aggressively to inflation. If we assume a standard objective function for monetary policy, our results suggest that the optimal response to inflation is more aggressive when private agents have imperfect information and overestimate the volatility of the inflation target than when private agents have full information.

Since our model is derived from the optimizing behavior of private agents, our framework can also be used to study the welfare effects of imperfect monetary policy credibility and transparency, for instance, using a linear-quadratic approximation of welfare in our model, following Benigno and Woodford (2003) and Altissimo, Cúrdia, and Rodríguez Palenzuela (2005). We plan to pursue this avenue in future work.

Appendix

Simulating the Model with Learning

The solution of the model is given by

$$\mathbf{z}_t = \mathbf{A}\,\mathbf{z}_{t-1} + \mathbf{B}\boldsymbol{\eta}_t, \tag{A1}$$

where $\mathbf{z}_t$ is a vector that includes the variables in the sticky price/wage model (thirteen equations), the Kalman filter variables $E_t\pi^*_{t+1}$, $E_t\varepsilon^r_{t+1}$, $E_t\pi^*_t$, and $E_t\varepsilon^r_t$ (four equations), the flexible price/wage model (nine equations), and the ten shock processes, including π^*_t and ε^r_t, while $\boldsymbol{\eta}_t$ is a vector that includes the ten innovations.

Under imperfect information, the shocks to the inflation target (η^*_t) and the monetary policy rule (η^r_t) are not directly observable by private agents. Instead, in each period *t,* private agents observe the interest rate R_t, use the Kalman filter to update their estimates of π^*_t and ε^r_t, and then adjust their expectations of future monetary policy, inflation, and output accordingly. As time passes, the observed interest rate differs from agents' expectations, so agents continue to update their information and adjust their expectations. To capture this process we feed in the change in agents' estimates of π^*_t and ε^r_t as new "shocks" in each period by calculating

$$\begin{aligned}
\begin{bmatrix}\hat{E}_t\eta^*_t\\ \hat{E}_t\eta^r_t\end{bmatrix} &= \begin{bmatrix}\hat{E}_t\pi^*_t\\ \hat{E}_t\varepsilon^r_t\end{bmatrix} - \begin{bmatrix}\hat{E}_{t-1}\pi^*_t\\ \hat{E}_{t-1}\varepsilon^r_t\end{bmatrix}\\
&= \mathbf{F}^{-1}\begin{bmatrix}\hat{E}_t\pi^*_{t+1}\\ \hat{E}_t\varepsilon^r_{t+1}\end{bmatrix} - \begin{bmatrix}\hat{E}_{t-1}\pi^*_t\\ \hat{E}_{t-1}\varepsilon^r_t\end{bmatrix}\\
&= \left[\mathbf{F}^{-1}(\mathbf{F}-\boldsymbol{\kappa}\mathbf{H}')-\mathbf{I}\right]\begin{bmatrix}\hat{E}_{t-1}\pi^*_t\\ \hat{E}_{t-1}\varepsilon^r_t\end{bmatrix} + \mathbf{F}^{-1}\boldsymbol{\kappa}\mathbf{H}'\begin{bmatrix}\pi^*_t\\ \varepsilon^r_t\end{bmatrix},
\end{aligned} \tag{A2}$$

and we add the shocks $E_t\eta^*_t$ and $E_t\eta^r_t$ in the innovation vector $\boldsymbol{\eta}_t$, and the forecasts $E_t\pi^*_t$ and $E_t\varepsilon^r_t$ among the shock processes in the vector $\mathbf{z}_t$. (These $E_t\pi^*_t$ and $E_t\varepsilon^r_t$ coincide with those from the Kalman filter.) This gives a total of twenty-six endogenous variables, twelve autoregressive shocks in the vector $\mathbf{z}_t$, and twelve innovations in the vector $\boldsymbol{\eta}_t$.

Finally, we need to modify the model solution (A1) to take into account the effect of learning on the endogenous variables: while the central bank responds to the true π^*_t and ε^r_t, private agents respond to

$E_t\pi_t^*$ and $E_t\varepsilon_t^r$. We do this by reshuffling the matrices **A** and **B** so that the columns corresponding to π_t^*, ε_t^r, η_t^*, and η_t^r in the private sector equations (all equations except the interest rate rule) are moved to the positions of $E_t\pi_t^*$, $E_t\varepsilon_t^r$, $E_t\eta_t^*$, and $E_t\eta_t^r$. Simulating the model with the learning shocks described above then gives the evolution of the economy.

References

Altissimo, F., V. Cúrdia and D. Rodríguez Palenzuela. 2005. "Linear-Quadratic Approximation to Optimal Policy: An Algorithm and Two Applications." Frankfurt: European Central Bank.

Andolfatto, D., S. Hendry and K. Moran. 2005. "Are Inflation Expectations Rational?" EconWPA macroeconomics paper 0501002. Simon Fraser University.

Aoki, K. and T. Kimura. 2007. "Uncertainty about Perceived Inflation Target and Monetary Policy." Discussion paper series 1: economic studies 2007-18. Frankfurt: Deutsche Bundesbank.

Beechey, M. 2004. "Excess Sensitivity and Volatility of Long Interest Rates: The Role of Limited Information in Bond Markets." Working paper 173. Stockholm: Sveriges Riksbank.

Beechey, M. and P. Österholm. 2007. "The Rise and Fall of U.S. Inflation Persistence." Finance and economics discussion paper 2007-26. Washington: Board of Governors of the Federal Reserve System.

Benati, L. 2006. "Investigating Inflation Persistence across Monetary Regimes." Working paper 851. Frankfurt: European Central Bank.

Benigno, P. and M. Woodford. 2003. "Optimal Monetary and Fiscal Policy: A Linear-Quadratic Approach." In *NBER Macroeconomics Annual*, edited by M. Gertler and K. Rogoff. MIT Press.

Bernanke, B.S., T. Laubach, F.S. Mishkin and A.S. Posen. 1999. *Inflation Targeting: Lessons from the International Experience*. Princeton University Press.

Calvo, G.A. 1983. "Staggered Prices in a Utility-Maximizing Framework." *Journal of Monetary Economics*, 12(3): 383–98.

Cateau, G. 2007. "Monetary Policy under Model and Data-Parameter Uncertainty." *Journal of Monetary Economics* 54(7): 2083–101.

Cecchetti, S.G. and M. Ehrmann. 1999. "Does Inflation Targeting Increase Output Volatility? An International Comparison of Policymakers' Preferences and Outcomes." Working paper 7426. Cambridge, Mass.: National Bureau of Economic Research.

Christiano, L.J., M. Eichenbaum and C.L. Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113(1): 1–45.

Clarida, R., J. Galí, and M. Gertler. 1999. "The Science of Monetary Policy: A New Keynesian Perspective." *Journal of Economic Literature* 37(4): 1661–707.

Cogley, T., G.E. Primiceri and T.J. Sargent. 2008. "Inflation-Gap Persistence in the U.S." Working paper 13749. Cambridge, Mass.: National Bureau of Economic Research.

Del Negro, M., F. Schorfheide, F. Smets and R. Wouters. 2005. "On the Fit and Forecasting Performance of New Keynesian Models." Working paper 491. Frankfurt: European Central Bank.

Dennis, R. 2006. "The Policy Preferences of the U.S. Federal Reserve." *Journal of Applied Econometrics* 21(1): 55–77.

Ellingsen, T. and U. Söderström. 2001. "Monetary Policy and Market Interest Rates." *American Economic Review* 91(5): 1594–607.

————. 2005. "Why Are Long Rates Sensitive to Monetary Policy?" Working paper 256. Milan: Bocconi University, IGIER.

Erceg, C.J. and A.T. Levin. 2003. "Imperfect Credibility and Inflation Persistence." *Journal of Monetary Economics* 50(4): 915–44.

Gaspar, V., F. Smets and D. Vestin. 2006a. "Adaptive Learning, Persistence, and Optimal Monetary Policy." Working paper 644. Frankfurt: European Central Bank.

————. 2006b. "Monetary Policy over Time." *Macroeconomic Dynamics* 10(2): 207–29.

————. 2006c. "Optimal Monetary Policy under Adaptive Learning." *Computing in Economics and Finance* 183.

Gürkaynak, R.S., A.T. Levin, A.N. Marder and E.T. Swanson. 2007. "Inflation Targeting and the Anchoring of Inflation Expectations in the Western Hemisphere." In *Monetary Policy under Inflation Targeting*, edited by F. S. Mishkin and K. Schmidt-Hebbel. Santiago: Central Bank of Chile.

Gürkaynak, R.S., A.T. Levin and E.T. Swanson. 2006. "Does Inflation Targeting Anchor Long-Run Inflation Expectations? Evidence from long-Term Bond Yields in the U.S., U.K., and Sweden." Working paper 2006-09. Federal Reserve Bank of San Francisco.

Gürkaynak, R.S., B. Sack and E. Swanson. 2005. "The Sensitivity of Long-Term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models." *American Economic Review* 95(1): 425–36.

Hamilton, J.D. 1994. *Time Series Analysis*. Princeton University Press.

Ireland, P.N. 2007. "Changes in the Federal Reserve's Inflation Target." *Journal of Money, Credit, and Banking* 39(8): 1851–82.

Kozicki, S. and P.A. Tinsley. 2005. "Permanent and Transitory Policy Shocks in an Empirical Macro Model with Asymmetric

Information." *Journal of Economic Dynamics and Control* 29(11): 1985–2015.

Leiderman, L. and L.E.O. Svensson. 1995. *Inflation Targets*. London: Centre for Economic Policy Research.

Levin, A.T., F.M. Natalucci and J.M. Piger. 2004. "The Macroeconomic Effects of Inflation Targeting." *Federal Reserve Bank of St. Louis Review* 86(4): 51–80.

Levin, A.T., A. Onatski, J.C. Williams and N. Williams. 2005. "Monetary Policy under Uncertainty in Micro-Founded Macroeconometric Models." In *NBER Macroeconomics Annual*, edited by M. Gertler and K. Rogoff. MIT Press.

McCallum, B.T. 1997. Comment on "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy," by J. J. Rotemberg and M. Woodford. In *NBER Macroeconomics Annual*, edited by B. S. Bernanke and J. J. Rotemberg. MIT Press.

Molnár, K. and S. Santoro. 2006. "Optimal Monetary Policy When Agents Are Learning." Computing in Economics and Finance 2006, paper 40. Society for Computational Economics.

Moran, K. 2005. "Learning and the Welfare Implications of Changing Inflation Targets." Working paper 05-11. Université Laval, CIRPÉE.

Nunes, R.C. 2005. "Learning the Inflation Target." Universitat Pompeu Fabra.

Onatski, A. and N. Williams. 2004. "Empirical and Policy Performance of a Forward-Looking Monetary Model." Princeton University.

Orphanides, A. and J.C. Williams. 2007. "Inflation Targeting with Imperfect Knowledge." *Federal Reserve Bank of San Francisco Economic Review* 1–23.

Rudebusch, G.D. 2001. "Is the Fed Too Timid? Monetary Policy in an Uncertain World." *Review of Economics and Statistics* 83(2): 203–17.

————. 2002. "Term Structure Evidence on Interest Rate Smoothing and Monetary Policy Inertia." *Journal of Monetary Economics* 49(6): 1161–87.

Rudebusch, G.D. and J.C. Williams. 2008. "Revealing the Secrets of the Temple: The Value of Publishing Central Bank Interest Rate Projections." In Campbell, John Y. (ed.), *Asset Prices and Monetary Policy*, edited by J. Y. Campbell. University of Chicago Press.

Smets, F. and R. Wouters. 2003. "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area." *Journal of the European Economic Association* 1(5): 1123–75.

———. 2005. "Comparing Shocks and Frictions in U.S. and Euro Area Business Cycles: A Bayesian DSGE Approach." *Journal of Applied Econometrics* 20(2), 161–83.

———. 2007. "Shocks and Frictions in U.S. Business Cycles: A Bayesian DSGE Approach." *American Economic Review* 9(3): 586–606.

Svensson, L.E.O. and M. Woodford. 2004. "Implementing optimal policy through inflation-forecast targeting." In *The Inflation-Targeting Debate*, edited by B. S. Bernanke and M. Woodford. University of Chicago Press.

Tinsley, P.A. 1999. "Short Rate Expectations, Term Premiums, and Central Bank Use of Derivatives to Reduce Policy Uncertainty." Finance and economics discussion paper 99-14. Washington: Board of Governors of the Federal Reserve System.

Woodford, M. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Learning, Endogenous Indexation, and Disinflation in the New-Keynesian Model

Volker Wieland
Goethe University of Frankfurt

Developing a better understanding of the costs of disinflation has long been an important objective for macroeconomic research. Since the 1980s, disinflation episodes and strategies have been studied extensively under the assumption of rational expectations. This assumption implies that central bank announcements regarding future policy plans can help achieve disinflation at little or no cost in terms of lost output in spite of the presence of price level rigidity. Many researchers consider this prediction too optimistic in light of historical experience. Thus, most models used for policy analysis today combine the rational expectations assumption with additional frictions that increase the cost of disinflation, such as exogenous backward-looking indexation of wages and producer prices.

The success of many inflation-targeting countries in lowering inflation in the 1990s provides a new set of case studies that can improve our understanding of inflation-output tradeoffs and serve as a testing ground for macroeconomic modeling. These experiences can serve as the basis for evaluating departures from the benchmark New-Keynesian model with rational expectations and exogenous indexation and investigate the desirability of alternative policy strategies. Chile, which in 1990 became the second country to adopt inflation targeting, constitutes a particularly interesting example as an increasing number of developing economies opt for inflation targeting. The Chilean

I am grateful to the Stanford Center for International Development, where I was a visiting professor while writing this paper. Helpful comments were provided by Klaus Schmidt-Hebbel, Vicente Tuesta, Gernot Müller, Nick Hope, Pete Klenow, John C. Williams, Cliff Tan, and Gustavo Leyva. I also thank Klaus Schmidt-Hebbel, Gustavo Leyva, and Fabian Gredig for providing me with data for Chile.

Monetary Policy under Uncertainty and Learning, edited by Klaus Schmidt-Hebbel and Carl E. Walsh, Santiago, Chile.

disinflation episode stands out as a very gradual disinflation achieved with temporary annual inflation targets.

In light of the Chilean experience, this paper examines the implications of two departures from the benchmark New-Keynesian model. First, I follow the recent literature on adaptive learning and replace the assumption of rational expectations with recursive least squares learning. Second, I introduce endogenous indexation by allowing firms to choose between backward-looking indexation and the central bank's announced target. At the start of the disinflation episode, indexation is complete and price-setters expect highly persistent inflation. As price-setting firms learn over time, they reassess the likelihood of announced inflation targets and adjust indexation of contracts accordingly.

The findings in this paper indicate that learning and endogenous indexation may reduce the costs of disinflation. A gradual disinflation approach can take advantage of these favorable dynamics to achieve the long-run target at lower output costs. An interesting new result is the finding that announcing and meeting annual targets for inflation results in lower disinflation costs relative to the announcement of a long-run inflation target that will only be met after many years of gradual disinflation. Model simulations of the actual targets announced by the Central Bank of Chile during the disinflation from 1990 to 2001 imply rather favorable learning and indexation dynamics.

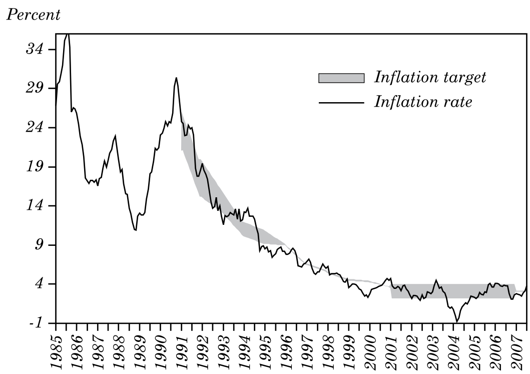
The paper proceeds as follows. Section 1 shortly summarizes several aspects of the Chilean disinflation process and the related literature. Section 2 compares traditional views with the New-Keynesian approach to understanding the costs of disinflation. In section 3, I introduce adaptive learning and endogenous indexation in the New-Keynesian model. Section 4 contrasts immediate and gradual disinflation strategies. In section 5 I formulate different sequences of annual inflation targets and evaluate their performance in implementing disinflation. Section 6 briefly discusses possible approaches to designing dynamically optimal policy, while section 7 concludes.

1. Inflation Targeting and Disinflation: Chile, 1991–2007

Inflation targeting started with public announcements of inflation targets in New Zealand and Chile in 1990. Since then, this monetary policy strategy has been implemented in many economies around the world, including developed countries such as the United Kingdom, Canada, Sweden, Norway, and Australia and an increasing number of developing countries. Many of these developing countries have been able to reduce inflation rates substantially following the adoption of inflation targeting, and they seem to have succeeded in stabilizing inflation at low to moderate rates. Fraga, Goldfajn, and Minella (2003), Corbo and Schmidt-Hebbel (2003), Corbo, Landerretche, and Schmidt-Hebbel (2002), and Mishkin and Schmidt-Hebbel (2001) provide empirical assessments of the performance of inflation targeting in a large number of diverse economies.

Given the increasing popularity of inflation targeting in developing countries, any lessons for policymakers that can be derived from Chile's experience are particularly useful. The Chilean disinflation stands out as a very gradual disinflation. The Central Bank's first official target, which was publicly announced in September 1990, was a range of 15 to 20 percent for annual consumer price index (CPI) inflation between December 1990 and December 1991. From 1991 to 1999, inflation target ranges and point targets were set on an annual basis for the following calendar year. Figure 1 reports the inflation targets (shaded area) along with actual inflation (solid line).

Figure 1. Inflation Targets and Actual Inflation in Chile, 1985-2007



Source: Schmidt-Hebbel and Werner (2002) updated to 2007 by Klaus Schmidt-Hebbel, Gustavo Leyva, and Fabian Gredig from the Central Bank of Chile.

Initially, many observers were skeptical about the importance of the Chilean Central Bank's strategic framework in achieving disinflation. They attributed much of the improvement to good luck in the form of exogenous developments concerning the exchange rate and raw material prices. Calvo and Mendoza (1999), for example, wrote that "factors other than stabilization policies have played an important role in Chilean economic performance, and the dynamics exhibited by key macroeconomic aggregates can be interpreted in part as an endogenous process of adjustment triggered by exogenous shocks." However, the amazing success of the Central Bank of Chile in meeting its annual inflation targets during the disinflation phase from 1990 to 2001 and its continued ability to keep inflation close to or within the target zone of 2 to 4 percent suggest that its strategic framework played an important role.

Aguirre and Schmidt-Hebbel (2007) argue that the short-term annual targets announced during the disinflation phase were observationally equivalent to hard policy targets in full-fledged inflation-targeting regimes. They provide some evidence in favor of this view. In spite of low initial policy credibility and widespread backward-looking price indexation in goods, labor, and financial markets, disinflation was achieved at relatively low costs in terms of associated output losses. Aguirre and Schmidt-Hebbel (2007) suggest that the Central Bank was able to overcome the consequences of backward-looking price indexation and related inflation inertia, and to influence private sector inflation expectations, by pursuing a forward-looking inflation target that served as an explicit nominal anchor. Similarly, Corbo, Landerretche, and Schmidt-Hebbel (2002) draw three main lessons from Chile's experience that should be of interest to other developing countries:

> First, initial progress in reducing inflation toward the target was slow as the public was learning about the Central Bank's true commitment to attaining the target. Second, the gradual phasing in of inflation targeting contributed to declining inflation by lowering inflation expectations and changing wage and price dynamics. Third, with respect to the speed of inflation reduction, a cold-turkey approach would have resulted in a larger sacrifice ratio stemming from higher unemployment during the early years of inflation targeting, when credibility was gradually being built up.

These conclusions suggest that learning by price-setting firms and changes in the degree of backward-looking indexation regarding wages and producer prices played an important role in shaping the costs of disinflation in Chile.[1]

More recently, researchers have developed and estimated sophisticated New-Keynesian dynamic general equilibrium models for policy analysis in Chile.[2] These models share the assumptions of rational expectations and exogenous backward-looking indexation with similar models developed for industrialized economies (see Christiano, Eichenbaum, and Evans, 2005). The New-Keynesian Phillips curve embedded in these models, however, does not seem to be stable. For example, Céspedes, Ochoa, and Soto (2005) report evidence of structural change in the late 1990s. This change is exhibited in a higher weight of expected future inflation—and a correspondingly lower weight of lagged inflation—when producers set their prices. For a sample from 1990 to 2000, they estimate a degree of backward-looking indexation around 0.85, which is essentially indistinguishable from the limiting case of complete indexation. With the sample extended to 2005, however, the degree of indexation declines to around 0.66.

The remainder of this paper explores departures from the standard New-Keynesian model by allowing for adaptive learning and endogenous indexation. I investigate whether the particular choice of inflation-targeting strategy may influence the costs of disinflation by increasing the speed of learning and reducing the degree of backward-looking indexation.

2. Disinflation and the New-Keynesian Phillips Curve

It is conventional wisdom among central bankers that conducting monetary policy so as to keep inflation constant at all times will induce fluctuations in aggregate real output. Historical experience such as the 1980s Volcker disinflation in the United States suggests that a permanent reduction in the rate of inflation cannot be achieved without a temporary decline of output below the economy's potential. Such a cost of disinflation is embedded in the traditional accelerationist Phillips curve:

$$\pi_t = \pi_{t-1} + \lambda x_t. \tag{1}$$

1. See also Herrera (2002) and Lefort and Schmidt-Hebbel (2002).

2. See Caputo, Liendo, and Medina (2007), Caputo, Medina, and Soto (2006), De Gregorio and Parrado (2006), and Céspedes, Ochoa, and Soto (2005).

Here, π_t denotes the rate of inflation and x_t the output gap (that is, the deviation of actual output from the economy's potential).

A simple experiment serves to illustrate the cost of disinflation. Assume that inflation in period $t = 1$ is equal to 1 percent and that the central bank aims to achieve price stability (that is, an inflation rate of zero percent) in period $t = 2$. Such a reduction in the rate of inflation requires a negative output gap of $-1/\lambda$ percent in period $t = 1$. In the absence of any future shocks that might push inflation up or down, inflation could then be held at zero from period 2 onward by keeping the output gap closed. Thus, the cumulative output loss in absolute terms that is required to achieve a reduction in inflation of 1 percentage point corresponds to $1/\lambda$ percent of total output.

In central bank circles, the cumulative output loss associated with a permanent reduction of the inflation rate by one percent is often referred to as the sacrifice ratio. If equation (1), the accelerationist Phillips curve, is treated as a structural relationship, then the associated sacrifice ratio is constant at $1/\lambda$ and invariant to policy design. In other words, no particular strategy or announcement by the central bank could help in changing the trade-off between output and inflation or in reducing the cumulative output cost of a disinflation. Nevertheless, a central bank that cares about stabilizing output and inflation would always opt for disinflating gradually and spreading the output loss over time.

2.1 The New-Keynesian Perspective on Disinflation

The traditional Phillips curve shown above lacks microeconomic foundations. Fortunately, the New-Keynesian paradigm offers an alternative model of inflation that is consistent with optimizing behavior and rational expectations formation by households and firms. However, the basic version of the New-Keynesian model has a very controversial property. In this model, the macroeconomic policy goals of stabilizing output and inflation do not come into conflict with each other (see Walsh, 2003; Woodford, 2003). This property is often referred to as the divine coincidence. It implies that disinflation can be achieved without any reduction in aggregate output. It is somewhat surprising that a model that incorporates long-lasting nominal rigidities exhibits such a property. To understand its origins, it is helpful to reiterate the elements of the model that drive price-setting and inflation dynamics.

The model is populated by a continuum of monopolistic firms that produce differentiated goods. Importantly, these firms cannot adjust product prices freely in every period. The basic version of the model relies on the mathematically convenient mechanism for modeling price ridigity, as introduced by Calvo (1983). It implies that firms have to wait for a signal to adjust prices. They receive such a signal with probability $1 - \theta$. Every firm that receives a price-setting signal solves a dynamic optimization problem to set its price optimally, taking into account the probabilistic constraint on future price-setting opportunities. A firm j that does not receive a price-setting signal leaves its price unchanged at the zero inflation steady state. Alternatively, if the steady-state rate of inflation, π^S, differs from zero, firm j lets its price grow with that steady-state rate, that is, $P_{j,t} = (1 + \pi^S)P_{j,t-1}$. In other words, firms that are not allowed to reoptimize their price are instead assumed to index to steady-state inflation. In solving their optimization problem, firms are assumed to form rational, model-consistent expectations.

A useful feature of this model is that it can be solved without explicitly tracking the distribution of prices across firms. Aggregation and log-linear approximation deliver a well-known, simple relationship between inflation, expected future inflation, and the output gap—the New-Keynesian Phillips curve:

$$\pi_t - \pi^S = \beta E_t[\pi_{t+1} - \pi^S] + \lambda x_t \tag{2}$$

Here, the output gap, x_t, denotes the difference between actual output and the level of output that would be achieved if prices were flexible. The parameter β refers to the discount factor. The slope parameter, λ, is a function of θ and β.[3]

Again, a simple experiment serves to assess the cost of disinflation. Suppose the central bank enters period $t = 1$ with an inflation target, π^*, equal to 1 percent. Since equation (3) is linear, the steady-state rate of inflation must be equal to the central bank's target, $\pi^S = \pi^*$. In period $t = 2$, the central bank announces a new target rate of zero percent inflation. Market participants would immediately incorporate the new target in their expectations for period $t = 3$. It would imply zero inflation in steady state. As a result, inflation in period $t = 2$

3. To be precise, the baseline version of the model (see Walsh, 2003) implies that λ is determined as follows: $\lambda = (1 - \theta)(1 - \beta\theta)\theta^{-1}(\sigma + \varphi)$. Here, σ^{-1} and φ represent the constant intertemporal elasticity of consumption and labor supply elasticity, respectively.

immediately drops to the new target rate. No reduction in the output gap, x_t, is required to achieve this outcome. Disinflation is costless. It is achieved by influencing market participants' expectations.

The model's implication of costless disinflation stands in contrast to historical experience. For this reason, researchers who have estimated New-Keynesian models using data from leading industrial economies have typically assumed an additional source of price rigidity. One possible approach is to introduce firms that apply a simple rule of thumb in price setting, as in Galí and Gertler (1999). An alternative approach assumes that some firms index prices to past inflation in those periods when they cannot adjust prices optimally (Christiano, Eichenbaum, and Evans, 2005).

Backward-looking indexation has become a popular assumption embedded in many empirically estimated dynamic stochastic general equilibrium (DSGE) models used for monetary policy analysis. Firms that do not receive a Calvo-style signal to adjust prices in the current period are assumed to implement instead a pricing rule based on past inflation, that is, $P_{j,t} = (1 + \pi_{t-1})P_{j,t-1}$. The share of firms that use backward-looking indexation, denoted by κ in the following discussion, is treated as exogenous. Consequently, the log-linear approximation of the New-Keynesian Phillips curve takes the following form:

$$\pi_t - \left(\kappa\pi_{t-1} + (1-\kappa)\pi^S\right) = \beta E_t\left[\pi_{t+1} - \left(\kappa\pi_t + (1-\kappa)\pi^S\right)\right] + \lambda x_t. \tag{3}$$

The current inflation rate then depends on a weighted average of past and expected future inflation. The weight is a function of the share of firms that implement backward-looking indexation:

$$\pi_t = \frac{\kappa}{1+\beta\kappa}\pi_{t-1} + \frac{\beta}{1+\beta\kappa}E_t\left[\pi_{t+1}\right] + \frac{\lambda}{1+\beta\kappa}x_t + \frac{(1-\kappa)(1-\beta)}{1+\beta\kappa}\pi^S. \tag{4}$$

In the limiting case of complete indexation, $\kappa = 1$, the inflation equation simplifies to

$$\pi_t = \frac{1}{1+\beta}\pi_{t-1} + \frac{\beta}{1+\beta}E_t\left[\pi_{t+1}\right] + \frac{\lambda}{1+\beta}x_t. \tag{5}$$

Interestingly, with complete indexation the current inflation rate is independent of steady-state inflation, π^s.

Equation (4) has been estimated for many countries. Estimates for Chile have been obtained by Céspedes, Ochoa, and Soto (2005), Caputo, Medina, and Soto (2006), and Caputo, Liendo, and Medina (2007). Céspedes, Ochoa, and Soto (2005) took care to account for time-variation in the inflation target. In this case, the last term in equation (4) is modified to $(1 - \kappa)(1 + \beta\kappa)^{-1}(\pi_t^* - \beta\pi^*_{t+1})$. As mentioned earlier, they report evidence of structural change. For a sample from 1990 to 2000, they estimate a degree of backward-looking indexation around 0.85, which is essentially indistinguishable from the limiting case of complete indexation. With the sample extended to 2005, however, the degree of indexation declines to around 0.66.

In this paper, I relax two important assumptions of the standard model—namely, the assumption of rational expectations and the assumption of exogenous backward-looking indexation. Relaxing these assumptions is important because of the empirical evidence regarding changes in the degree of inflation persistence during and following the disinflation in Chile. The reduction in inflation persistence may well be due to changes in price setters' beliefs or changes in the degree of backward-looking indexation. I thus depart from the assumption of rational expectations by considering adaptive learning. This follows the lead of Marcet and Sargent (1989), Evans and Honkapohja (2001), Orphanides and Williams (2006a, 2006b), and Gaspar, Smets, and Vestin (2006a, 2006b). A further innovation is rendering the share of firms that implement backward-looking indexation endogenous. In particular, I allow firms to choose between the central bank's inflation target and past inflation as possible indexes. This choice of index is made according to the likelihood that the chosen index better matches the mean of the observed inflation distribution. Firms thus aim to choose the index that seems to provide a better estimate of steady-state inflation.

3. Adaptive Learning and Endogenous Indexation

As shown above, expectations play a key role in determining inflation dynamics. Since the 1980s, research on monetary policy has relied on the assumption of rational expectations and explored its implications for policy design. A drawback of the assumption of rational expectations is that it imputes an unrealistic extent of knowledge to market participants. An interesting alternative approach is adaptive or least-squares learning, which assumes that economic agents behave like econometricians in forming expectations and estimate reduced-form inflation equations. Under certain assumptions, adaptive learning may converge to rational expectations in the long run.

Following the influential contribution by Evans and Honkapohja (2001), Orphanides and Williams (2006a, 2006b) and Gaspar, Smets, and Vestin (2006a, 2006b) have studied monetary policy design with price-setting firms that form their expectations about future inflation in a least-squares fashion. Motivated by this line of research, I assume that price-setting firms estimate the following regression for inflation:

$$\pi_t = \gamma_t \pi_{t-1} + \varepsilon_t. \tag{6}$$

The parameter γ_t carries a time subscript to allow for episodes with high and low degrees of inflation persistence. I make this assumption because the model will endogenously generate a time-varying degree of inflation persistence. Incorporating this time variation in price setters' perceived inflation equation ensures that expectations formation is consistent with equilibrium outcomes. γ_t is believed to follow a random walk with the variance of innovations denoted by σ^γ. Recursive estimation then implies the following updating equations for the price setters' point estimate of the inflation persistence parameter, c_t, and its variance, Σ_t:

$$\begin{aligned} c_t &= c_{t-1} + (\pi_{t-1}^2 \Sigma_{t-1} + \sigma^\varepsilon)^{-1} \Sigma_{t-1} \pi_{t-1} (\pi_t - c_{t-1} \pi_{t-1}), \\ \Sigma_t &= \Sigma_{t-1} - (\pi_{t-1}^2 \Sigma_{t-1} + \sigma^\varepsilon)^{-1} \Sigma_{t-1}^2 \pi_{t-1}^2 + \sigma^\gamma. \end{aligned} \tag{7}$$

For a derivation of these updating equations using the Kalman filter, see Harvey (1992). The updating equations are also consistent with Bayes' rule under the assumption of normally distributed shocks and beliefs (see Zellner, 1971). In the adaptive learning literature, researchers typically choose from a variety of learning specifications. Branch and Evans (2006) provide a useful exposition of alternative approaches and investigate how well they fit survey expectations.

Given equations (6) and (7), the price setters' expectation of future inflation under least-squares learning, $E_t^{LS}\left[\pi_{t+1}\right]$, corresponds to

$$E_t^{LS}\left[\pi_{t+1}\right] = c_{t-1}\,\pi_t. \tag{8}$$

Here, I follow Gaspar, Smets, and Vestin (2006a, 2006b) in assuming that $E_t^{LS}\left[\pi_{t+1}\right]$ is based on the estimate c_{t-1}, which does not yet incorporate the most recent inflation observation, π_t.[4] Using equation (8) to substitute out expected future inflation in equation (4) yields the following reduced-form inflation equation:

$$\pi_t = \frac{\kappa}{1+\beta(\kappa - c_{t-1})}\pi_{t-1} + \frac{\lambda}{1+\beta(\kappa - c_{t-1})}x_t + \frac{(1-\kappa)(1-\beta)}{1+\beta(\kappa - c_{t-1})}\pi^S. \tag{9}$$

Adaptive learning in the form of the time-varying estimate, c_{t-1}, influences the observed degree of inflation persistence. In addition, the degree of persistence depends on central bank policy.

3.1 Introducing Endogenous Indexation

So far, the degree of backward-looking indexation, κ, has been treated as constant and exogenous. A novel contribution of this paper is to allow for an endogenous determination of a time-varying share of firms that apply backward-looking indexation. I assume that firms would like to pick an index that is a good estimate of steady-state inflation. They have two options. One option is the central bank's announced inflation target, π^*. If the central bank delivers on its promise, then steady-state inflation will be equal to the target. The other option is the most recent observation of inflation, π_{t-1}. If the central bank does not aim to control inflation, the inflation rate will follow a random walk, and past inflation will be the best estimate of future inflation.

Every time firms obtain a new observation on inflation, they investigate whether the target or past inflation better matches the mean of the observed inflation distribution. The probability that the announced inflation target corresponds to the mean of the observed inflation distribution is denoted $s_t = \text{Prob}(\pi^S = \pi^*)$. When a new observation becomes available, s_t is updated as follows:

$$s_{t+1} = \frac{s_t e^{-0.5(\pi_t - \pi^*)^2}}{s_t e^{-0.5(\pi_t - \pi^*)^2} + (1-s_t)e^{-0.5(\pi_t - \pi_{t-1})^2}}. \tag{10}$$

4. Alternatively, one could either use only lagged information, that is $E_t^{LS}[\pi_{t+1}] = c_{t-1}^2\,\pi_{t-1}$, or incorporate current inflation in the estimate of the persistence parameter, $E_t^{LS}[\pi_{t+1}] = c_t\pi_t$. The latter specification would require solving a more complicated fixed-point problem.

This updating equation is consistent with Bayes' rule given normal shocks and beliefs.[5]

Firms cannot switch indexes at all times. They are allowed to make a choice regarding the index at the same time as they receive a Calvo-style signal that allows them to adjust their current price optimally. The probability of such a signal is $1 - \theta$. A firm that has received such a signal will then consider whether to switch the index that will apply to its pricing rule in the periods without Calvo signals. One possibility would be to assume that firms switch from backward-looking indexation to the central bank's target as soon as the probability s_t has moved above 0.5 and switch back if this probability falls slightly below 0.5. Such an assumption would seem reasonable in the unlikely case that the index can be switched at zero cost.

Instead, it is assumed that firms only choose to switch the index when there is overwhelming evidence in favor of such a change. Specifically, I introduce a trigger probability, $\bar{S}$. If the firm's current choice of index is π_{t-1}, it will switch to π^* once s_t exceeds $\bar{S}$. Similarly, if the current choice of indexation rate is π^*, the firm will switch back to π_{t-1} if $1 - s_t$ (the probability of π_{t-1}) exceeds the same trigger value. All firms face the same information regarding inflation, so s_t is symmetric across firms. Since the probability of a Calvo signal is $1 - \theta$, a share of $1 - \theta$ firms switches the rate of indexation at any point in time given that there is overwhelming evidence in favor of such a shift.

Finally, the degree of indexation, κ_t, is allowed to vary between complete indexation (that is, $\kappa_t = 1$) and a minimal value of $\underline{\kappa}$ (that is, $\kappa_t \in [\underline{\kappa}, 1]$).[6] Thus, κ_t is governed by the following process:

$$\kappa_t = \begin{cases} \theta\kappa_{t-1} & \text{if } s_t > \bar{S} \text{ and } \kappa_t \geq \underline{\kappa} \\ 1 - \theta(1 - \kappa_{t-1}) & \text{if } (1 - s_t) > \bar{S} \\ \kappa_{t-1} & \text{otherwise} \end{cases} \tag{11}$$

Every period in which s_t exceeds the trigger probability, a share of $1 - \theta$ firms switches from backward-looking indexation to the central bank's target, while a share of θ firms sticks with the past inflation rate.

5. See Wieland (2000a).

6. I maintain a minimal amount of exogenous indexation to ensure that lagged inflation remains a determinant of the equilibrium inflation process under rational expectations. As a result, the learning model uses the correct reduced-form inflation equation under rational expectations.

Since the share of firms using backward-looking indexation varies over time, the reduced-form inflation equation (9) needs to be rewritten as follows:

$$\pi_t = \frac{\kappa_{t-1}}{1+\beta(\kappa_{t-1}-c_{t-1})}\pi_{t-1} + \frac{\lambda}{1+\beta(\kappa_{t-1}-c_{t-1})}x_t + \frac{(1-\kappa_{t-1})(1-\beta)}{1+\beta(\kappa_{t-1}-c_{t-1})}\pi^S. \tag{12}$$

As a short-hand $\delta_{(1,2,3),t}$ denotes the time-varying, reduced-form parameters. Accordingly, the reduced-form inflation equation may be written as

$$\pi_t = \delta_{1,t}\,\pi_{t-1} + \delta_{2,t}\,x_t + \delta_{3,t}. \tag{13}$$

To be able to study disinflation under alternative targeting strategies, it is still necessary to describe the central bank's objectives and the determination of the output gap x_t.

4. Inflation Targeting: Immediate versus Gradual Disinflation

A central bank that has adopted an inflation-targeting strategy is typically assumed to pursue a policy that minimizes the following per-period loss function:

$$l(\pi_t, x_t) = (\pi_t - \pi^*)^2 + \alpha x_t^2. \tag{14}$$

The parameter α refers to the central bank's relative preference for stabilizing output versus inflation.

Two simplifying assumptions keep the technical analysis manageable: the central bank directly controls the output gap, x_t, and it observes the key parameters of the inflation equation as well as the price setters' beliefs regarding inflation persistence, c_{t-1}. Thus, the central bank can take into account the parameters $\delta_{(1,2,3),t}$ of equation (13) in designing its policy. The central bank is not allowed, however, to exploit the dynamic learning process of the price-setters in conducting policy.[7] Under these assumptions, the central bank's dynamic optimization problem corresponds to:

$$\min_{x_t} E_t\left[\sum_{t=1}^{\infty}\beta^{t-1}\left(\pi_t-\pi^*\right)^2+\alpha x_t^2\right], \tag{15}$$

subject to $\pi_t = \delta_{1,t}\,\pi_{t-1} + \delta_{2,t}\,x_t + \delta_{3,t}$.

The extreme cases are strict inflation targeting, $\alpha = 0$, and strict output stabilization, $\alpha \to \infty$. Strict output stabilization would imply that the central bank always aims to set the output gap, x_t, equal to zero. Consequently, the dynamics of inflation would be governed exclusively by the time-varying parameter, $\delta_{1,t}$, which depends in turn on the degree of backward-looking indexation and the price setters' beliefs regarding inflation persistence. If $\delta_{1,t}$ ever exceeded unity, inflation would spiral out of control. In contrast, strict inflation targeting would ensure that the inflation target is met at all times for any perceived degree of inflation persistence. The resulting output gap policy corresponds to

$$x_t = -\delta_{4,t}\left(\delta_{1,t}\pi_{t-1}+\delta_{3,t}-\pi^*\right). \tag{16}$$

with $\delta_{4,t} = \delta_{2,t}^{-1}$. With a zero inflation target, $\delta_{3,t}$ would also be equal to zero.

In the intermediate case, α is positive but not infinite. Such central bank preferences are often called flexible inflation targeting. Under this policy, the output gap falls between the two extremes implied by strict inflation targeting and strict output stabilization, that is, $0 < \delta_{4,t} < \delta_{2,t}^{-1}$. Orphanides and Wieland (2000) provide an analytical formula for the case of $\delta_{1,t} = 1$. Dynamically optimal policies for alternative values of $\delta_{1,t}$ may be computed numerically with the algorithm provided in that paper.[8]

7. I discuss such an ambitious proposal in the last section of the paper. Gaspar, Smets, and Vestin (2006a, 2006b) refer to a central bank with this capability as sophisticated.

8. The matlab code is available from www.volkerwieland.com.

4.1 Model Parameterization and Initial Conditions

Having specified a very stylized but complete macroeconomic model, the next step is to evaluate alternative disinflation strategies. Initial conditions for the disinflation are defined as follows: (i) initial inflation is set at 20 percent, $\pi = 0.2$, similar to the average inflation rate of Chile prior to the start of inflation targeting; (ii) initially all firms implement backward-looking indexation, $\kappa_0 = 1$; and (iii) perceived inflation persistence indicates a unit root in inflation, that is, $c_0 = 1$. Given these initial conditions, the reduced-form inflation equation (13) simplifies to

$$\pi_t = \pi_{t-1} + \lambda x_t, \tag{17}$$

corresponding exactly to equation (1), the accelerationist Phillips curve discussed in section 1. It follows that these initial conditions represent an equilibrium if policy aims exclusively at stabilizing output, that is, if $x_0 = 0$. The parameter values used in the subsequent simulations are summarized in table 1.

Table 1. Parameter Values and Initial Beliefs

Parameter	*Value*	*Economic interpretation*
β	0.99	Discount factor
λ	0.5	Slope of Phillips curve
κ_t	$\kappa_0 = 1$	Degree of indexation to $t - 1$ inflation
c_t	$c_0 = 1$	Price setters' initial belief regarding inflation persistence
Σ_t	$\Sigma_0 = 100$	Price setters' initial variance
s_t	$s_0 = 0.1$	Price/index setters' initial belief regarding prob ($\pi^S = \pi^*$)
π_0, π^*	0.2 / 0	Initial inflation: 0.2; long-run inflation target: 0
$\underline{\kappa}$	0.05	Degree of minimal exogenous indexation
θ	0.5	Probability of no price- or index-adjustment signal
$\bar{S}$	0.8	Trigger probability for switching the rate for indexation
σ	2^{-4}	Variance of noise (added later)
σ_γ	10	Belief regarding variability of γ

Source: Author's calculations.

4.2 Immediate versus Gradual Disinflation

The initial conditions summarized above set the stage for the entry of an independent inflation-targeting central bank.[9] This central bank faces very high initial costs of disinflation. The analysis starts by contrasting the immediate disinflation approach that would be implemented under strict inflation targeting with a more gradual approach consistent with a positive weight on output in the central bank's preferences.

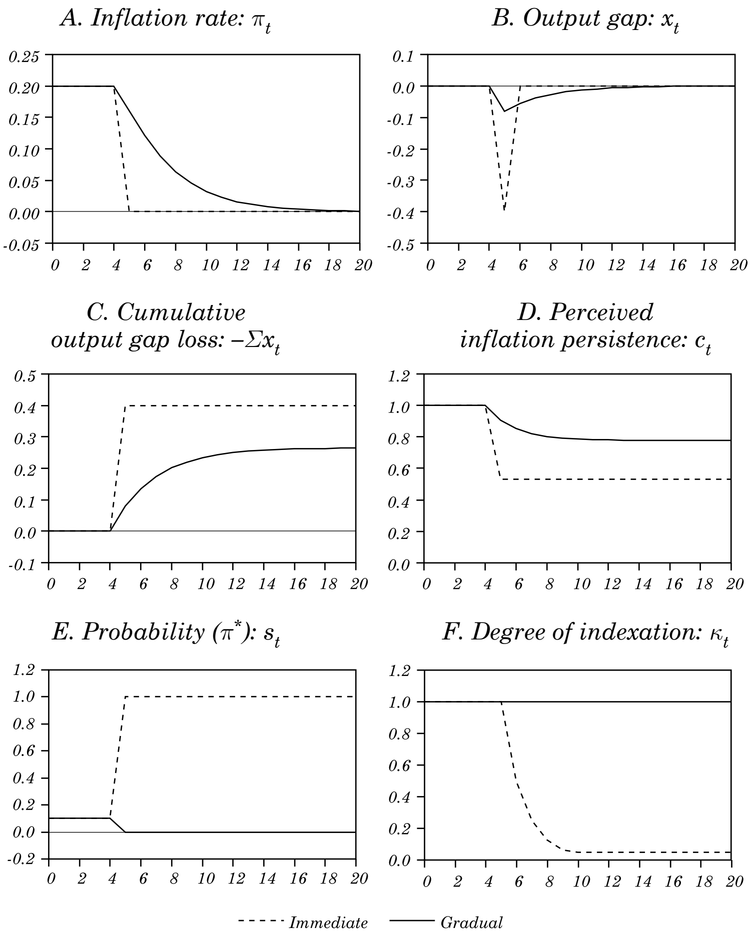
The optimal policy coefficient under strict inflation targeting corresponds to the inverse of the slope of the reduced-form inflation equation and equals $\delta_{4,0} = \delta_{2,0}^{-1} = 2$. In the model, this policy would achieve the inflation target of zero percent within one period but, such an immediate disinflation would result in an output loss of 40 percent in the same period. This outcome is shown by the dotted line in figure 2. In period 5, the central bank introduces a new inflation target of zero percent. The cumulative output loss required to disinflate by 20 percentage points is also realized in period 5. While this approach can be simulated in this simple model, such an immense reduction of total output would not be implementable in practice.

The dramatic experience of immediate disinflation induces price setters to revise their estimates of the inflation persistence parameter, c_t, from 1.0 to about 0.5 (panel D). Furthermore, the probability s_t, which is initially set at 0.1, jumps to 1.0. In other words, the immediate reduction in inflation convinces firms that the central bank's inflation target constitutes a better estimate of the mean of the inflation distribution than the past realization of inflation. Thus, from period 6 onward, the probability s_t exceeds the trigger value $\bar{S}$ (panel E), and firms that receive a Calvo signal will abandon backward-looking indexation and instead choose the central bank's target as their index. Since the probability of such a signal is $1 - \theta$, a share of θ firms continues to implement backward-looking indexation. Thus, κ_t declines over time to the minimum exogenous degree of indexation, $\underline{\kappa}$ (panel F).

A strict inflation-targeting strategy fails to take advantage of the reduction in the cost of disinflation stemming from the decline

9. Sargent, Williams, and Zha (2006) provide a fascinating account of the implications of learning for inflation and stabilization when money growth and inflation are determined by the government's budget constraint rather than by an independent central bank.

Figure 2. Immediate versus Gradual Disinflation



Source: Author's calculations.

in perceived inflation persistence and backward-looking indexation. The reason is simply that the disinflation is completed prior to these favorable developments. Instead, a gradual disinflation strategy might be able to profit from such developments and achieve disinflation at lower output costs. A gradual disinflation strategy is optimal if central bank preferences incorporate output stability—that is, a positive weight α in the loss function (equation 14). In this case, the response parameter, δ_4, in the policy function (equation 16) must be positive but below δ_2^{-1}.

To simulate a gradual disinflation, I set $\delta_{4,t} = \delta_{2,t}/(1+\delta_{2,t}^2)$. Initially, the policy response coefficient, $\delta_{4,t}$, corresponds to 0.4, which is one-fifth of the policy response needed to meet the target immediately. The resulting outcome is depicted in figure 2, with the disinflation again starting in period 5. The initial output decline is much smaller, but it will be sustained for a much longer time than in the case of immediate disinflation. The inflation rate declines gradually. By period 15, inflation is within 0.5 percentage points of the long-run target of zero. If a period in the model is treated as a year, this ten-year disinflation is broadly similar to the Chilean experience between 1991 and 2001.

The cumulative sum of output gap losses is much smaller under the gradual approach than under strict inflation targeting. The cumulative output loss converges to about 26 percent of annual output spread over more than ten years. The reason for the decline in the sacrifice ratio from 2.0 in the case of strict inflation targeting to about 1.3 in the case of gradual disinflation is to be found in adaptive learning. As price-setters observe the fall in the inflation rate, they revise their estimate of inflation persistence downward. This reduction in c_t from 1.0 to about 0.8 adds disinflationary impetus and reduces the costs of disinflation. While the decline in perceived inflation persistence is much smaller under gradual than under immediate disinflation, the gradual approach can take advantage of the resulting reduction in disinflation costs.

With regard to the degree of backward-looking indexation, firms see no reason to switch from backward-looking indexation to the announced inflation target. The announced target is just too far away and progress toward it too slow to change the probability weights on lagged inflation versus the announced target. As a result, endogenous indexation does not come into play in terms of reducing the costs of disinflation under such a gradual disinflation strategy.

5. Inflation Targeting: Temporary Inflation Targets

Two important aspects of the Chilean disinflation strategy were its gradual nature and its use of temporary annual inflation targets. Having shown that the gradual approach helps reduce disinflation costs by taking advantage of the reduction in perceived inflation persistence, I now extend the analysis to consider the effect of announcing temporary targets. In the Chilean case, these temporary targets appear to have been pursued quite vigorously. This section thus investigates whether such temporary targets, π_t^*, could have an additional beneficial effect on learning and the degree of indexation and thereby lower the costs of disinflation further.

With temporary targets, the New-Keynesian Phillips curve needs to be slightly modified:

$$\pi_t = \frac{\kappa}{1+\beta\kappa}\pi_{t-1} + \frac{\beta}{1+\beta\kappa}E_t\left[\pi_{t+1}\right] + \frac{\lambda}{1+\beta\kappa}x_t + \frac{(1-\kappa)}{1+\beta\kappa}\left(\pi_t^* - \beta\pi_{t+1}^*\right). \quad (18)$$

Accordingly, the reduced-form inflation equation with adaptive learning and endogenous indexation corresponds to

$$\begin{aligned}\pi_t &= \frac{\kappa_{t-1}}{1+\beta(\kappa_{t-1}-c_{t-1})}\pi_{t-1} + \frac{\lambda}{1+\beta(\kappa_{t-1}-c_{t-1})}x_t \\ &\quad + \frac{(1-\kappa_{t-1})}{1+\beta(\kappa_{t-1}-c_{t-1})}(\pi_t^* - \beta\pi_{t+1}^*) \\ &= \delta_{1,t}\pi_{t-1} + \delta_{2,t}x_t + \delta_{3,t}.\end{aligned} \quad (19)$$

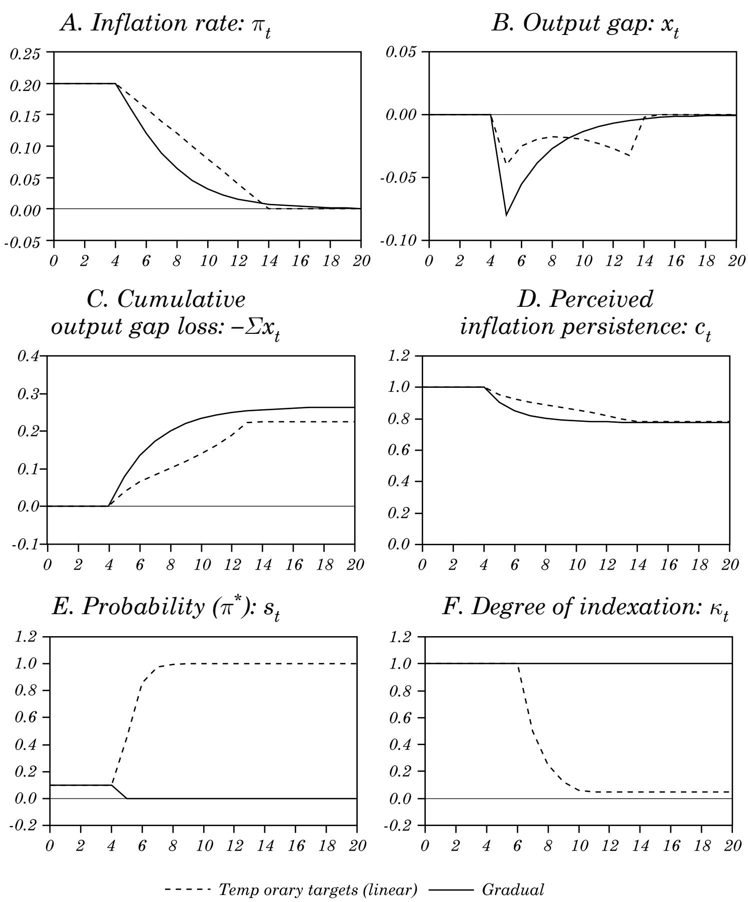
As a first example, consider a gradual, linear reduction in the inflation target by 2 percentage points per year. The long-run target of zero percent inflation is then reached in year 14, ten years after the start of disinflation. I assume that the central bank pursues these annual targets as actively as possible. In other words, the central bank implements strict inflation targeting with respect to temporary targets. After deciding on next year's inflation target, the central bank acts to meet this target. Thus, it pursues the following output gap policy:

$$x_t = -\delta_{4,t}\left(\delta_{1,t}\pi_{t-1} + \delta_{3,t} - \pi_t^*\right), \quad (20)$$

with $\delta_{4,t} = \delta_{2,t}^{-1}$, and $\delta_{(1,2,3)}$ consistent with equation (19).

The disinflation performance with temporary annual targets is shown by the dotted line in figure 3. It compares with the gradual disinflation (that is, the solid line) shown previously in figure 2. In both cases, the parameter governing the perceived degree of inflation persistence, *c*, declines toward a value of 0.8 (panel D). This decline occurs slightly faster under the gradual disinflation because inflation is initially reduced more quickly than the linear reduction implied by the annual targets.

Figure 3. Temporary Inflation Targets



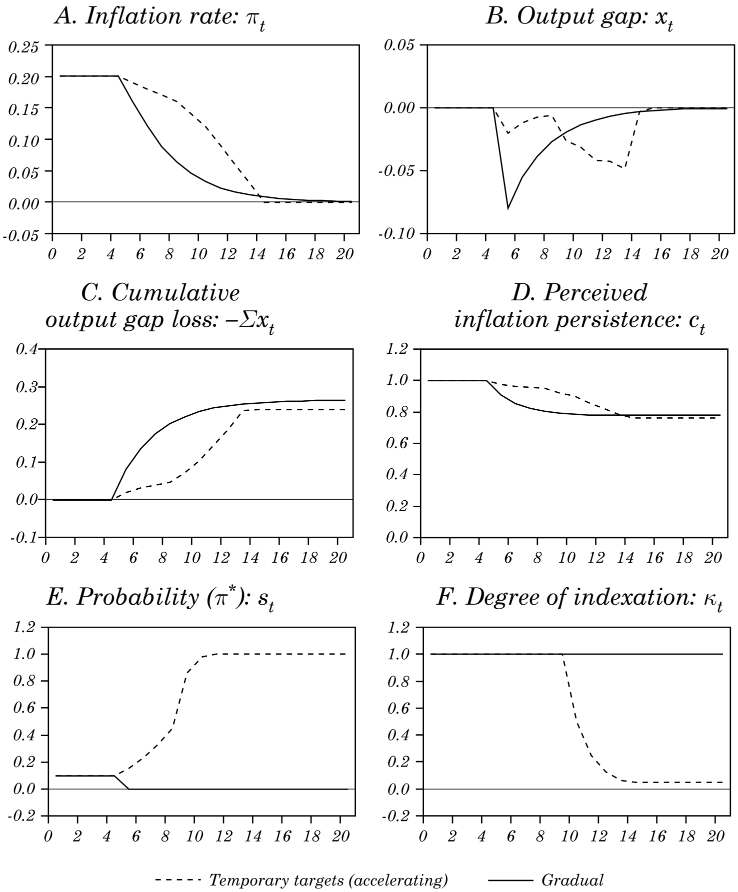
Source: Author's calculations.

An important difference arises with respect to the degree of backward-looking indexation. By announcing and meeting the temporary annual inflation targets, the central bank succeeds in convincing firms that they are better off choosing the central bank's target as an index for the pricing rule applied in those periods without Calvo-style optimal price-adjustment signals. The probability s_t that the central bank's target(s) will represent the mean of the inflation distribution rises quickly (panel E). It exceeds the trigger probability $\overline{S}$ of 0.8 by the second year of the disinflation. Every year from then on, a share of $1 - \theta$ of the firms that previously applied backward-looking indexation switches to using the central bank's targets. As a result, the degree of backward-looking indexation declines fairly rapidly and approaches the minimum level $\underline{\kappa}$ by year 11.

Unlike the gradual disinflation strategy with a long-run target, the strategy with temporary annual targets allows the central bank to take advantage of the endogenous reduction in backward-looking indexation. Firms change their behavior because they can already observe during the first few years of the disinflation that the central bank means to achieve its announced targets. Consequently, the output losses associated with disinflation are lower with annual targets. The cumulative output loss, (panel C) converges to 22 percent of output, that is 4 percent lower than in the case of the gradual disinflation. The sacrifice ratio is reduced to 1.1. Further substantial gains in terms of stabilization performance will accrue in the future. Given the substantial reduction in backward-looking indexation, the central bank will be able to reduce variations in inflation in the event of unexpected shocks at much lower cost in terms of output variability.

Next, I explore three alternative parameterizations of the sequence of annual inflation targets: targets that imply accelerating disinflation; targets that imply decelerating disinflation; and the annual targets set in Chile from 1991 to 2001. In the first case, shown in figure 4 the reduction in the central bank's annual targets accelerates over time (dotted line). The central bank initally lowers the inflation target by one percentage point per year. Starting in year 9, the fifth year of the disinflation, the inflation target is lowered by two percentage points per year. From year 11 onward, the target is lowered by three percentage points per year. The long-run target of zero percent is reached in year 14, after a ten-year disinflation process. Relative to the disinflation with linearly declining targets, accelerating targets initially imply a slower decline in inflation. The output gap incurred during the disinflation increases over time in absolute value. The total cost of disinflation—that is, the cumulative output gap—remains smaller than with the gradual disinflation strategy (solid line) but larger than with linearly declining targets. The cumulative output gap reaches 24 percent, versus 22 percent with linearly-declining targets. Because of the slow pace of disinflation in the first few years, price-setting firms take longer to become convinced that they are better off using the central bank's target as an index for their pricing rules in periods without Calvo-style signals. The probability s_t (panel E) rises slowly and takes five years to exceed the trigger value of 0.8. Only from year 10 onward do those firms that receive Calvo signals start switching from backward-looking indexation to the central bank's targets.

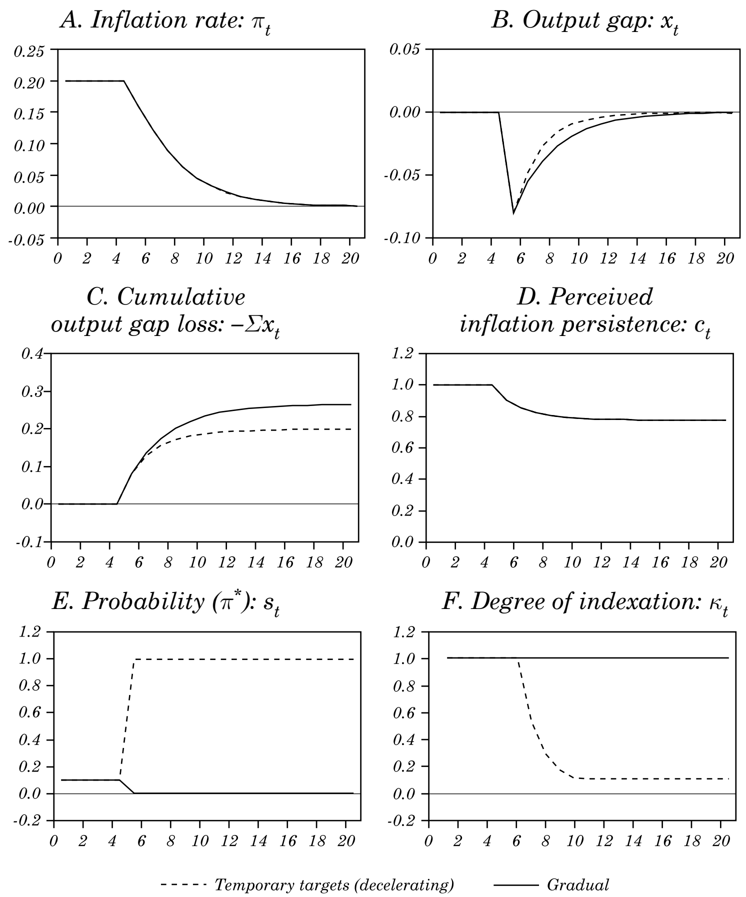
Figure 4. Accelerating Disinflation with Temporary Targets



Source: Author's calculations.

Figure 5 shows the simulation with decelerating targets. In the first year of disinflation, year 5, the central bank aims to lower inflation by 4 percentage points to 16 percent. The speed of disinflation declines in subsequent years. These annual inflation targets (dotted line) are set to be identical to the inflation path that is realized under the gradual disinflation with a long-run target (solid line). Thus, the actual path of inflation (panel A) coincides under these two scenarios. This parameterization is particularly interesting because it provides a ceteris paribus assessment of the reduction in disinflation costs that is achieved by announcing temporary annual

Figure 5. Decelerating Disinflation with Temporary Targets



Source: Author's calculations.

targets. As shown in panel B, the output gap associated with the disinflation with temporary targets (dotted line) is at all times equal to or smaller than (in absolute value) the output gap under gradual disinflation with a long-run target. The total cost of disinflation comes to 20 percent of output—that is, another 2 percent lower than with linearly declining targets. The sacrifice ratio associated with a disinflation from 20 percent to zero inflation is unity. Announcing and achieving the reduction of inflation by 4 percentage points in the first year of the disinflation convinces price-setting firms that the central bank means business. As a result, the probability s_t rises rapidly and firms soon start to abandon the practice of backward-looking indexation.

The annual targets set by the Chilean Central Bank between 1991 and 2001 also implied a decelerating disinflation. In 1990 inflation was substantially above 20 percent. The announced target for 1990 of 15–20 percent thus indicated a significant reduction with the start of the inflation-targeting strategy. Table 2 reports the announced target ranges and point targets, as well as the midpoints of these ranges. From 2001 onward, the Central Bank has aimed to keep inflation within a target range of 2 to 4 percent.

Figure 6 reports a simulation of a disinflation in the New-Keynesian model with adaptive learning and endogenous indexation using the midpoints of the Chilean target ranges from 1991 to 2001. [10] The initial conditions are the same as in the preceding simulations shown in figures 2 to 5. The midpoints of the Chilean target ranges are implemented from year 5 through year 15. To render the cost of disinflation incurred by the pursuit of the Chilean targets in the model comparable to the preceding simulations, I added a further reduction in the inflation target. In period 16, the target is reduced by an additional 3 percentage points so as to reach a long-run target of zero inflation.

The total cost of disinflation in terms of the cumulative output gap loss amounts to 18 percent of gross domestic product (GDP) spread over twelve years (panel C). The sacrifice ratio is 0.9, which is lower than in the simulation with decelerating targets shown in figure 5. This reduction is possible for the following reasons. The initial disinflation steps in years 5, 6, and 7 are vigorous enough to reduce the perceived degree of inflation persistence (panel D) and

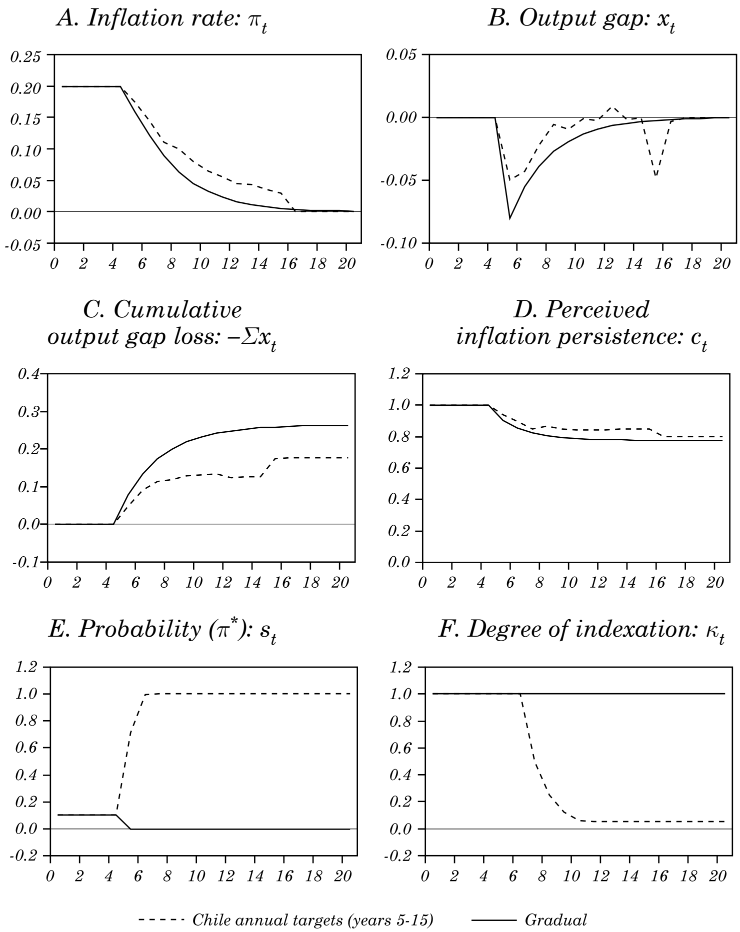
10. I disregard the potential effects of target ranges; see Orphanides and Wieland (2000) for an analysis of such nonlinearities.

Table 2. Chile's Inflation Targets: 1991–2001

Year of target	*1991*	*1992*	*1993*	*1994*	*1995*	*1996*	*1997*	*1998*	*1999*	*2000*	*2001*
Year in model	5	6	7	8	9	10	11	12	13	14	15
Range	15–20	13–16	10–12	9–11	8.0	6.5	5.5	4.5	4.3	3.5	2–4
Midpoint	17.5	14.5	11.0	10.0	8.0	6.5	5.5	4.5	4.3	3.5	3.0

Source: Central Bank of Chile.

Figure 6. Chile's Inflation Targets: 1991–2001 (Years 5 to 15)



Source: Author's calculations.

a. From 2001 onward, the Central Bank of Chile pursued an inflation target zone of 2 to 4 percent with a midpoint of 3 percent. For comparability with the preceding evaluation of disinflation costs, I have added a further 3 percent disinflation step in year 16 to achieve a long-run target of zero inflation.

to raise the probability s_t beyond the trigger level, $\bar{S}$. The degree of backward-looking indexation therefore declines over the course of the disinflation. However, the disinflation stretches out for a longer period than in figure 5 and thereby benefits even more from the reduction in inflation persistence and indexation.

The baseline version of the New-Keynesian model does not include structural shocks in the inflation equation. Such shocks are often added either to capture the presence of measurement error or to reflect missing variables or other sources of rigidity. I now proceed to introduce random shocks in the New-Keynesian Phillips curve:

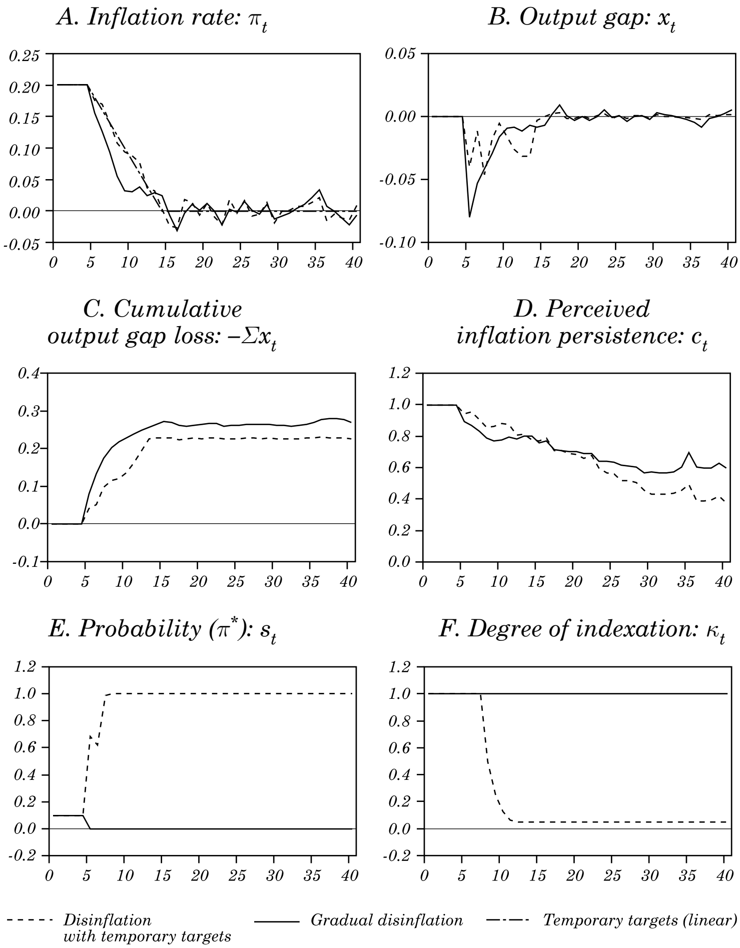
$$\begin{aligned}\pi_t = {} & \frac{\kappa}{1+\beta\kappa}\pi_{t-1} + \frac{\beta}{1+\beta\kappa}E_t\left[\pi_{t+1}\right] + \frac{\lambda}{1+\beta\kappa}x_t \\ & + \frac{(1-\kappa)}{1+\beta\kappa}\left(\pi_t^* - \beta\pi_{t+1}^*\right) + \eta_t.\end{aligned} \tag{21}$$

The shocks are denoted by η_t and are normally distributed with zero mean and variance $\sigma_\eta = 2^{-4}$. The timing of expectations formation, policy actions, and shocks is such that the shocks are realized after time t expectations have been formed and policy has been set. The shocks thus introduce noise in inflation that cannot be avoided by contemporaneous policy actions. However, in the period following the shock, the central bank will act to minimize further consequences from these variations that would occur as a result of the intrinsic persistence of inflation. To this end, the central bank induces offsetting variations in the output gap.

The fluctuations of inflation and output that result from random shocks and subsequent policy responses have an important influence on the dynamics of learning and endogenous indexation. On the one hand, such shocks imply that the central bank never meets its target exactly. Firms may therefore find it more difficult to assess whether it is better to use past inflation or the central bank's target as an index for their pricing rules in periods without Calvo signals. On the other hand, the fact that the central bank will set policy to counter the consequences of unforeseen shocks to inflation will generate information regarding the degree of inflation persistence and induce adaptive learning. Fluctuations may thus increase the speed of learning and reduce inflation persistence, and the costs of disinflation may decline further.

Figure 7 shows dynamic simulations with a particular draw of random shocks, η. The length of time covered is forty years, rather than twenty as in the preceding figures. The figure compares the outcome under a gradual disinflation with a long-run target (solid line) with a disinflation based on linearly declining annual targets (dash-dotted line). Panel A reports the actual inflation rates, which exhibit some random fluctuations, together with the annual targets (dotted line).

Figure 7. Shocks Accelerate Learning and Perceived Persistence Declines

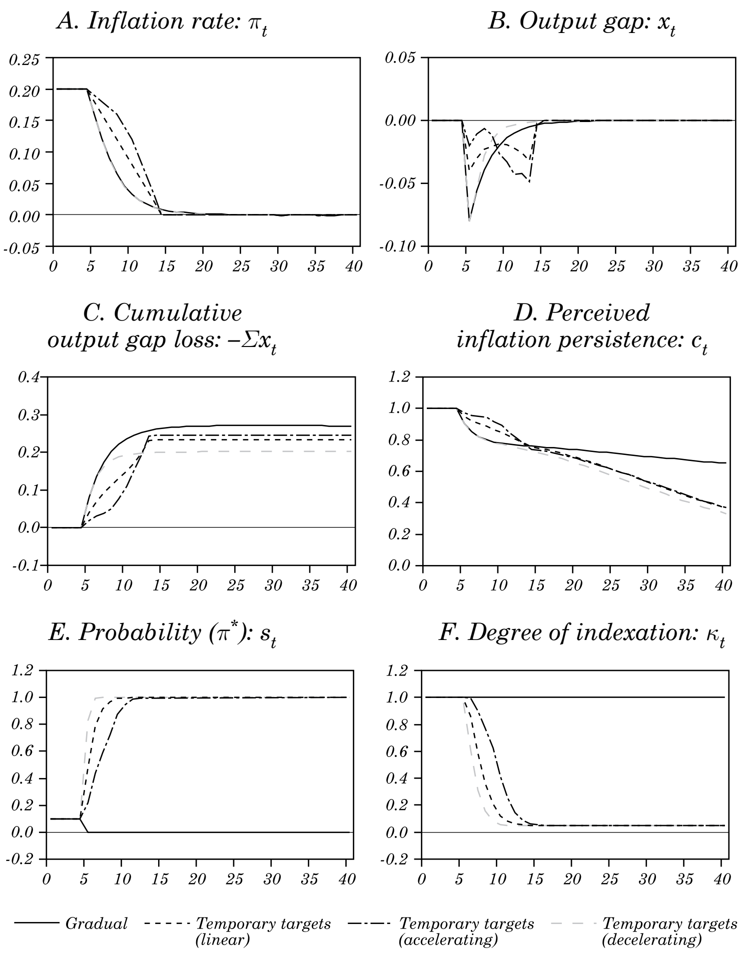


Source: Author's calculations.

Two aspects of these stochastic simulations are of particular interest. Panel D shows that the perceived degree of inflation persistence continues to decline even after the disinflation process has been completed. It is the policy response to the consequences of unforeseen shocks that stabilizes inflation fluctuations and drives down price setters' estimates of the persistence parameter, c_t. This decline is much more pronounced in the simulation with annual targets. By year 40, it reaches 0.4, while it is still at 0.6 in the gradual disinflation with long-run target. The reason is that the structural persistence from indexation is ultimately much smaller in the simulation with annual targets. The central bank's announcement and achievement of these targets has convinced firms to switch from backward-looking indexation to using the target rates. The probability s_t measuring the usefulness of central bank targets for indexation does not increase as smoothly as in the absence of unforeseeable random shocks. In figure 3, panel E, the probability s_t rises rapidly and smoothly above the trigger level in the simulation with linearly decline targets. In figure 7, panel E, it moves up and down a little bit before rising further above the trigger level. This finding shows that the switch from backward-looking indexation to the central bank targets is influenced by the particular series of shocks.

Figure 7 only reports the outcomes for a single draw of shocks. The strategy with temporary inflation targets need not always outperform the gradual disinflation strategy in terms of output losses. To shed further light on the likely outcomes, I simulated a thousand series of shocks drawn from a normal distribution and compute averages across these thousand simulations. The averages are reported in figure 8, which shows averages for the gradual disinflation with a long-run target (solid line), with linearly-declining annual targets (dotted line), with decelerating targets (dash-gray line) and with accelerating targets (dash-dotted line). The results are quite similar to the simulation without shocks, although they are not the same because of the nonlinearity resulting from adaptive learning and indexation. The ranking of speeds of disinflation (panel A) and cumulative output losses (panel C) remains unchanged. The perceived degree of inflation persistence reaches 0.4 for all three types of temporary targets by year 40. After many more years, it converges to a small but positive value consistent with the persistence implied by the minimum degree of backward-looking indexation under rational expectations. The increase in the probability s_t, (panel E) is fastest with decelerating targets and slowest with accelerating targets. As a result, the degree of backward-looking indexation declines most quickly with decelerating targets and most slowly with accelerating targets. In the case of a gradual disinflation with long-run targets, backward-looking indexation remains complete.

Figure 8. Averages over a Thousand Simulations



Source: Author's calculations.

6. A "Sophisticated" Central Bank versus One That Learns

These findings suggest that the performance of monetary policy could be improved further by allowing the central bank to observe and exploit the nonlinear dynamics stemming from adaptive learning and endogenous indexation—that is, equations (7), (10), and (11), in the design of dynamically optimal policy. Gaspar, Smets, and Vestin (2006a) study such an optimal policy problem with adaptive learning, but without endogenous indexation. They introduce the label "sophisticated" for a central bank that is capable of exploiting learning dynamics. In my model, such a sophisticated central bank would solve the following dynamic optimization problem:

$$\min_{x_t} E_t\left[\sum_{t=1}^{\infty}\beta^{t-1}\left(\pi_t-\pi^*\right)^2+\alpha x_t^2\right], \tag{22}$$

subject to $\pi_t = \delta_{1,t}\pi_{t-1} + \delta_{2,t}x_t + \delta_{3,t} + \eta_t$ and equations (7), (8), (10), and (11).

The optimal policy is nonlinear because it takes into account the nonlinearities arising from recursive estimation of the degree of inflation persistence—that is, equations (7) and (8)—and endogenous indexation—that is, equations (10) and (11).

Following Gaspar, Smets, and Vestin (2006a, 2006b), the central bank's choice variable is assumed to be the output gap and the central bank is assumed to aim at a long-run inflation target. An alternative approach, inspired by the present paper, would be to use annual inflation targets as the central bank's choice variable. A particular choice of temporary target would then automatically imply a given output gap according to the strict inflation-targeting policy shown by equation (20).

The optimization problem defined by (22) corresponds to a nonlinear dynamic programming problem with four state variables: (π_{t-1}, c_{t-1}, Σ_{t-1}, s_{t-1}). Numerical approximation of such a problem is complicated but within the reach of current methodology. However, optimal policy design here relies on rather courageous assumptions regarding the central bank's knowledge of private sector expectations formation. The central bank is assumed not only to observe the private sector's beliefs, but also to know the exact learning dynamics. The policy that could be implemented by such an extremely knowledgeable central bank might provide a useful benchmark for model-based comparison, but it does not represent a strategy that could be implemented in practice. I propose instead an alternative approach to policy design under uncertainty that can be pursued under more realistic informational assumptions.

Optimal policy design that could be implemented with the information available to central banks in practice takes recourse to learning. In this case, the central bank would learn about inflation dynamics by recursively estimating the relevant parameters of the reduced-form inflation equations (13) or (19). Contrary to the price-setting firms in the model, which were assumed to simply estimate a regression of inflation on its own lag, the central bank can spend more resources on learning. Certainly, central bank econometricians regularly estimate Phillips curves that include the effect of policy on inflation via the output gap, x_t.

In the model studied in this paper, central bank learning could be applied to the reduced-form inflation equation consistent with adaptive learning and endogenous indexation—that is,

$$\pi_t = \delta_{1,t}\,\pi_{t-t} + \delta_{2,t}\,x_t + \delta_{3,t} + \eta_t. \tag{23}$$

Central bank beliefs regarding the three time-varying parameters may be summarized by the vector $\mathbf{d}_t = (d_{1,t}, d_{2,t}, d_{3,t})$ and associated covariance matrix $\mathbf{\Sigma}_{d,t}$.[11]

$$\text{var}\left[\begin{pmatrix} d_{1,t} \\ d_{2,t} \\ d_{3,t} \end{pmatrix}\right] = \mathbf{\Sigma}_d = \begin{pmatrix} v_t^1 & v_t^{12} & v_t^{13} \\ v_t^{12} & v_t^2 & v_t^{23} \\ v_t^{13} & v_t^{23} & v_t^3 \end{pmatrix}. \tag{24}$$

The vector of state variables that characterize central bank beliefs contains nine variables, the three means, three variances, and three covariances. The associated updating equations for recursive least squares with time-varying parameters correspond to the following:[12]

$$\begin{pmatrix} d_{1,t} \\ d_{2,t} \\ d_{3,t} \end{pmatrix} = \begin{pmatrix} d_{1,t-1} \\ d_{2,t-1} \\ d_{3,t-1} \end{pmatrix} + \mathbf{\Sigma}_{t-1}\mathbf{X}_t F^{-1}\,(\pi_t - d_{1,t-1}\pi_{t-1} - d_{2,t-1}x_t - d_{3,t-1}), \tag{25}$$

11. See Wieland (2006). Related work on central bank learning in this context includes Cogley, Colacito, and Sargent (2007), Ellison (2006), Svensson and Williams (2007), and Wieland (2000a,2000b).

12. For a derivation of the updating equations using Bayes' rule or the Kalman filter, see Zellner (1971) and Harvey (1992), respectively.

$$\boldsymbol{\Sigma}_{d,t} = \boldsymbol{\Sigma}_{d,t-1} - \boldsymbol{\Sigma}_{d,t-1}\mathbf{X}_t F^{-1}\mathbf{X}_t' \boldsymbol{\Sigma}_{d,t-1} + \boldsymbol{\sigma}_d,$$

where $\mathbf{X}_t' = (\pi_{t-1}\, x_t\; 1)$. F refers to the conditional variance of inflation and , $F = \mathbf{X}_t \boldsymbol{\Sigma}_{d,t-1} \mathbf{X}_t' + \sigma^{\eta}$.

The information requirements for such a learning central bank are much less stringent than for the sophisticated central bank discussed above. Only inflation and output observations are needed. Potential output could be subsumed in the time-varying intercept. A fruitful area for future research would be to reassess the disinflation policies in the preceding section under the assumption that the central bank learns about the time-varying parameters governing the inflation process in this manner. Wieland (2000a, 2000b, 2006) and Beck and Wieland (2002) compute optimal learning policies for such problems with up to two unknown parameters and compare their performance to passive learning policies that do not take into account the central bank's own updating equations in optimization. At the least, policy design under passive learning could be applied to the policy problem in this paper.

7. Conclusions and Extensions

This paper has shown that inflation-targeting strategies can lower the costs of disinflation and future inflation stabilization. I have explored two channels through which such a reduction may take place: adaptive learning and endogenous indexation. Arguably, both channels may have played an important role in Chile's disinflation experience.

If market participants learn adaptively rather than form rational expectations, then history matters. As the central bank acts to bring inflation under control, market participants will observe the consequences of these actions and revise their beliefs regarding the degree of inflation persistence. Over time, adaptive learning lowers the cost of disinflation. A gradual approach to disinflation can take advantage of this beneficial effect.

Endogenous indexation implies that price-setting firms are allowed to choose between past inflation and the central bank's target as an index for their pricing rule in periods without Calvo-style signals to set prices optimally. Firms assess the likelihood that announced inflation targets determine steady-state inflation and adjust the indexation of contracts accordingly. A strategy of announcing and achieving short-term targets for inflation is able to influence the degree of backward-looking indexation. It implies that firms are able to observe fairly quickly whether the central bank acts to meet the targets it proclaims. When the central bank follows through on its commitments, the likelihood that firms switch from backward-looking indexation to the central bank's announced targets rises. Short-term annual targets that are pursued aggressively help reduce the degree of indexation more effectively than a strategy with a long-run target that is achieved only gradually.

This analysis suggests that dynamic general equilibrium models estimated under the assumptions of rational expectations and an exogenous, constant degree of backward-looking indexation may misjudge the costs of disinflation in two ways. First, the assumption of rational expectations may overstate the central bank's power to influence the costs of disinflation through words alone, whether they be announcements or verbal commitments. Learning implies that announcements need to be followed by action to convince market participants. The resulting reduction in inflation persistence is influenced by policy actions, as well as economic shocks. Second, the assumption of exogenous indexation may lead to model estimates that overstate the cost of disinflation and inflation-output trade-offs. Endogenous reductions in the degree of backward-looking indexation as inflation rates decline to a low level consistent with announced targets would present the central bank with more favorable trade-offs.

This research presents a number of interesting and potentially important possible extensions. These extensions concern the optimal design of monetary policy, the formation of expectations, the role of the interest rate, the role of the exchange rate, and the degree of openness of the economy. With regard to dynamically optimal policy design, two possible approaches were proposed in section 6 of the paper, including the derivation of the dynamically optimal policy that takes into account the nonlinear learning dynamics present in the model. Although such a policy relies on unrealistic informational assumptions, it would form a useful benchmark for comparison with practically implementable policies, such as the policy with central bank learning proposed in section 6.

As to the formation of expectations, it would be useful to evaluate the implications of alternative adaptive learning specifications (see Branch and Evans, 2006; Milani, 2007) for the cost of disinflation. It would also be interesting to study endogenous indexation under rational expectations. The quantitative effects of endogenous indexation could thus be studied separately from those stemming from adaptive learning.

The model consider here is very stylized. The central bank has been assumed to control the output gap directly. Instead, the transmission from the central bank's primary policy instrument (namely, the nominal short-term interest rate) and the output gap could be modeled explicitly. In other words, the model can be extended to include the log-linearized Euler equation of households—that is, the New-Keynesian IS curve. This extension would support exploration of a host of new questions regarding the design of interest rate rules and the conditions for stability under learning (see also Llosa and Tuesta, 2007).

Finally, Chile, like many inflation-targeting countries, is a small open economy. During the disinflation in Chile, favorable shocks to the exchange rate and the terms of trade may have played an important role in cushioning the economy. These effects could be examined by extending the analysis of learning and endogenous indexation conducted in this paper to a small open economy. In an open economy, further practical questions arise such as whether to target domestic inflation or CPI inflation and how to account for the exchange rate in interest rate policy.

References

Aguirre, A. and K. Schmidt-Hebbel. 2007. "Monetary Policy in Transition from Partial to Full-Fledged Inflation Targeting in Chile." In *Monetary Policy in Emerging Markets and Other Developing Countries,* edited by Nicoletta Batini. Nova Science Publishers.

Branch, W. and G.W. Evans. 2006. "A Simple Recursive Forecasting Model." *Economics Letters* 91(2): 158–66.

Beck, G. and V. Wieland. 2002. "Learning and Control in a Changing Economic Environment." *Journal of Economic Dynamics and Control* 26(9–10): 1359–77.

Calvo, G. 1983. "Staggered Prices in a Utility-Maximizing Framework." *Journal of Monetary Economics* 12(3): 383–98.

Calvo, G. and E. Mendoza. 1999. "Empirical Puzzles of Chilean Stabilization Policy." In *Chile: Recent Policy Lessons and Emerging Challenges*, edited by Guillermo Perry and Danny M. Leipziger. Washington: World Bank.

Caputo, R., F. Liendo and J.P. Medina. 2007. "New Keynesian Models for Chile in the Inflation-Targeting Period." In *Monetary Policy under Inflation Targeting*, edited by Frederic Mishkin and Klaus Schmidt-Hebbel. Santiago: Central Bank of Chile.

Caputo, R., J.P. Medina and C. Soto. 2006. "Nominal Rigidities, Indexation, and Inflation Persistence in Chile: A Structural Investigation." Santiago: Central Bank of Chile.

Christiano, L., M. Eichenbaum and C. Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113(1): 1–45.

Céspedes, L., M. Ochoa and C. Soto. 2005. "The New Keynesian Phillips Curve in an Emerging Market Economy: The Case of Chile." Working paper 355. Santiago: Central Bank of Chile.

Cogley, T., R. Colacito and T.J. Sargent. 2007. "Benefits from U.S. Monetary Policy Experimentation in the Days of Samuelson and Solow and Lucas." *Journal of Money. Credit, and Banking* 39(S1): 67–99.

Corbo, V., O. Landerretche and K. Schmidt-Hebbel. 2002. "Does Inflation Targeting Make a Difference?" In *Inflation Targeting: Design, Performance, Challenges*, edited by Norman Loayza and Raimundo Soto. Santiago: Central Bank of Chile.

Corbo, V. and K. Schmidt-Hebbel. 2003. "Inflation Targeting: The Latin American Experience." In *Latin American Macroeconomic*

Reform: The Second Stage, edited by Jose-Antonio González, Vittorio Corbo, Ann O. Krueger, and Aaron Tornell. University of Chicago Press.

De Gregorio, J. and E. Parrado. 2006. "Overshooting Meets Inflation Targeting." Working paper 394. Santiago: Central Bank of Chile.

Ellison, M. 2006. "The Learning Cost of Interest Rate Reversals." *Journal of Monetary Economics* 53(8): 1895–907.

Evans, G. and S. Honkapohja. 2001. *Learning and Expectations in Macroeconomics*. Princeton University Press.

Fraga, A., I. Goldfajn and A. Minella. 2003. "Inflation Targeting in Emerging Market Economies." Working paper 10019. Cambridge, Mass.: National Bureau of Economic Research.

Galí, J. and M. Gertler. 1999. "Inflation Dynamics: A Structural Econometric Analysis." *Journal of Monetary Economics* 44(2): 195–222.

Gaspar, V., F. Smets and D. Vestin. 2006a. "Adaptive Learning, Persistence and Optimal Monetary Policy." *Journal of the European Economic Association* 4(2–3): 376–85.

————. 2006b. "Optimal Monetary Policy under Adaptive Learning." *Computing in Economics and Finance* 183.

Harvey, A.C. 1992. *Time Series Models*, 2d ed. MIT Press.

Herrera, L.O. 2002. "Indexation, Inflation Inertia, and the Sacrifice Coefficient." In *Indexation, Inflation, and Monetary Policy*, edited by Fernando Lefort and Klaus Schmidt-Hebbel. Santiago: Central Bank of Chile.

Lefort, F. and K. Schmidt-Hebbel. 2002. "Indexation, Inflation, and Monetary Policy: An Overview." In *Indexation, Inflation, and Monetary Policy*, edited by Fernando Lefort and Klaus Schmidt-Hebbel. Santiago: Central Bank of Chile.

Llosa, L.G. and V. Tuesta. 2007. "Learning about Monetary Policy Rules When the Cost Channel Matters." Working paper 2007-014. Lima: Central Reserve Bank of Peru.

Marcet, A. and T. Sargent. 1989. "Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models." *Journal of Economic Theory* 48(2): 337–68.

Milani, F. 2007. "Expectations, Learning, and Macroeconomic Persistence." *Journal of Monetary Economics* 54(7): 2065–82.

Mishkin, F. and K. Schmidt-Hebbel. 2001. "One Decade of Inflation Targeting: What Do We Know and What Do We Need to Know?" Working paper 8397. Cambridge, Mass.: National Bureau of Economic Research.

Orphanides, A. and V. Wieland. 2000. "Inflation Zone Targeting." *European Economic Review* 44(7): 1351–87.

Orphanides, A. and J.C. Williams. 2006a. "Inflation Targeting under Imperfect Knowledge." Working paper 2006-14. Federal Reserve Bank of San Francisco.

————. 2006b. "Monetary Policy with Imperfect Knowledge." *Journal of the European Economic Association* 4(2–3): 366–75.

Sargent, T., N. Williams and T. Zha. 2006. "The Conquest of South American Inflation." Working paper 12606. Cambridge, Mass.: National Bureau of Economic Research.

Schmidt-Hebbel, K. and A. Werner 2002. "Inflation Targeting in Brazil, Chile and Mexico: Performance, Credibility and the Exchange Rate." Working paper 171. Santiago: Central Bank of Chile.

Svensson, L.E.O. and N. Williams. 2007. "Bayesian and Adaptive Optimal Policy under Model Uncertainty." Working paper 13414. Cambridge, Mass.: National Bureau of Economic Research.

Walsh, C. 2003. *Monetary Theory and Policy*, 2nd edition. MIT Press.

————. 2007. "Announcements and the Role of Policy Guidance." Paper prepared for the Thirty-Second Annual Economic Policy Conference. Federal Reserve Bank of St. Louis, 19–20 October.

Wieland, V. 2000a. "Learning by Doing and the Value of Optimal Experimentation." *Journal of Economic Dynamics and Control* 24(4): 501–34.

————. 2000b. "Monetary Policy, Parameter Uncertainty, and Optimal Learning." *Journal of Monetary Economics* 46(1): 199–228.

————. 2006. "Monetary Policy and Uncertainty about the Natural Unemployment Rate: Brainard-Style Conservatism versus Experimental Activism." *Advances in Macroeconomics* 6(1): 1288–97.

Woodford, M. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Zellner, A. 1971. *An Introduction to Bayesian Inference in Econometrics*. Hoboken, N.J.: John Wiley and Sons.

SOURCES OF UNCERTAINTY IN CONDUCTING MONETARY POLICY IN CHILE

Felipe Morandé
Universidad de Chile

Mauricio Tejada
Universidad de Chile

Monetary policy is made in an environment of substantial uncertainty. Consequently, academic researchers have sought to formally demonstrate the implications of uncertainty, as well as the ways in which central banks can manage it. The theoretical literature on uncertainty distinguishes between three types: additive uncertainty refers to central banks' lack of knowledge on the shocks the economy will face in the future; multiplicative uncertainty represents the lack of knowledge, or the erroneous knowledge, on one or more parameters of the behavioral model of the economy; and data uncertainty is associated with the fact that the information used by the central bank at the time policy decisions are made could be incorrect or could incompletely reflect the actual state of the economy. The objective of this paper is to review the quantitative relevance of these three types of uncertainty in the Central Bank of Chile's monetary policy. The paper is divided into two parts: the first covers the problem of data uncertainty and focuses on the output gap estimates for the full-fledged inflation-targeting period (1999 onward); the second centers on additive and multiplicative uncertainty for the period 1990–2006, with a special emphasis on the period after 1999.

We are thankful for the comments and suggestions received from Klaus Schmidt-Hebbel and Rodrigo Valdés. We are also grateful for the comments made on a previous version of this paper by participants attending the Central Bank of Chile conference where this paper was presented and to the participants of seminars held at the Department of Economics of Universidad de Chile and at the Economics Institute of Pontificia Universidad Católica de Chile.

Monetary Policy under Uncertainty and Learning, edited by Klaus Schmidt-Hebbel and Carl E. Walsh, Santiago, Chile.

Our analysis of data uncertainty focuses on the output gap because of its importance in forecasting inflation and because only preliminary figures for real output (real-time data) are available when monetary decisions are made. Also, the estimation of the output trend (part of the output gap) depends on statistical filters applied to output series, which contain these preliminary figures. For our exercise, we use several well-known univariate filters: the Hodrick-Prescott (HP) filter, the Baxter-King (BK) filter, the Christiano-Fitzgerald (CF) filter, the quadratic trend, and the Clark method based on the unobserved components model. To analyze their reliability and statistical accuracy with real-time data, we follow the methodology proposed by Orphanides and van Norden (1999). We find that revisions of the output gap in the case of Chile are important and persistent, and that correlations between the final data output gap and the real-time data output gap are relatively low. Nonetheless, the Clark method produces the best results, implying that caution should be used when evaluating the business cycle with real-time data and that using popular filters, like HP, could be misleading.

To evaluate the empirical importance of additive and multiplicative uncertainty, we use the methodology proposed by Zhang and Semmler (2005) to estimate behavioral equations for the Chilean economy with time-varying parameters and shocks with state-dependent variance (two states), which follow a first-order Markov process. To estimate behavioral equations, we use a slightly modified version of the forward-looking specification of Svensson (2000) and Al-Eyd and Karasulu (2008) for the equations that govern the behavior of a small open economy—namely, aggregate demand, the Phillips curve, and the real uncovered interest parity condition. We also use a technique from Kim (1993) to decompose total uncertainty, measured through the conditional variance of the forecast error, into two components: that associated with multiplicative uncertainty and that associated with additive uncertainty. We find that for all the behavioral equations of the economy, the uncertainty of shocks (that is, additive uncertainty) has been the most important factor in explaining total uncertainty. Moreover, the estimations support the hypothesis of state-dependent variances, as well as the hypothesis that these states could be considered periods of high and low volatility in the shocks. Also, total uncertainty of both the output gap and the inflation rate has declined over time, and the period of greatest stability coincides with the establishment of the full-fledge inflation-targeting framework for the conduct of monetary policy.[1]

1. This period also coincides with the establishment of the structural surplus rule for the conduct of fiscal policy and with a highly stable international context.

The paper is organized as follows. Section 1 reviews the literature on the types of uncertainty faced by central banks, the implications for the conduct of monetary policy, and the way uncertainty is usually modeled empirically. In section 2, we analyze the quantitative relevance of data uncertainty, focusing on the output gap estimates. Section 3 explores the importance of additive and multiplicative uncertainty in the models typically used to study the effects of monetary policy. Finally, concluding remarks are presented in section 4.

1. Monetary Policy and Uncertainty

In the last few years, academic researchers have become increasingly interested in formally demonstrating how central banks can deal with uncertainty (Schellekens, 2002; Feldstein, 2003). Some papers study the distinct types of uncertainty faced by central banks, which introduce important challenges in the modeling of monetary policy, and their implications for the behavior of the monetary authority. This group of studies includes Isard, Laxton, and Eliasson (1999), Martin and Salmon (1999), Svensson (1999), Wieland (2000), Meyer, Swanson, and Wieland (2001), Tetlow and von zur Muehlen (2001), Giannoni (2002), Orphanides and Williams (2002), and Söderström (2002). Other papers propose different strategies for managing uncertainty, such as robust monetary policy rules and learning mechanisms. Examples include Craine (1979), Holly and Hughes Hallett (1989), Basar and Salomon (1990), Bertocchi and Spagat (1993), Balvers and Cosimano (1994), Sargent (1998), Onatski and Stock (2002), and Wieland (2000).

Feldstein (2003) argues that central banks typically face four types of uncertainty: uncertainty about the current and future states of the economy, uncertainty about how the economy operates, uncertainty of individuals about their personal futures, and uncertainty about the impact of potential future monetary policies. However, the most common classification defines three types of uncertainty: additive uncertainty, multiplicative uncertainty, and data uncertainty.[2] Additive uncertainty represents the component of a forecast error associated with the outcome of an exogenous variable in the system (the regression model error). This type of uncertainty captures central

2. Another type of uncertainty considered in the literature, but not analyzed in this paper, is uncertainty about the probability distributions over possible events, known as Knightian uncertainty.

banks' lack of knowledge on the future shocks that the economy will face (Zhang and Semmler, 2005; de Grauwe and Senegas, 2006). Multiplicative (or parameter) uncertainty, in turn, represents the lack of knowledge, or the erroneous knowledge, of one or more parameters of the behavioral model of the economy (and its agents). Hall and others (1999) claim that this type of uncertainty can occur for several reasons, including the stochastic nature of the parameters, measurement errors in the data used to estimate the model, and structural changes. The distinction between additive and multiplicative uncertainty is based on the assumption that the true behavioral model of the economy is known. The limitation of this assumption is that total uncertainty, which could also result from misspecification of the model, is underestimated, so the results of any efforts to quantify this uncertainty using a particular specification of the behavioral model of the economy should be undertaken with caution.[3] Finally, data uncertainty is associated with the fact that the information used by the central bank at the time policy decisions are made could be incorrect or could incompletely reflect the actual state of the economy (Orphanides and van Norden, 1999). When these types of uncertainty are combined, they weigh heavily on policymakers (Rudebush, 2001). If policymakers have no knowledge of the actual state of the economy (regardless of whether the uncertainty lies in the data or in the behavior of the economy), they must base their decisions on expected outcomes. This could generate dilemmas in the adoption of an adequate policy if the outcome is not clear (for example, whether the central banks should react more aggressively or more passively).

Phillips (1954) and Theil (1964) were the first to introduce the idea of additive uncertainty, and their contributions have led to the expansion of the literature in this area. Phillips (1954), in studying whether the stabilization policy recommendations of simple models based on multipliers are appropriate and under what conditions this might be the case, showed that in a system that is automatically regulated (with flexible prices and interest rates), monetary policy could be a suitable instrument for stabilizing the economy, or at least for maintaining the economic system close to its desired values.

3. Although part of the existing literature defines multiplicative uncertainty as the lack of knowledge on the parameters and on the model, the distinction between the two is important from a practical point of view. If the distinction is not made, it is not possible to separate the concepts of additive and multiplicative uncertainty, given that any specification error affects both the regression error and the magnitude of the parameters (bias).

Monetary policy should also be able to deal with all but the most severe shocks. Theil (1964) expanded on Phillips (1954) by assuming that the policymakers choose their policy by maximizing a quadratic expected utility. He found that in a world where there is only uncertainty in shocks, policymakers could conduct their policy as if there were total certainty regarding the possible outcomes of the economy. This result is known as certainty equivalence and has important implications for monetary policy.

The period in which Philllips and Theil were working was marked by a high degree of confidence in econometric modeling, such that any error could be eliminated in the estimation of structural models, except that associated with additive uncertainty. However, the principle of certainty equivalence is valid only under certain conditions, specifically those pertaining in a linear-quadratic world. The policy implications could therefore differ depending on the assumptions adopted regarding the behavior of the central bank (that is, its loss function). Walsh (2004) finds that optimal monetary policy rules, derived from a quadratic loss function for the central bank, are robust under this type of uncertainty and do not require that the monetary authority change its rule in the presence of shocks. However, simple Taylor reaction functions can generate important increases in the central bank's loss function depending on whether, based on particular situations, they require changes in the central bank's behavior. Sack (2000) estimates and simulates a vector autoregression (VAR) model for the U.S. economy under different assumptions. He finds that if the only source of uncertainty is additive, the U.S. Federal Reserve should behave more aggressively than it does in practice. He argues that other types of uncertainty, such as multiplicative, generate greater gradualism in the Federal Reserve's monetary policy.

Holt (1962) was the first to analyze multiplicative uncertainty (uncertainty in the parameters). While exploring linear decision rules for stabilization and growth, he shows that policymakers are only able to apply an active stabilization policy when they can adequately anticipate the implications of the policies they adopt. Otherwise, they would contribute more to the instability of the economic system than to its stability. If the way in which the economy reacts is uncertain—that is, if the parameters of the behavioral model of the economy are uncertain—then the performance of monetary policy could be seriously affected. The certainty equivalence principle is not valid in this context, and the central bank should consider this type of uncertainty when making policy decisions.

Brainard (1967) uses a quadratic utility function for the policymaker, similar to that of Theil (1964), to study the effect of uncertainty in shocks and parameters. He finds that the certainty equivalence principle is valid if the only source of uncertainty is associated with shocks. However, when the economy's reaction to policy actions is unknown (that is, when the model feedback parameters are uncertain), the central bank's behavior is seriously affected and it becomes optimal to respond more cautiously to changes in the economic system. This result has important practical implications for the conduct of monetary policy, since it indicates that it could be optimal for policymakers not to expect to completely eliminate the gap between the observed objective variable and its target value, in a given period. This could be interpreted as a justification for a gradual monetary policy. Although Brainard's (1967) result is quite intuitive and is widely discussed in the literature (see Blinder, 1998), it cannot be generalized. Papers such as Martin and Salmon (1999) and Sack (2000) may provide some empirical validity to Brainard's (1967) work, but other studies show that the results depend on the model specification.[4] For example, Söderström (2002) shows that when the coefficients of the lagged variables in the model are subject to uncertainty, the optimal policy could be for the central bank to react more aggressively.[5]

The study of data uncertainty is relatively new in the literature on monetary policy. Academics and policymakers have only recently invested resources in studying the properties of real-time data and the implications for policy decisions (Bernhardsen and others, 2005). Croushore and Stark (2001) were the first to construct a database that provides a snapshot of the macroeconomic data available at any given point of time in the past, with the objective of showing the implications of forecasting with revised and real-time data. In their database, they refer to the data for a particular date as vintage and the collection of the vintages as the real-time data set. This methodology has been used in various empirical applications, which primarily focus on developed countries. Examples of studies exploring the implications of real-time

4. Both Martin and Salmon (1999) and Sack (2000) estimate a VAR model, the former for England and the latter for the United States. They show that incorporating multiplicative uncertainty in the model could explain the preference for gradualism in the central bank's behavior.

5. Other examples in support of the argument that multiplicative uncertainty does not necessarily lead the central bank to behave more cautiously can be found in Giannoni (2002) and Gonzalez and Rodriguez (2004).

data for monetary policy include Orphanides and van Norden (1999) and Orphanides (2001).[6] This literature highlights that the moment at which the data are obtained, their availability and their reliability for the empirical evaluation of policy rules are crucial for monetary policy performance, since they condition the decisions of the policymakers (Ghysels, Swanson, and Callan, 2002). In this regard, Rudebush (2001) and Bernhardsen and others (2005) argue that the new information that central banks obtain between two policy meetings does not justify drastic changes in the policy instrument, which can lead to very slow responses to particular economic events.

One of the variables that summarize the actual state of the economy—and that is thus crucial for monetary policy decisions—is the output gap. If potential output measures are not reliable, policy decisions may fail to react to the true economic conditions and may instead reflect measurement error. Orphanides and van Norden (1999) argue that the output gap is associated with important components of uncertainty, since central banks typically face at least three types of problems when evaluating the business cycle with real-time data. First, output data are revised continuously. Second, different methods of estimating potential output generally provide different results. When trend output is used as a proxy, different filtering procedures also yield a variety of outcomes; this problem is particularly critical with the end-of-sample estimates that are precisely the most relevant for policy decisions.[7] Third, a future evaluation of output data may indicate that the economy has experienced a structural change, which might not have been revealed by real-time data.

To illustrate these concepts, we consider the following economic model based on Zhang and Semmler (2005), which is standard in the literature of optimal rules of monetary policy:

$$\min_{\{\mathbf{u}_t\}_0^\infty} E_0 \sum_{t=0}^{\infty} \rho^t L(\mathbf{x}_t, \mathbf{u}_t), \tag{1}$$

subject to

$$\mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t, \varepsilon_t), \tag{2}$$

6. For an excellent literature review for the case of the United States, see Kozicki (2004).

7. Kuttner (1994) and St-Amant and van Norden (1998) find substantial differences in the estimation of trends using final output data and different estimation methods.

where ρ is the discount factor bounded between 0 and 1, $L(\mathbf{x}_t, \mathbf{u}_t)$ is the loss function of an economic agent (in this case, the central bank), $\mathbf{x}_t$ is the vector of state variables, $\mathbf{u}_t$ is the vector of control variables (the policy instrument), ε_t is the vector of shocks, and E_0 is the mathematical expectation operator based on the initial values of the state variables. This kind of model represents the basic framework of monetary policy analysis and control used by Clarida, Galí, and Gertler (1999), Svensson (1997, 1999), and Beck and Wieland (2002), where the constraints in equation (2) are the Phillips curve, the IS curve, and the interest rate parity condition (Svensson, 2000).

Given the state equations in (2), the central bank's problem consists in deriving a path for its instrument (the control variable $\mathbf{u}_t$) that satisfies equation (1). The question that arises, however, is whether the state equations can be correctly specified with time series estimates. Given the previous discussion, the answer to this question is no, since these equations can be subject to a high degree of uncertainty caused by shocks (ε_t), as well as to parameter uncertainty and data uncertainty. This is particularly important since the optimal monetary policy rule is derived from the solution of the previous problem.[8] This rule therefore depends on the parameters of the state equations. If the model parameters are uncertain, the estimated "optimal" monetary policy rule could be unreliable.

The brief literature review presented in this section shows that the different types of uncertainty (namely, additive, multiplicative, and data uncertainty) have important and different implications for the conduct of monetary policy. When the economy is faced with additive uncertainty, or uncertainty about the shocks it will face, the central bank could potentially behave as if it has total certainty about the results of its policy; this is known as the certainty equivalence principle. This result, however, depends on the type of assumptions adopted regarding the behavior of the central bank (its preferences) and the structure of the economy, since this principle is only valid in a linear-quadratic world and depends on whether the monetary authority behaves optimally. With regard to multiplicative uncertainty, or uncertainty in the parameters, the fact that the central bank does not know how the economy reacts to its policies would, in principle, justify a preference for a more gradual monetary policy. There is no consensus on this result, however, and the literature shows that the assumptions that are adopted in a particular model can lead to

8. See, for example, Svensson (1999).

different implications, including a preference for a more aggressive policy response. Finally, data uncertainty arises when the data are unknown at the moment policy decisions are made, when they contain measurement errors (resulting from previous revisions), or when they are unobservable. Policy decisions are seriously conditioned to the available information. Nevertheless, sudden changes in policy when a new set of information becomes known may not be justified, since the actual information could present an erroneous notion of the actual state of the economy. The literature has sought monetary policy rules that are immune to this type of uncertainty, for example, by using output growth rates or unemployment level rates rather than the gap with respect to their natural values.

2. Data Uncertainty: The Output Gap

To analyze the quantitative relevance of data uncertainty in the case of Chile, we focus on the output gap—defined as the difference between actual (measured) gross domestic product (GDP) and its trend—for the period from 2000 to 2006. We chose this period for two reasons: (1) the availability of historical data taken from the output series publications at each moment in time; and (2) this is the period in which the Central Bank of Chile adopted a full-fledged inflation-targeting scheme to conduct its monetary policy. We use real-time data (that is, data available to the Central Bank at the time policy decisions were made) and various well-known methods to estimate the output trend. For each method, we analyze the behavior of the end-of-sample output gap estimates, which are relevant for policy decisions, as well as the revisions of these estimates across time. We present the statistical properties of the revisions and verify the reliability of the estimates for each method.

The section is divided into two subsections. The first describes the methodological issues related to the construction of the output gap with real-time data and the detrending methods; the second presents the estimation results and their implications.

2.1 Methodological Issues

Monetary policy decisions are typically based on real-time data, which are classified as preliminary data (Bernhardsen and others, 2005). This is also true, to a lesser degree, of very old historical data. The preliminary nature of the data calls for it to be in constant revision. As suggested by the Central Bank of Chile[9], the data revision is motivated by factors such as the inclusion of new basic information (resulting from new sources of information or the improvement of these sources); the recalculation of the estimates (that is, revisions attributed to new estimates);[10] methodological improvements (reflecting changes in statistical methods, concepts, definitions, or classification); and error correction (either in the basic sources or in the calculations). One of the variables that encompasses the actual state of the economy and that is key for monetary policy decisions is the output gap. At the time policy decisions are made, this variable is estimated using preliminary output data, so it is necessary to assess the degree of reliability of these estimates.[11] For this assessment, we use real-time data to replicate the available information for the policymakers at every point in time. We thus simulate the actual environment of the monetary policy setting process (Ghlysels, Swanson, and Callan, 2002).

To analyze the reliability and statistical accuracy of the output gap estimates commonly used in the literature, we follow the methodology proposed by Orphanides and van Norden (1999). This consists of measuring, at each point in time, the degree to which the output gap estimates vary when the data are revised using different output gap estimation methods. This allows us to capture the effects caused by data revisions and the misspecification of statistical models used to estimate the output trend. The advantage of this methodological approach is that it does not require a priori assumptions on the true structure of the economy or on the process that generated the observed output time series. This approach has certain limitations: the analysis of the data revisions is based on a comparison of the output level observed at the end of the sample at every point in time with the "final" output, but there could still be measurement errors.

Orphanides and van Norden (1999) base their approach on two key definitions: the final and the real-time estimates of the output gap. The final estimate of the output gap is simply the difference between the last available vintage of output data and its trend (obtained via a detrending method). The real-time estimate, in turn, is a time series consisting of the last observed estimate of the output gap constructed

9. *Monetary Policy Report*, September 2004.

10.The recalculation of the estimates refers to the updating of either seasonal factors or the base period used in the constant price estimates.

11. If the output gap measures are not reliable, it could be advantageous, in some situations, for the central bank to base their monetary policy decisions on information on output growth (Orphanides and others, 2000; Bernhardsen and others, 2005).

as the difference between the observed output for each point in time (each vintage) and its trend. The real-time estimate for each period t contains all the revisions available up to that period and represents the estimate that the central bank may have calculated at the time policy decisions were made. Formally, assuming that we have access to the observed output series published at each point in time during N periods, we would have a matrix of the form $(\mathbf{y}^1, \mathbf{y}^2, \ldots, \mathbf{y}^N)$, where each $\mathbf{y}^i$ (with $i = 1, \ldots, N$) is a column vector that contains the output time series and where each column is an observation (row) shorter than the one that follows it.[12] If $f^{dt}(\cdot)$ is a function that detrends the time series $\mathbf{y}$, the final estimate of the output gap is given by

$$\text{GAP}^{final} = \ln(\mathbf{y}^N) - \ln[f^{dt}(\mathbf{y}^N)]. \tag{3}$$

If we then define the function $l(\cdot)$ as one that extracts the last real observation of the column vector $\mathbf{y}^i$, we have the real-time estimate of the output gap:

$$\begin{aligned}\text{GAP}^{real\text{-}time} = {} & \ln[\, l(\mathbf{y}^1), l(\mathbf{y}^2), \ldots, l(\mathbf{y}^N)\,] \\ & - \ln\{\, l[f^{dt}(\mathbf{y}^1)], l[f^{dt}(\mathbf{y}^2)], \ldots, l[f^{dt}(\mathbf{y}^N)]\,\}'.\end{aligned} \tag{4}$$

The difference between the final output gaps and the real-time output gaps represents the total revision of the estimates at each point in time. The statistical properties of these series of revisions inform our evaluation of the reliability and accuracy of the output gap estimates. For the estimates drawn from equations (3) and (4), it is necessary to define the function $f^{dt}(\cdot)$ (the detrending method), given that in practice neither the true potential output of the economy nor its data-generating process are known. This is important since these methods generally provide quite different results. In the case of Chile, Gallego and Johnson (2001) find that the set of methods used to estimate the trend component of output provide a wide range of estimates. The method chosen thus constitutes a source of uncertainty in addition to the revisions in the data.

A detrending method decomposes real output y_t (measured in logarithms) into two components: trend (y_t^T) and cycle (y_t^C), such that $y_t = y_t^T + y_t^C$. We consider five alternative univariate methods that

12. In the matrix $(\mathbf{y}^1, \mathbf{y}^2, \ldots, \mathbf{y}^N)$, we consider the missing observations as imaginary numbers.

are widely used in the literature: the Hodrick-Prescott filter; the Baxter-King filter; the Christiano-Fitzgerald filter; the quadratic trend; and Clark's method based on the unobservable components model.[13] Table 1 summarizes these methods and the models on which they are based. We focus only on univariate techniques of detrending, since the use of multivariate techniques requires the compilation of information on the data that is not revised (in real time) for each possible regressor in the model. Hence, the conclusions that are derived from the analysis correspond only to the evaluation of the univariate filters used here and cannot be applied to other alternative methods such as those used by the Central Bank of Chile and in some other papers for Chile (see Gredig, 2007; Fuentes, Gredig, and Larraín, 2007).[14]

The Hodrick-Prescott (HP) filter is one of the most popular detrending methods. It is based on choosing the trend that minimizes the variance of the cyclical component of the series, and it is subject to penalization for variations in the second difference of the cyclical growth component (Hodrick and Prescott, 1997). Both the Baxter-King (BK) filter and the Christiano-Fitzgerald (CF) filter are based on smoothing the series through the use of weighted moving averages. The fundamental difference between the two, for the case of symmetric filters as considered in this paper, lies in the choice of the objective function that defines the weights (Baxter and King, 1999; Christiano and Fitzgerald, 2003). Moreover, the Christiano-Fitzgerald filter imposes the restriction that the filter weights add up to zero when unit roots are considered. The quadratic trend, in turn, is a method of deterministic components that assumes that the behavior of the trend series is triggered by a second-order polynomial. This method is thus flexible at the moment of detecting slow trend changes.[15] Finally, the unobserved components model allows us to specify the data-generating processes for the output time series and use these to identify the trend and

13. See Orphanides and van Norden (1999) for an extensive revision of the detrending methods and its principal advantages and disadvantages. See Gallego and Johnson (2001) for an interesting compilation of the use of the five methods in different central banks.

14. The approach currently used by the Central Bank of Chile to estimate the output gap is based on the production function.

15. Its simplicity has made it quite valuable for empirical applications related to monetary policy (for example, Clarida, Galí, and Gertler, 1998), but its use has generated much controversy based on the argument that better modeling of output requires stochastic components in the model.

Table 1. Alternative Methods of Calculating the Output Trend

Method	*Calculation*
Hodrick-Prescott (λ = 1,600)	$y_t^T = \arg\min \sum_{t=1}^{T} \left[(y_t - y_t^T)^2 + \lambda (\Delta^2 y_{t+1}^T) \right]$
Baxter-King (6, 32)[a]	$y_t^T = \sum_{c=1}^{q+1} \omega^{BK}(1,c) y_{t+1-c} + \sum_{c=2}^{q+1} \omega^{BK}(1,c) y_{t+c-1}$ $t = q+1,...,n-q$
Christiano-Fitzgerald (6, 32, 1, 0, 0)[b]	$y_t^T = \sum_{c=1}^{q+1} \omega^{CF}(1,c) y_{t+1-c} + \sum_{c=2}^{q+1} \omega^{CF}(1,c) y_{t+c-1}$ $t = q+1,...,n-q$
Quadratic trend	$y_t = \alpha + \beta t + \gamma t^2 + y_t^C$
Clark (unobserved components)	$y_t = y_t^T + y_t^C$ $y_t^T = g_{t-1} + y_{t-1}^T + \nu_t$ $g_t = g_{t-1} + \omega_t$ $y_t^C = \delta_1 y_{t-1}^C + \delta_2 y_{t-2}^C + e_t$

Source: Authors' calculations.
a. The numbers 6 and 32 represent the minimum and maximum of the desired oscillation period for quarterly data.
b. The numbers 6 and 32 have the same interpretation as in the Baxter-King filter. The numbers 1 ,0, 0 represent the existence of unit roots, without drift and symmetric filter, respectively.

cyclical components. Clark (1987) proposes a model in which he assumes that the trend component follows a random walk process with drift and the cyclical component follows an AR(2) process. The main advantage of this type of model is that it allows a richer short-term dynamic specification for the model.

2.2 Results

The output data observed at each point in time were constructed using data compiled from the monthly publications (bulletins) of the Central Bank of Chile. We constructed an output series for each new statistical entry in which a new output record was published, incorporating the revisions of the data published before.[16] For the quantitative evaluation of uncertainty in the output gap estimates, we consider the period between the first quarter of 2000 and the last quarter of 2006, although the output gap estimates were calculated based on information since 1986.[17] Hence, the first time series we use covers the period between the first quarter of 1986 and the first quarter of 2000. The series that follows contains an additional quarter not included in the previous series, and this occurs successively until the last series, which comprises the complete period from the first quarter of 1986 to the last quarter of 2006. Each output series was seasonally adjusted using the X-12-ARIMA procedure employed by the Central Bank of Chile. The series thus reflect both the revisions and the reestimation of seasonal factors. Finally, the series published in the last quarter of 2006 is our final output series, although we are aware that this series contains data that will be revised in the future.

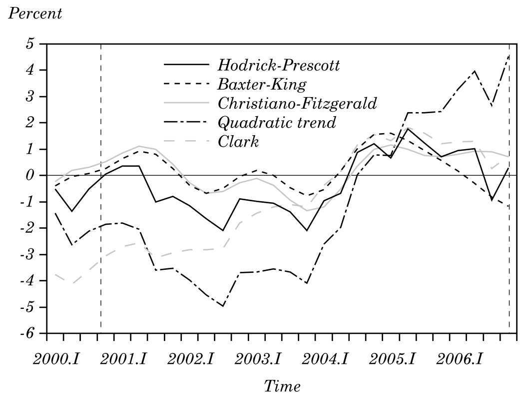
The compilation of the information described above produced a total of twenty-eight output series for each point in time. We apply the five detrending methods to each of these estimates to calculate the output gap. Following the methodology applied by Orphanides and van Norden (1999), our final estimates are the output gap for the last available series and our real-time estimates are the series constructed with the last observation of each of the output gaps estimated with the twenty-eight series. Figures 1 and 2 illustrate these estimates using final and real-time data.

As the figures show, most of the estimations generated by the different detrending methods behave similarly in terms of their trajectories. This is true for both the estimations using final data and those using real-time data. The only exception is the estimation of the output gap based on the quadratic trend. Despite the comovements observed in the different series, however, the magnitude of the changes varies considerably among methods. The different methods also produce a wide range of output gap estimates. The average difference between the highest and lowest estimates is 6 percent with final data and 12 percent with real-time data. The order of magnitude of these differences

16. In some cases, the revisions were observed for one or two quarters back, while in others, such as the periods with base changes, the revisions were performed on the full series. The Central Bank revised the national accounts and changed the base year on two separate occasions during the sample period. The first time was in the fourth quarter of 2001, when the base year changed from 1986 to 1996 prices, and the second time was in the last quarter of 2006, when the base year changed to 2003. (The vertical dotted lines in figures 1 to 3 show these changes.)

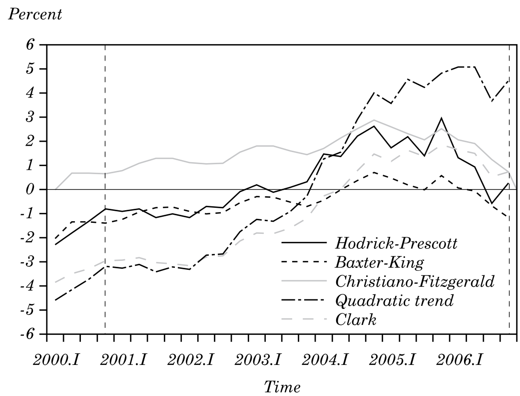
17. For a statistical filter to produce reasonable results, we need at least one complete cycle in the series, which implies that long time series are necessary.

Figure 1. Output Gap Estimates for the Chilean Economy with Final Data



Source: Authors' calculations.

Figure 2. Output Gap Estimates for the Chilean Economy with Real-Time Data

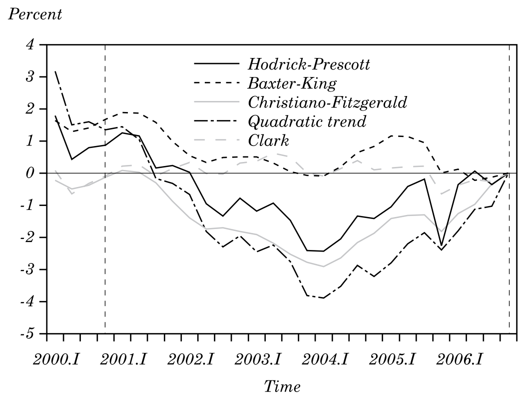


Source: Authors' calculations.

is considerable since they are much greater than the difference between the highest and the lowest points of the business cycle within the period considered (approximately 5 percent for both types of data and for a majority of filters). The average dispersion between methods is also important, reaching 2.3 percent with final data and 4.3 percent with real-time data. In addition, the estimations using final data tend to be clustered between the fourth quarter of 2004 and the third quarter of

2005. These estimates remain relatively close toward the end of the period of analysis, with the exception of the output gap based on the quadratic trend. This latter pattern is not observed with real-time estimates. To provide a qualitative idea of the importance of data revision, figure 3 shows the difference between the estimates with final data and those with real-time data for the five detrending methods. This difference represents the total revision in the output gap.

Figure 3. Total Revisions in the Output Gap for the Chilean Economy



Source: Authors' calculations.

The figure reveals that the magnitude of the revisions is also important and differs substantially among the filters used, with an average dispersion of revisions among different measures of 2.8 percent. The most extreme cases are observed in early 2004, when revisions of the HP, CF, and quadratic trend methods were the most important in the entire sample. This is due to the fact that these filters do not adequately capture the turning point of the output gap in that period (see figures 1 and 2), and it suggests that real-time estimates were imprecise. This is not the case for the BK and Clark methods, which present practically null revisions at that same point in time. Rather, the most important revisions for these last two filters are seen at the beginning of the sample. To better understand the differences between the estimates with final data and those with real-time data, we present descriptive statistics of the output gap estimates and of the revisions for the five filters in tables 2 and 3, respectively. Figure 4 shows the time behavior of all these estimates.

Table 2. Descriptive Statistics of the Output Gap Measures calculated with Final and Real-Time Data

Filter and data	*Mean*	*Absolute value*	*Standard deviation*	*Minimum*	*Maximum*	*Correlation*
Hodrick-Prescott						
Final estimates	–0.003	0.010	0.011	–0.021	0.018	1.000
Real-time estimates	0.002	0.012	0.014	–0.023	0.030	0.611
Baxter-King						
Final estimates	0.002	0.006	0.007	–0.012	0.016	1.000
Real-time estimates	–0.005	0.007	0.007	–0.020	0.007	0.561
Christiano-Fitzgerald						
Final estimates	0.002	0.007	0.008	–0.013	0.012	1.000
Real-time estimates	0.015	0.015	0.007	0.000	0.029	0.203
Quadratic trend						
Final estimates	–0.012	0.028	0.029	–0.050	0.045	1.000
Real-time estimates	0.001	0.031	0.035	–0.046	0.051	0.841
Clark						
Final estimates	–0.010	0.019	0.020	–0.041	0.018	1.000
Real-time estimates	–0.011	0.020	0.020	–0.039	0.019	0.988

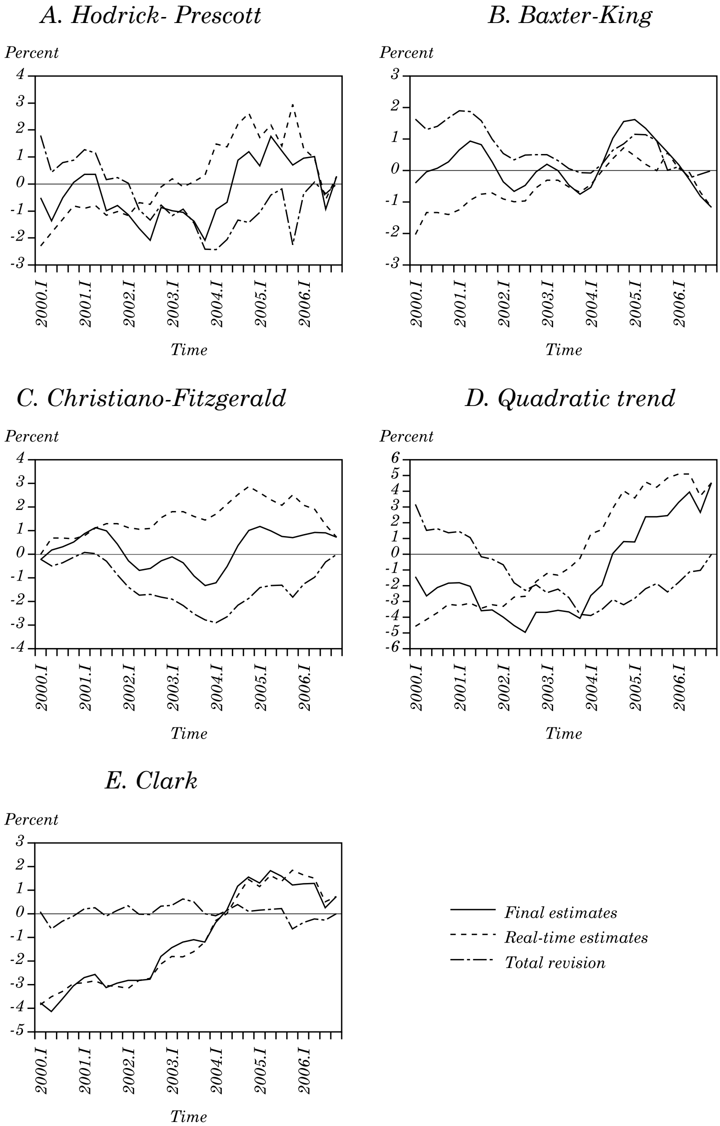
Source: Authors' calculations.

Table 3. Descriptive Statistics of the Total Revisions in the Output Gap

Filter	*Mean*	*Absolute value*	*Standard deviation*	*Minimum*	*Maximum*	*AR(1)*
Hodrick-Prescott	−0.005	0.010	0.011	−0.024	0.018	0.700
Baxter-King	0.007	0.007	0.007	−0.002	0.019	0.875
Christiano-Fitzgerald	−0.013	0.013	0.009	−0.029	0.001	0.939
Quadratic trend	−0.013	0.020	0.019	−0.039	0.032	0.842
Clark	0.000	0.002	0.003	−0.006	0.006	0.473

Source: Authors' calculations.

Figure 4. Estimation of the Output Gap and Total Revisions using Final and Real-Time Data for the Five Alternative Filters



Source: Authors' calculations.

Comparing the results in tables 2 and 3 shows that, on average, total revisions are similar to or greater than the output gap estimates for all filters used.[18] Something similar occurs with the average gap in absolute value. This confirms the previous discussion, since the revisions are always significant in magnitude regardless of whether the economy is in a recession or is expanding. With respect to the minimum and maximum points of the cycle, the estimations with final and real-time data coincide with the minimum values of the gap only in the case of the Clark method (see figure 4, panel E), while the estimations coincide with the maximum values for the BK filters, the quadratic trend, and the Clark method (see figure 4, panels B, D, and E). This suggests that most of the methods fail to identify the magnitude of the recessive periods. The last column of table 2 shows the correlation coefficients between final data estimates and real-time data estimates for each filter. The highest correlations are observed for the Clark and the quadratic trend methods (over 0.8), while the CF and BK filters produce the lowest correlations. Another important element is the degree of persistence of the revisions, since as the revisions persist over time, the discrepancies between the final and real-time estimates tend to remain or disappear slowly over time. The last column of table 3 reports the estimated first-order autocorrelation coefficients for total revisions, which indicate that these revisions are highly persistent, with the exception of the Clark filter.

We have yet to address the issue of whether the output gap measures constructed with real-time data are reliable.[19] Since the different methods vary substantially with respect to the size of the cyclical component, it is more convenient to compare the reliability of the real-time estimates using independent scale measures. Table 4 presents the reliability measures used by Orphanides and van Norden (1999). In the first column, we present the correlation between final and real-time series for each detrending method. The other three indicators in table 4 measure in different ways the relative importance of the revisions (the ideal value for these three indicators is zero). The N/S indicator is the ratio of the standard deviation of the revision to that of the final estimate of the output gap and approximates the

18. This result is qualitatively similar to that found by Orphanides and van Norden (1999) for the U.S. economy.

19. We define reliability in terms of quantifying the difference between the final estimates and the real-time estimates. Our measures thus do not indicate anything regarding the reliability of each method as a tool for the estimation of the true output gap (Bernhardsen and others, 2005).

noise-to-signal ratio. The OPSING indicator shows the frequency with which the real-time estimates of the output gap reveal a different sign than the final estimates. Finally, the XSIZE indicator shows the frequency with which the absolute value of the revision exceeds the absolute value of the final estimates of the output gap. The Clark and the quadratic trend methods reveal smaller noise levels and smaller frequencies in observations with errors in the sign and with significant size in the revision. The CF filter performs the worst under these reliability measures.

Table 4. Descriptive Statistics of the Reliability Measures for Alternative Different Filters[a]

Filter	*Correlation*	*N/S*	*OPSIGN*	*XSIZE*
Hodrick-Prescott	0.611	1.055	0.286	0.500
Baxter-King	0.560	0.902	0.321	0.536
Christiano-Fitzgerald	0.203	1.229	0.393	0.750
Quadratic trend	0.841	0.650	0.071	0.214
Clark	0.988	0.156	0.000	0.036

Source: Authors' calculations.
a. The first column presents the correlation between the final and real-time series for each detrending method. The N/S indicator is the ratio of the standard deviation of the revision to the standard deviation of the final estimate of the output gap; it approximates the noise-to-signal ratio. The OPSING indicator shows the frequency with which the real-time estimates of the output gap reveal a different sign than the final estimates. The XSIZE indicator shows the frequency with which the absolute value of the revision exceeds the absolute value of the final estimates of the output gap.

In sum, the above results show that, in general, revisions of the output gap seem to be important and persistent for the period considered and that the correlations between the final and real-time estimates of the output gap are relatively low. Nonetheless, the Clark method provides the most favorable statistics. The analysis also reveals that the Clark method is the most reliable with real-time data.[20] Comparing our results with those of Orphanides and van Norden (1999) for the U.S. economy, we find that the different reliability measures generally produce similar values. These results

20. As a robustness test, we also calculated the reliability measures in real time using output gap estimations with unadjusted and seasonally adjusted data through the seasonal dummy variables. Our conclusions do not change (for details, see appendix A). This exercise was done to verify whether the reestimation of the seasonal factors, which is not present in the unadjusted data and is constant when we use seasonal dummy variables, substantially influences our results.

imply that caution should be used when assessing the level of the real-time estimates of the output gap, at least with the methodologies used here. Additionally, our results should be considered a lower bound for measurement errors that could be present in the output gap estimates, because comparisons are made with respect to a measure of the final output gap that could contain unrevised data.

3. Additive and Multiplicative Uncertainty

To focus on the empirical importance of additive and multiplicative uncertainty, we use data for the 1990 to 2006 period but with emphasis on the 1999–2006 subsample, the full-fledged inflation-targeting period. We adopt a slightly modified version of the forward-looking specification of Svensson (2000) and Al-Eyd and Karasulu (2008) to estimate the behavioral equations of a small open economy, as is the case of Chile (namely, aggregate demand, the Phillips curve, and the real uncovered interest parity condition). As in Zhang and Semmler (2005), we do not include a monetary policy rule in this specification, given that the paper's objective is to analyze the primary sources of uncertainty faced by the Central Bank, which is associated with the structure and behavior of the economy.[21] To capture the sources of uncertainty, we estimate the model with time-varying parameters, assuming that shocks have state-dependent variances (two states) and that their behavior follows a first-order Markov process. This strategy allows us to decompose the conditional variance of the forecast error into two components: one associated with the parameters (multiplicative uncertainty) and one with the shocks in the model (additive uncertainty).

3.1 Methodological Issues

The existing literature on additive and multiplicative uncertainty typically uses models that explicitly consider the stochastic volatility potentially present in the errors (heteroskedasticity) and time-varying parameters (Zhang and Semmler, 2005). The papers that explicitly address parameter uncertainty include Cogley and Sargent (2002), who study the inflation dynamics of the United States in the postwar period using a Bayesian VAR with time-varying parameters. Another

21. Moreover, the optimal monetary policy feedback parameters will depend on the structure and behavior of the economy.

example is Semmler, Greiner, and Zhang (2005), who estimate the Phillips curve and a monetary policy Taylor rule for the euro area, also with time-varying parameters. Both works find substantial changes in the model parameters. However, even though the evidence encountered when using models with time-varying parameters points to the existence of important degrees of uncertainty, this analysis cannot be separated from additive uncertainty in the modeling process. When additive uncertainty is not considered, volatility in the parameters could be exaggerated when it is indeed captured (Sims, 2002). Sims and Zha (2006), who study regime changes in the dynamics of the U.S. economy, find evidence in favor of stable model dynamics but unstable variance of the disturbances. In response, Cogley and Sargent (2005) modify their original model considering time-varying parameters and stochastic volatility; they also find the existence of regime changes. More recent examples of the estimation of Taylor rules with time-varying parameters and stochastic volatility can be found in Kim and Nelson (2006) and Zampolli (2006).

To incorporate both additive and multiplicative uncertainty, we follow the approach used by Zhang and Semmler (2005). We use a model with time-varying parameters and shocks characterized by state-dependent variance. In contrast to Cogley and Sargent (2005), who assume that the variance of the shocks changes each period, we assume that the variance has only two states (high and low) and follows a Markov process, as in Sims and Zha (2006).[22] This specification, besides having the advantage of dealing with both types of uncertainty in the same model, allows the decomposition of the variance of the forecast error into two components: one associated with additive uncertainty and one with multiplicative uncertainty (Kim, 1993).

The specification we use for the behavioral equations of the economy is a slightly modified version of the specification of Svensson (2000) and Al-Eyd and Karasulu (2008); it is a neo-Keynesian version for a small open economy comprising the IS curve (aggregate demand), the short-run supply curve (Phillip's curve), and the real uncovered interest parity condition (UIP). We diverge from these authors however, in allowing deviations of the UIP because of imperfections in the capital markets, capital controls, speculative bubbles, and so forth. As is usual in the modern dynamic stochastic general equilibrium (DSGE) literature, the deviations in the UIP are modeled by introducing a

22. These authors assume that the variance of the regression errors follows a Markov process with three states.

backward-looking component in the original specification of Svensson (2000) and Al-Eyd and Karasulu (2008). The behavioral equations of the economy can thus be written as

$$y_t = \theta_{1,t}\, y_{t-1} + \theta_{2,t}\, E_t[y_{t+1}] + \theta_{3,t}\, r_{t-1} + \theta_{4,t}\, q_{t-1} + \varepsilon_t^d, \tag{5}$$

$$\pi_t = \phi_{1,t}\, \pi_{t-1} + \phi_{2,t}\, E_t[\pi_{t+1}] + \phi_{3,t}\, y_{t-1} + \phi_{4,t}\, q_t + \varepsilon_t^s, \tag{6}$$

and

$$q_t = \gamma_{1,t}\, E_t[q_{t+1}] + \gamma_{2,t}\,(r_t - r_t^f) + \gamma_{3,t}\, q_{t-1} + \upsilon_t, \tag{7}$$

where y_t represents the real output gap, π_t is the inflation rate, r_t is the short-term real interest rate, q_t is the real exchange rate, and r_t^f is the foreign real interest rate, observed in period t. The terms $E_t[y_{t+1}]$, $E_t[\pi_{t+1}]$ and $E_t[q_{t+1}]$ represent the expectations for period $t + 1$ of the output gap, the inflation rate, and the real exchange rate, respectively, conditional on the information available at period t (E_t is the expectations operator). The terms ε_t^d, ε_t^s and υ_t are shocks with state-dependent variances. The first two are aggregate demand and supply shocks, respectively, and the third is associated with the exchange market. As described by Al-Eyd and Karasulu (2008), this last disturbance term could be interpreted as a risk premium that captures the effects of unobservables, such as the exchange market sentiments. Finally, $\theta_{i,t}$ (with $i = 1, 2, 3, 4$), $\phi_{i,t}$ (with $i = 1, 2, 3, 4$), and $\gamma_{i,t}$ (with $i = 1, 2, 3$) are the time-varying parameters.

Two interesting observations can be made about this specification. First, the explicit inclusion of the exchange rate in the modeling process is relevant for an economy such as Chile, whose Central Bank uses inflation targeting as a monetary policy framework. Relative to the closed economy models, the specification introduces an important additional transmission channel of monetary policy and incorporates the external shock effect on the domestic economy. Second, the specification incorporates both forward-looking and backward-looking terms (hybrid model), for which there is empirical backing at least in the case of the Phillips curve (Caputo, Liendo, and Medina, 2006, and Céspedes, Ochoa, and Soto, 2005). Forward-looking terms can be justified by appealing to sticky price models of the Calvo (1983) type, whose wage-setting (or price-setting) mechanism is built in for a share of Chilean labor contracts.

The inclusion of forward-looking components, however, introduces the problem of how the components are measured or approximated, a choice that can have important implications for estimation properties (namely, consistency). The literature proposes various ways to deal with these variables and the most appropriate estimation techniques in each case. An obvious option is to use ex post data, that is, to approximate the expectation variables with their respective observed future values. While this option is operationally simple, it generates an endogeneity bias in the estimation of the model parameters, which leads to inconsistent estimates (Kim and Nelson, 2006).[23]

Galí and Gertler (1999), Roberts (2001), and Galí, Gertler, and López-Salido (2005) propose a methodology to address the endogeneity problem using ex post data for the forward-looking component of the model and instrumentalizing expectations through generalized method of moments (GMM) estimation. The use of GMM techniques to estimate the Phillips curve and the forward-looking Taylor rules is very common in the literature.[24] Along these lines, Kim (2004, 2006) proposes the application of instrumental variables for the estimation with endogenous regressors, using time-varying parameter models and regime changes. This methodological proposal solves the endogeneity problem by applying the Kalman filter in a two-stage Heckman (1976) estimation.[25] The specification of the behavioral equations in equations (5) to (7) can be rewritten in a state-space form under Kim's (2004, 2006) methodology as follows:

$$\begin{aligned}
&x_t = \mathbf{w}_t'\beta_{1,t} + \mathbf{v}_t'\beta_{2,t} + \varepsilon_t, && \varepsilon_t \sim N(0, \sigma^2_{\varepsilon,S_t});\\
&\beta_t = \beta_{t-1} + \eta_t, && \eta_t \sim N(\mathbf{0}, \mathbf{Q}_\eta);\\
&\mathbf{v}_t = \mathbf{Z}_t'\delta_t + \xi_t, && \xi_t \sim N(\mathbf{0}, \mathbf{Q}_\xi);\\
&\delta_t = \delta_{t-1} + \kappa_t, && \kappa_t \sim N(\mathbf{0}, \mathbf{Q}_\kappa);\\
&\sigma^2_{\varepsilon,S_t} = \sigma^2_{\varepsilon,0} + (\sigma^2_{\varepsilon,1} - \sigma^2_{\varepsilon,0})S_t, && \sigma^2_{\varepsilon,1} > \sigma^2_{\varepsilon,0};
\end{aligned} \qquad (8)$$

23. This is relevant because one of our objectives is precisely to study parameter uncertainty. Another straight-forward option is to use data from expectation surveys to construct a proxy variable of expectations (Roberts, 1995). This alternative has two potential problems: the first is associated with the availability of long time series for the estimation; the second is survey measurement error.

24. Several papers apply this methodology to Chile, including Céspedes, Ochoa, and Soto (2005), who estimates a hybrid Phillips curve, and Corbo (2002), who estimates a reaction function for the Central Bank.

25. Kim and Nelson (2006) use this methodology to estimate a forward-looking Taylor rule with ex post data for the United States.

where x_t represents a vector of state variables (y_t, π_t, and q_t for aggregate demand, the Phillip's curve, and the UIP, respectively), $\mathbf{w}_t$ is the vector of explanatory variables that are assumed to be exogenous or predetermined (y_{t-1}, r_{t-1}, and q_{t-1} for aggregate demand, π_{t-1}, y_{t-1}, and q_t for the Phillip's curve, and $r_t - r_t^f$ and q_{t-1} for the UIP), $\mathbf{v}_t$ is a vector of endogenous explanatory variables, which are correlated with the model errors ε_t (y_{t+1}, π_{t+1}, and q_{t+1}, respectively), $\mathbf{Z}_t$ is a vector of instrumental variables, $\boldsymbol{\beta}_t = (\boldsymbol{\beta}_{1,t}, \boldsymbol{\beta}_{2,t})'$ and $\boldsymbol{\delta}_t$ are vectors of time-varying parameters, $\boldsymbol{\eta}_t$, $\boldsymbol{\xi}_t$, and $\boldsymbol{\kappa}_t$ are Gaussian errors with a matrix of variances $\mathbf{Q}_i$ (with $i = \boldsymbol{\eta}, \boldsymbol{\xi}, \boldsymbol{\kappa}$), and S_t is an unobservable indicator variable that is equal to one in the high-volatility state and zero otherwise. We assume that the variance of errors ε_t present two states with transition probabilities that follow a Markov process and that are expressed as $\Pr[S_t = 1 \mid S_{t-1} = 1] = p$ and $\Pr[S_t = 0 \mid S_{t-1} = 0] = q$.

Kim (2006) proposes specifying the endogeneity in the model assuming that the correlation between the error term, ε_t, and the standardized forecast error associated with the endogenous variables, ξ_t^* (that is, the prediction error associated with the rational expectations of the agents), is constant and equal to ρ. On the other hand, in an earlier work that considers state-dependent variance of the errors, Kim (2004) suggests that this correlation will also be state dependent. The model error can thus be rewritten as

$$\varepsilon_t = \xi_t^{*}{}'\rho_{S_t}\sigma_{\varepsilon,S_t} + \sqrt{1-\rho_{S_t}{}'\rho_{S_t}}\,\sigma_{\varepsilon,S_t}\omega_t,$$

with $\omega_t \sim \mathrm{N}(0, 1)$. Using this last expression we can write the first equation of model (8) as

$$x_t = \mathbf{w}_t{}'\boldsymbol{\beta}_{1,t} + \mathbf{v}_t{}'\boldsymbol{\beta}_{2,t} + \xi_t^{*\prime}\rho_{S_t}\sigma_{\varepsilon,S_t} + \sqrt{1-\rho_{S_t}{}'\rho_{S_t}}\,\sigma_{\varepsilon,S_t}\omega_t, \tag{9}$$

with $\omega_t \sim \mathrm{N}(0, 1)$, where $\rho_{S_t} = \rho_0 + (\rho_1 - \rho_0)S_t$ and S_t is the same indicator variable defined above. In this last equation, the model error is independent of $\mathbf{v}_t$ and ξ_t^*. Hence, the estimation generates parameters that are consistent. For the estimation, Kim (2004, 2006) proposes the following two-stage procedure. The first stage consists in estimating a model that instrumentalizes the endogenous variables using the maximum log-likelihood method based on the error forecast and the conventional Kalman filter. That is,

$$
\begin{aligned}
&\mathbf{v}_t = \mathbf{Z}_t'\boldsymbol{\delta}_t + \boldsymbol{\xi}_t, && \boldsymbol{\xi}_t \sim \mathrm{N}\,(\mathbf{0}, \mathbf{Q}_\xi);\\
&\boldsymbol{\delta}_t = \boldsymbol{\delta}_{t-1} + \boldsymbol{\kappa}_t, && \boldsymbol{\kappa}_t \sim \mathrm{N}\,(\mathbf{0}, \mathbf{Q}_\kappa).
\end{aligned}
\tag{10}
$$

The standardized forecast error of $\mathbf{v}_t$ is then calculated as

$$
\xi_t^* = \mathbf{Q}_{\xi,t|t-1}^{-1/2}\left(\mathbf{v}_t - \mathbf{Z}_t{}'\boldsymbol{\delta}_{t|t-1}\right),
$$

for all $t = 1, 2,\ldots, T$. The second stage consists in using the calculated forecast error to estimate the following model using maximum log-likelihood techniques that combine the Kalman filter and the expectation-maximization (EM) algorithm proposed by Hamilton (1989, 1990):[26]

$$
\begin{aligned}
&x_t = \mathbf{w}_t{}'\boldsymbol{\beta}_{1,t} + \mathbf{v}_t{}'\boldsymbol{\beta}_{2,t} + \xi_t^{*\prime}\rho_{S_t}\sigma_{\varepsilon,S_t} + \sqrt{1-\rho_{S_t}{}'\rho_{S_t}}\,\sigma_{\varepsilon,S_t}\omega_t, && \omega_t \sim \mathrm{N}\,(0,1);\\
&\boldsymbol{\beta}_t = \boldsymbol{\beta}_{t-1} + \boldsymbol{\eta}_t, && \boldsymbol{\eta}_t \sim \mathrm{N}\,(\mathbf{0}, \mathbf{Q}_\eta);\\
&\sigma^2_{\varepsilon,S_t} = \sigma^2_{\varepsilon,0} + (\sigma^2_{\varepsilon,1} - \sigma^2_{\varepsilon,0})S_t, && \sigma^2_{\varepsilon,1} > \sigma^2_{\varepsilon,0};\\
&\rho_{S_t} = \rho_0 + (\rho_1 - \rho_0)S_t.
\end{aligned}
\tag{11}
$$

Finally, Kim (1993) suggests a procedure, using specification (8), to decompose the conditional variance of the forecast error (f) into two components: f^1, the conditional variance resulting from changes (lack of knowledge) in the model parameters (that is, multiplicative uncertainty) and f^2, the conditional variance given the heteroskedasticity in the error term (additive uncertainty).[27]

26. The estimation algorithm is presented in appendixes B, C, and D. A potential limitation of this methodology for estimating the behavioral equations of the economy is that Kim (2004, 2006) assumes that the shocks associated with each equation are independent of each other, and he thus does not take advantage of the information contained in the possible correlations (that is, common states). In other words, the methodology permits the estimation of each equation separately, so the different states of the shocks will not necessarily coincide for the three equations. Zhang and Semmler (2005) find very different occurrence probabilities for each state of the shocks depending on whether they are dealing with aggregate demand or the Phillips curve, indicating that the states in the model do not coincide in the same time period.

27. In this paper, Kim (1993) identifies the sources of uncertainty and their importance associated with the process of monetary creation in the United States.

In this procedure, Kim exploits the informational structure of the model related to the probability distributions in the different states. The conditional variance stemming from multiplicative uncertainty depends on the state in a previous period, while the conditional variance from additive uncertainty depends on the state in the current period. This decomposition is quite useful since it gives us the percentage of the total variance of the forecast error that is caused by each source of uncertainty. Formally,[28]

$$f_t = f_t^1 + f_t^2$$
$$\rightarrow \begin{cases} f_t^1 = (\mathbf{w}_{t-1}, \mathbf{v}_{t-1}) \left\{ \sum_{i=0}^{1} \Pr\left[S_t = i \mid \psi_{t-1}\right] \left[\mathbf{P}_{t|t-1}^i + \left(\tilde{\boldsymbol{\beta}}_{t|t-1} - \boldsymbol{\beta}_{t|t-1}^i\right)\left(\tilde{\boldsymbol{\beta}}_{t|t-1} - \boldsymbol{\beta}_{t|t-1}^i\right)' \right] \right\} (\mathbf{w}_{t-1}, \mathbf{v}_{t-1})' \\ f_t^2 = \sigma_{\varepsilon, S_t}^2 = \sigma_{\varepsilon,0}^2 + (\sigma_{\varepsilon,1}^2 - \sigma_{\varepsilon,0}^2)\Pr\left[S_t = 1 \mid \psi_{t-1}\right] \end{cases} \quad (12)$$

where

$$\tilde{\boldsymbol{\beta}}_{t|t-1} = \sum_{i=0}^{1} \Pr\left[S_t = i \mid \psi_{t-1}\right] \boldsymbol{\beta}_{t|t-1}^i$$

and where $\mathbf{P}_{t|t-1}^i$ is the variance-covariance matrix of $\boldsymbol{\beta}_{t|t-1}^i$ in state i.

3.2 Results

To estimate equation (8), we use quarterly data for the period beginning in the first quarter of 1990 and ending in the last quarter of 2006. The output gap, y_t, is the difference between the observed GDP and its trend, calculated using the HP filter. We use the HP filter because it is one of the most commonly used filters in the literature and it thus allows us to compare our results with those of other papers that estimate behavioral equations for Chile. Although the Clark

28. For details on the formal derivation of the decomposition of the conditional variance of the forecast error, see Kim and Nelson (1999).

filter behaves best with real-time data, according to the results in the previous section, this does not imply that it is the best filter to estimate the "true" output trend. Additionally, our measure of the output gap is "final" output, based on an output series that ends in 2006. Thus, the uncertainty associated with data revisions is not included in the types of uncertainty analyzed in this section.[29] The quarterly inflation rate, π_t, is measured as the quarterly variation of the underlying consumer price index excluding regulated prices and prices of fuel and of some perishable goods such as fruits and vegetables (CPIX). As in Céspedes, Ochoa, and Soto (2005), we use the CPI variation instead of the implicit deflator variation of the GDP since the latter is measured with considerable noise in the case of Chile and is strongly influenced by variations in the terms of trade. Also, the Central Bank's inflation target is expressed in terms of CPI variation. In the case of the real exchange rate, q_t, we chose the bilateral exchange rate index with the United States. Finally, the foreign and domestic short-term interest rates, r_t and r_t^f, are defined as the monetary policy rates of Chile and the United States, respectively. All the previous data were obtained from the Central Bank's database. Table 5 shows the parameters estimated using Heckman's two-stage procedure detailed in Kim (2004, 2006).[30] The parameters presented in this table are not structural parameters of the model.

We would like to highlight two interesting results. The first is that variances of shocks confirm that there are two states in the three behavioral equations: a high-volatility state and a low-volatility state. For the aggregate demand estimations, the variance of shocks in the high-volatility state is substantially greater than in the low-volatility state (0.48 versus 0.05). The difference between these variances for the Phillips curve is just as large (0.54 and 0.03 in the high and low-volatility states, respectively). We obtain similar results in the case of the UIP (3.75 versus 2.45), although the magnitude of the difference is not as large as in the previous two cases. All the variances are statistically significant, with the exception of the variance associated with the high-volatility state of the Phillips curve. Finally, while the variances of shocks for the UIP do not differ significantly, the size of the variances is considerable compared with those found for aggregate

29. The way detrending is done may affect the estimations, so we run a robustness analysis below.

30. In applying the Kalman filter in the evaluation of the likelihood function, we eliminated twelve observations at the beginning of the sample owing to the presence of nonstationary time series in the model; see Kim and Nelson (1999).

Table 5. Estimation of the Behavioral Equations

Aggregate demand			*Phillips curve*			*Real uncovered interest parity*		
Parameter	*Estimated value*	*Standard deviation*	*Parameter*	*Estimated value*	*Standard deviation*	*Parameter*	*Estimated value*	*Standard deviation*
p	0.6571	0.5267	p	0.6639	1.5101	p	0.9479	3.2325
q	0.6586	0.0644	q	0.8475	0.0501	q	1.0000	0.0001
$\sigma_{\eta_1^\theta}$	0.0697	0.2565	$\sigma_{\eta_1^\phi}$	2.4407	1.0338	$\sigma_{\eta_1^\gamma}$	0.0000	0.0026
$\sigma_{\eta_2^\theta}$	0.0797	0.2441	$\sigma_{\eta_2^\phi}$	1.2700	0.8449	$\sigma_{\eta_2^\gamma}$	0.0007	2.3801
$\sigma_{\eta_3^\theta}$	0.2942	0.2540	$\sigma_{\eta_3^\phi}$	0.0000	0.0001	$\sigma_{\eta_3^\gamma}$	0.0000	0.0029
$\sigma_{\eta_4^\theta}$	0.0002	0.0002	$\sigma_{\eta_4^\phi}$	1.6518	0.9554			
$\sigma_{\varepsilon,0}$	0.0570	0.0098	$\sigma_{\varepsilon,0}$	0.0329	0.0084	$\sigma_{\upsilon,0}$	2.4467	0.2295
$\sigma_{\varepsilon,1}$	0.4806	0.2347	$\sigma_{\varepsilon,1}$	0.5497	1.2718	$\sigma_{\upsilon,1}$	3.7539	0.1850
ρ_0	0.5123	0.1594	ρ_0	0.0010	0.2473	ρ_0	0.4924	0.2057
ρ_1	0.6324	0.1892	ρ_1	0.4705	0.1446	ρ_1	1.0000	0.2750
Log-likelihood	−64.026			−80.389			−109.64	

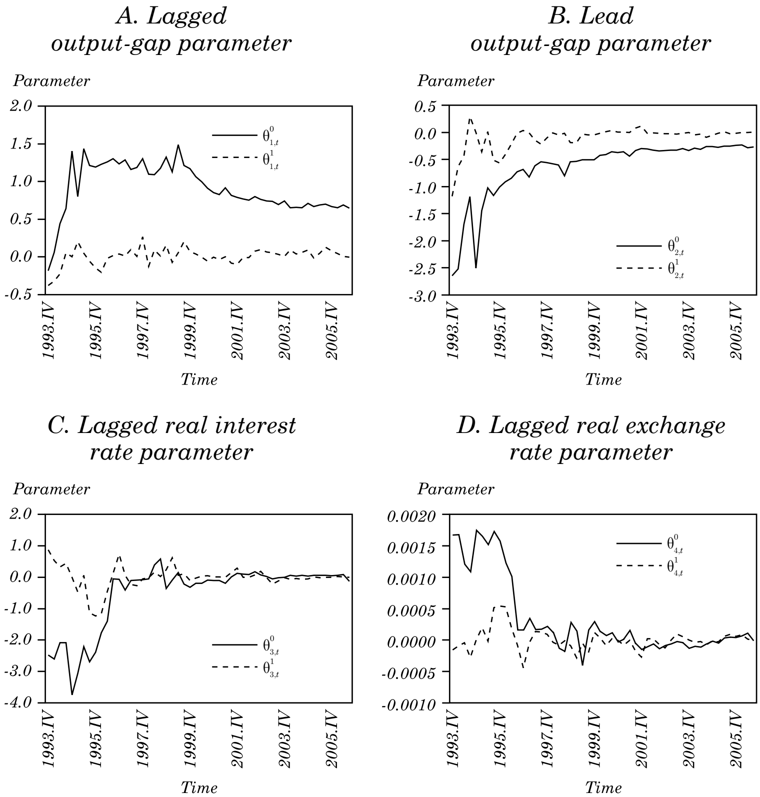
Source: Authors' calculations.

demand and the Phillips curve. The second interesting result is the existing correlation between the shocks of the behavioral equations and the errors in the economic agents' expectations, which also vary substantially with the states. In particular, the results suggest that agents tend to commit crucial errors in their forecasts in high-volatility states of the shocks. This fact is particularly true for the Phillips curve, where such correlation varies between 0.001 and 0.470 for both states, and for the real uncovered interest parity condition (0.49 versus 1.00). In the case of aggregate demand, there is also an important correlation in the high-volatility state. Nonetheless, the difference between the correlations of the two states is less evident than in the previous two cases. Also, the correlation coefficients are highly significant for all cases except the one associated with the low-volatility state of the shocks in the Phillips curve.

Figures 5 to 7 show the behavior over time of the structural parameters of the equations estimated in table 5. There are two series in each figure, which correspond to the relevant values of the parameters in each possible state of shocks in the model (that is, high volatility and low volatility). In the case of the aggregate demand parameters (figure 5), there are two clearly defined periods. The first period, which ends in 1999, is marked by high instability and substantial differences between the parameters of the two states associated with the demand shocks. During this period, the average probability that the economy was in a high-volatility state was 0.82, and the macroeconomic context was characterized by a substantial range of variation in the annual GDP growth rate (from 15 percent to below 6 percent) and by high inflation rates. The second period (from 1999 onward) saw a significant reduction in instability, as well as in the differences of the parameters with respect to the state of the shocks, with the exception of the parameter associated with the output gap's degree of persistence. The average probability that the economy was in a high-volatility state was only 0.10. These results suggest that the multiplicative uncertainty associated with aggregate demand tends to decline over time. Also, the output gap's degree of persistence ($\theta_{1,t}$) and its response to changes in relative prices ($\theta_{4,t}$) have declined over time, while the contrary has occurred with the degree of response to expectations ($\theta_{2,t}$) and the monetary policy interest rate ($\theta_{3,t}$). This is consistent with the logic of the inflation-targeting framework.[31]

31. In 1999, the full-fledged inflation-targeting framework was established.

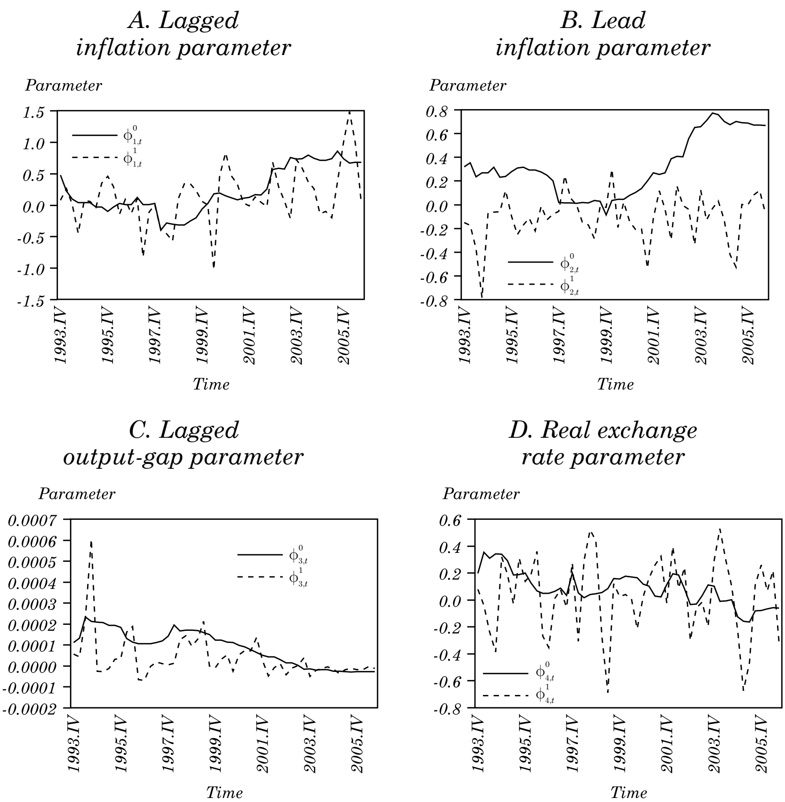
Figure 5. Time-Varying Parameters Estimated for the Aggregate Demand



Source: Authors' calculations.

The parameters of the Phillips curve show a significant dependence on the state of the supply shocks (see figure 6). In periods of high volatility, the parameters tend to show high instability, while they are much more stable in periods of low volatility. Unlike the results of the aggregate demand parameters, this dependence was maintained throughout the entire period. The state of shocks is thus key to explaining greater or lower degrees of uncertainty in the Phillips curve parameters. A high-volatility state of shocks prevailed throughout most of the 1990s (with an average probability of 0.9), so the relevant parameters in that period were those of the high-volatility state. In the most recent period (1999 onward), the average probability was only 0.06. Figure 6 also reveals that when the economy experiences a relatively calm period with respect to the supply shocks, persistence of the inflation rate ($\phi_{1,t}$) and the importance of expectations in the determination of the inflation rate ($\phi_{2,t}$) are greater. This happens toward the end of the period of analysis. The trend is lower in the case of the response of inflation to the business cycle ($\phi_{3,t}$) and to variations in the real exchange rate ($\phi_{4,t}$). When the supply shocks are highly volatile, however, there is no definite trend for the Phillips curve parameters.

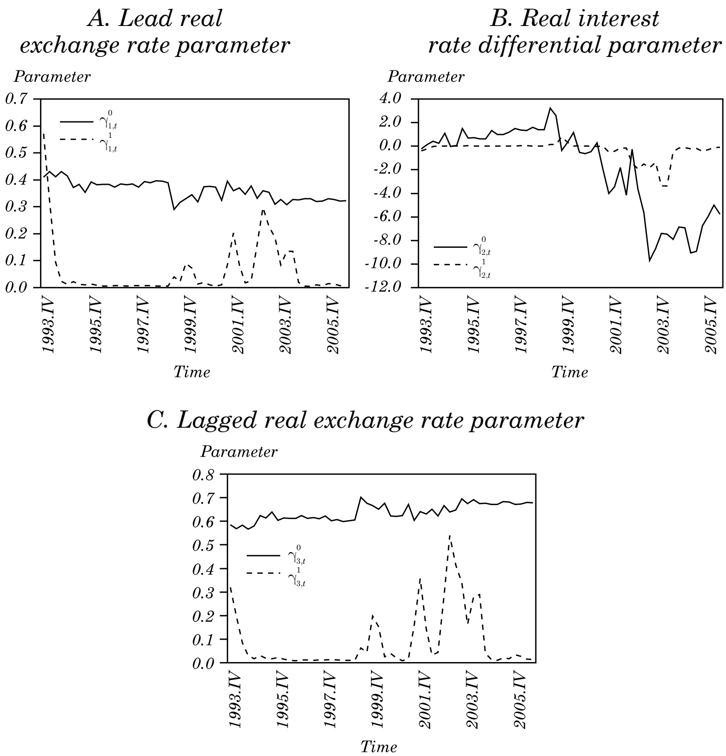
Figure 6. Time-Varying Parameters Estimated for the Phillips Curve



Source: Authors' calculations.

Finally, parameters associated with the UIP show substantial differences depending on the state of shocks (see figure 7). There is no defined trend in any of the cases. Moreover, in the entire period of analysis, the UIP parameters are more stable in the low-volatility state than in the high-volatility state. In this latter state there are two defined periods: one covering the decade of the 1990s, during which the parameters showed greater stability, and another from 2000 onward, in which the parameters increased their variability and magnitude substantially, in comparison with the first period. This change could be explained by the adoption of a completely flexible exchange rate scheme in 1999. Also, the estimations suggest that the economy was experiencing a high-volatility state of shocks in the entire period, since the occurrence probability of this state did not fall below 0.7 at any time.

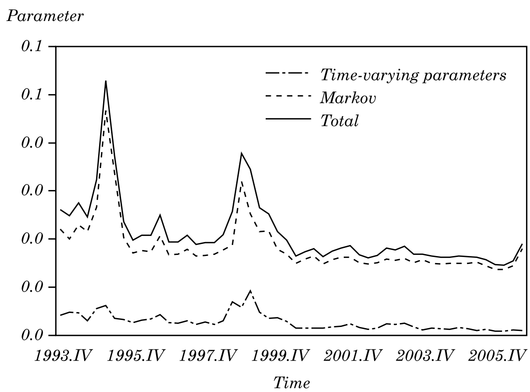
Figure 7. Time-Varying Parameters Estimated for the Real Uncovered Interest Parity



Source: Authors' calculations.

Based on the estimated parameters presented in table 5, we calculated the decomposition of the conditional variance of the forecast error. Figure 8 shows the decomposition for the set of equations associated with aggregate demand. Total uncertainty in the output gap (aggregate demand) equation has been relatively high throughout the entire period (with the output gap measured as the percentage deviation of output with respect to its trend). The forecast error variance was 0.021, on average, of which 87.6 percent was explained by uncertainty in the demand shocks and 12.4 percent by instabilities in the model parameters (see table 6). Total uncertainty registered significant spikes (almost twice the average) in the mid-1990s and in 1998–99. After 2000, however, total uncertainty declined by a little over 30 percent relative to the average observed between 1993 and 1999. We obtained similar results with the contributions of additive and multiplicative uncertainty to total uncertainty. While parameter instability contributed approximately 15 percent to total uncertainty throughout the 1990s, this contribution decreased to less than 10 percent in the period after 2000.

Figure 8. Decomposition of the Conditional Variance of the Forecast Error of the Output Gap



Source: Authors' calculations.

The decomposition of the conditional variance of the forecast error for the inflation rate equation (that is, the Phillips curve) is shown in figure 9. Results in this case are similar to those found for the output gap with respect to magnitude and behavior

Table 6. Decomposition of the Conditional Variance of the Forecast Error of the Output Gap

	Conditional variance of the forecast error			*Percentage*	
Period	*Time-varying parameters*	*Markov*	*Total*	*Time-varying parameters*	*Markov*
1993–99	0.00407	0.02173	0.02580	15.7	84.3
2000–06	0.00160	0.01535	0.01696	9.3	90.7
Total sample	0.00279	0.01842	0.02121	12.4	87.6

Source: Authors' calculations.

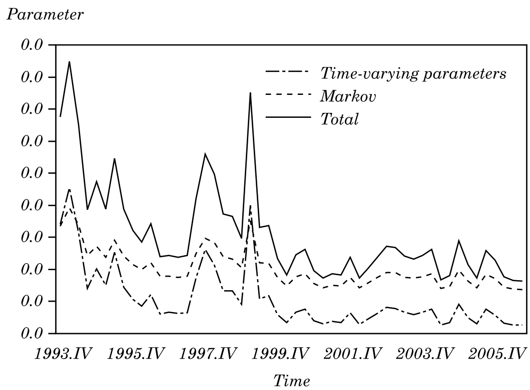
Table 7. Decomposition of the Conditional Variance of the Forecast Error of the Inflation Rate

	Conditional variance of the forecast error			*Percentage*	
Period	*Time-varying parameters*	*Markov*	*Total*	*Time-varying parameters*	*Markov*
1993–99	0.00852	0.01235	0.02087	37.5	62.5
2000–06	0.00260	0.00818	0.01078	23.2	76.8
Total sample	0.00545	0.01019	0.01563	30.1	69.9

Source: Authors' calculations.

(principally for the decade of the 90s). Total uncertainty associated with the inflation rate was 0.015, on average, for the entire period, of which 69.9 percent is explained by uncertainty in the supply shocks and 30.1 percent by parameter instability (see table 7). As in the case of the output gap, the two recurrent periods of high uncertainty are in the mid-1990s and 1998–99, when uncertainty reached levels more than twice the observed average for the entire period. Although additive uncertainty explains the largest share of total uncertainty for the whole period, the contribution pattern is briefly reverted during Asian crisis, when parameter uncertainty is most relevant. Total inflation uncertainty decreased over time, as in the case of the output gap, while the contribution of additive uncertainty increased with time.

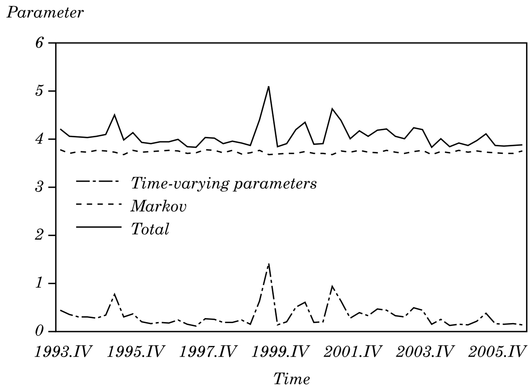
Figure 9. Decomposition of the Conditional Variance of the Forecast Error of the Inflation Rate



Source: Authors' calculations.

Finally, figure 10 presents the decomposition of the conditional variance of the forecast error associated with the real exchange rate equation. Total uncertainty, measured by the variance, was quite important throughout the period (approximately 4.1, on average) and is basically explained (92 percent) by uncertainty in the UIP shocks or uncertainty in the risk premium that captures the effects of the unobservables of the exchange market sentiments. Total uncertainty does not follow a defined pattern over time (see table 8).

Figure 10. Decomposition of the Conditional Variance of the Forecast Error of the Real Exchange Rate



Source: Authors' calculations.

In sum, overall uncertainty is dominated by additive uncertainty in all three sets of equations (namely, the output gap, inflation, and the real exchange rate). Moreover, the results of estimating the behavioral equations (aggregate demand and aggregate supply) suggest that the variance of shocks is state dependent and that such states could be considered as high-volatility periods and low-volatility periods in the shocks. For these two sets of equations, total uncertainty has consistently declined in the current decade, resulting in a rather long period of stability (so far) that coincides with the establishment of a full-fledged inflation-targeting framework for the conduct of the Chilean monetary policy and an explicit rule for setting fiscal policy. In the 1990s, in contrast, total uncertainty increased substantially in the output gap and the inflation rate, with a clear division into the two states in the variance of shocks. This also indicates that during these periods the Chilean economy experienced a high-volatility state of shocks. Finally, uncertainty in the real exchange rate is basically explained by the exchange market shocks, and it has not decreased over time like inflation and the output gap.

We use bootstrapping to verify whether the differences between the variance of the forecast error due to additive uncertainty and that due to multiplicative uncertainty are statistically significant and whether the assumption of Gaussian errors in the estimation

Table 8. Decomposition of the Conditional Variance of the Forecast Error of the Real Exchange Rate

	Conditional variance of the forecast error			*Percentage*	
Period	*Time-varying parameters*	*Markov*	*Total*	*Time-varying parameters*	*Markov*
1993–99	0.32816	3.73569	4.06385	8.8	92.3
2000–06	0.32701	3.72663	4.05364	8.8	92.1
Total sample	0.32756	3.73099	4.05855	8.8	92.2

Source: Authors' calculations.

introduces significant biases.[32] The most important findings of this exercise can be summarized as follows (for details on the results see appendix E): first, while the average bootstrap estimates differ from estimates based on the assumption of Gaussian errors, the bias does not seem to be important in magnitude; and second, the bootstrap estimations confirm the observed trends in total uncertainty (figures 8 to 10), as well as the statistical significance of the differences in the decomposition of the variance.

To conclude this subsection, we present a robustness analysis for the decomposition of the forecast error variance. In section 2, we found evidence of important differences in the estimation of the output gap when we tested five output detrending methods. Given that the aggregate demand and the Phillips curve equations contemplate an output gap measure for their estimation, measurement errors in the estimation of this variable will be part of the additive and multiplicative uncertainty without any possibility of discrimination.[33] Tables 9 and 10 show the results of the decomposition of uncertainty into its two sources, additive and multiplicative, for these two equations and for each of the five filters used in section 2. The first row of both tables shows the decomposition presented in the analysis of this subsection, where the gap was calculated using the HP filter; this represents our benchmark. In the case of the output gap (table 9), total uncertainty is generally quite similar for all filters, and differences arise in the contribution of each type of uncertainty to total uncertainty, as expected. However, all the detrending methods maintain additive uncertainty as the most important source of uncertainty (with a contribution ranging from a minimum of 84.7 percent with the BK filter and a maximum of 90.0 percent with the Clark filter). With respect to the inflation rate (table 10), the difference among filters can be observed in both the estimation of total uncertainty and the contributions of each type of uncertainty to the total. In the former

32. Our bootstrap resampling followed the methodologies of Stoffer and Wall (1991) and Psaradakis (1998) for state-space models using the Kalman filter and for the sampling of errors with Markov regime changes, respectively.

33. When the measurement error is associated with the dependent variable, as in the case of aggregate demand, the estimated parameters will still be unbiased and consistent. The measurement error will be captured by the regression error. When the measurement error is associated with one or more independent variables, as is the case with the Phillips curve, the parameters will be biased and inconsistent. Although the measurement error has different effects depending on the type of variable on which it operates, this could have implications for the decomposition of uncertainty (through the error or the magnitude of the parameters).

Table 9. Robustness Analysis for the Decomposition of the Conditional Variance of the Forecast Error of the Output Gap

	Conditional variance of the forecast error			*Percentage*	
Filter	*Time-varying parameters*	*Markov*	*Total*	*Time-varying parameters*	*Markov*
Hodrick-Prescott	0.00279	0.01842	0.02121	13.2	86.8
Baxter-King	0.00314	0.01734	0.02048	15.3	84.7
Christiano-Fitzgerald	0.00304	0.01733	0.02037	14.9	85.1
Quadratic trend	0.00287	0.01901	0.02189	13.1	86.9
Clark	0.00200	0.01803	0.02003	10.0	90.0

Source: Authors' calculations.

Table 10. Robustness Analysis for the Decomposition of the Conditional Variance of the Forecast Error of the Inflation Rate

	Conditional variance of the forecast error			*Percentage*	
Filter	*Time-varying parameters*	*Markov*	*Total*	*Time-varying parameters*	*Markov*
Hodrick-Prescott	0.00545	0.01019	0.01563	34.8	65.2
Baxter-King	0.00385	0.00988	0.01374	28.0	72.0
Christiano-Fitzgerald	0.00393	0.01006	0.01398	28.1	71.9
Quadratic trend	0.00761	0.01514	0.02274	33.4	66.6
Clark	0.00504	0.01397	0.01901	26.5	73.5

Source: Authors' calculations.

case, the estimations are in the range of 0.01374 and 0.02274 with the BK filter and the quadratic trend, respectively, while the contribution of additive uncertainty varies from 66.6 percent with the BK filter to 73.5 percent with the Clark filter. In this case, additive uncertainty again explains total uncertainty of inflation, regardless of the method used to estimate the output gap. These results strengthen our earlier conclusions regarding the importance of additive uncertainty for the Chilean economy.

4. Final Remarks

Current macroeconomic policy in Chile is world-class. The Central Bank of Chile has been operating within a full-fledged inflation-targeting framework since 1999–2000, while fiscal policy has been bounded by an explicit budget rule since 2001 that eliminates procyclical influences. As a result, inflation has remained within the target range most of the time, and economic activity has grown steadily between 2 and 6 percent annually (with no recessions or booms whatsoever). This stable period appears in our findings in the sense that overall uncertainty concerning monetary policy declined in the first seven years of the current decade. Moreover, uncertainty attributed to shocks has played a greater role, while uncertainty linked to unstable parameters has diminished, in the case of both inflation and the output gap, as could be expected. However, the prominence of additive uncertainty characterizes the entire period, including both the tranquil current decade and the more volatile 1990s. This means that investigating the (stochastic) nature of shocks affecting the Chilean economy should be high on the research agenda of the Central Bank.

The full-fledged inflation-targeting scheme applied since 1999 incorporated a floating exchange rate and no explicit or implicit target for the exchange rate (as was loosely the case during most of the 1990s). This important policy innovation has left the exchange rate as the main adjustment variable—a sort of fuse. This feature shows in our results: parameters in the exchange rate equation are less stable in the current decade than they were in the 1990s.

Our findings assume that there is no model uncertainty, so the only uncertainties relevant for the conduct of monetary policy are those in the shocks and parameters. Our results must therefore be interpreted with caution. To analyze uncertainty in the model, we could estimate the behavioral equations of the economy using the methodology presented in this paper but different specifications. This approach could be used to verify whether the decomposition of the uncertainty found here holds.[34] We leave this exercise pending for future research.

Finally, results on uncertainty about the quality and completeness of output gap data indicate that using the Hodrick-Prescott filter based on real-time data could be misleading. The Central Bank of Chile should thus consider a wide spectrum of filters for detrending real activity data.[35] More importantly, an ample menu of proxy variables should be employed to check the economy's temperature when making monetary policy decisions. The literature suggests that monetary policy rules that consider, for example, output growth rates or unemployment level rates (as opposed to the output gap) are more "immune" to this type of uncertainty.

34. This exercise was done only with the UIP under two specifications: the original equation of Svensson (2000) and Al-Eyd and Karasulu (2008) and the equation that includes the backward-looking term to allow deviations from the parity (presented here). We found that although the behavior of the parameters and the magnitude of total uncertainty change significantly, the decomposition of the uncertainty is not altered (additive uncertainty is maintained as the principal factor of uncertainty).

35. It should also use some alternative methodologies for estimating potential output, as it currently does.

Appendix A

Robustness Test for the Reliability of Real-Time Estimates using Seasonally Unadjusted Data and Seasonal Dummies

The tables presented in this appendix provide additional details on the results obtained in the estimation of the output gap with real-time data using seasonally unadjusted data and seasonally adjusted data through seasonal dummy variables.

Table A1. Descriptive Statistics of the Total Revisions in the Output Gap Using Seasonally Unadjusted Data

Filter	*Mean*	*Standard deviation*	*Minimum*	*Maximum*	*AR(1)*
Hodrick-Prescott	–0.005	0.015	–0.036	0.031	0.331
Baxter-King	0.006	0.007	–0.008	0.023	0.722
Christiano-Fitzgerald	–0.013	0.009	–0.029	0.005	0.836
Quadratic trend	–0.011	0.021	–0.050	0.033	0.676
Clark	0.001	0.006	–0.014	0.010	0.023

Source: Authors' calculations.

Table A2. Descriptive Statistics of the Reliability Measures for the Alternative Distinct Filters Using Seasonally Unadjusted Data[a]

Filter	*Correlation*	*N/S*	*OPSIGN*	*XSIZE*
Hodrick-Prescott	0.773	0.754	0.286	0.536
Baxter-King	0.529	0.958	0.286	0.464
Christiano-Fitzgerald	0.244	1.290	0.393	0.821
Quadratic trend	0.846	0.642	0.179	0.393
Clark	0.963	0.290	0.036	0.107

Source: Authors' calculations.

a. The first column presents the correlation between the final and real-time series for each detrending method. The N/S indicator is the ratio of the standard deviation of the revision to the standard deviation of the final estimate of the output gap; it approximates the noise-to-signal ratio. The OPSING indicator shows the frequency with which the real-time estimates of the output gap reveal a different sign than the final estimates. The XSIZE indicator shows the frequency with which the absolute value of the revision exceeds the absolute value of the final estimates of the output gap.

Table A3. Descriptive Statistics of the Total Revisions in the Output Gap Using Seasonal Dummies

Filter	*Mean*	*Standard deviation*	*Minimum*	*Maximum*	*AR(1)*
Hodrick-Prescott	0.002	0.017	–0.034	0.031	0.260
Baxter-King	0.008	0.007	–0.002	0.019	0.874
Christiano-Fitzgerald	–0.011	0.010	–0.029	0.002	0.942
Quadratic trend	–0.004	0.024	–0.051	0.046	0.521
Clark	0.005	0.007	–0.013	0.017	–0.063

Source: Authors' calculations.

Table A4. Descriptive Statistics of the Reliability Measures for the Alternative Filters Using Seasonal Dummies[a]

Filter	*Correlation*	*N/S*	*OPSIGN*	*XSIZE*
Hodrick-Prescott	0.413	1.044	0.321	0.429
Baxter-King	0.646	0.772	0.321	0.500
Christiano-Fitzgerald	0.312	1.031	0.357	0.571
Quadratic trend	0.745	0.771	0.179	0.321
Clark	0.932	0.367	0.071	0.214

Source: Authors' calculations.

a. The first column presents the correlation between the final and real-time series for each detrending method. The N/S indicator is the ratio of the standard deviation of the revision to the standard deviation of the final estimate of the output gap; it approximates the noise-to-signal ratio. The OPSING indicator shows the frequency with which the real-time estimates of the output gap reveal a different sign than the final estimates. The XSIZE indicator shows the frequency with which the absolute value of the revision exceeds the absolute value of the final estimates of the output gap.

Appendix B
Estimation Based on the Kalman Filter and the EM Algorithm

Our estimation approach follows the two-stage procedure proposed by Kim (2004, 2006). The first stage, described in the main text, consists in estimating a model that instrumentalizes the endogenous variables using the maximum log-likelihood method based on the error forecast and the conventional Kalman filter. The second stage is based on maximum log-likelihood techniques that combine the Kalman filter and the EM algorithm proposed by Hamilton (1989, 1990). The latter estimation is defined by the following series of equations (Kim and Nelson, 1999):

Kalman Filter

$$\beta_{t|t-1}^{(i,j)}, \mathbf{P}_{t|t-1}^{(i,j)}, \tau_{t|t-1}^{(i,j)}, f_{t|t-1}^{(i,j)}, \mathbf{H}_{t|t-1}^{(i,j)}.$$

Hamilton's EM Algorithm

$$\Pr\left[S_t, S_{t-1} \mid \psi_{t-1}\right] = \Pr\left[S_t, S_{t-1}\right] \Pr\left[S_{t-1} \mid \psi_{t-1}\right];$$

$$f(x_t \mid \psi_{t-1}) = \sum_{S_t} \sum_{S_{t-1}} f(x_t \mid S_t, S_{t-1}, \psi_{t-1}) \Pr\left[S_t, S_{t-1} \mid \psi_{t-1}\right];$$

$$l(\theta) = l(\theta) + \ln\left[f(x_t \mid \psi_{t-1})\right];$$

$$\Pr\left[S_t, S_{t-1} \mid \psi_{t-1}\right] = \frac{f(x_t, S_t, S_{t-1}, \psi_{t-1})}{f(x_t \mid \psi_{t-1})} = \frac{f(x_t \mid S_t, S_{t-1}, \psi_{t-1}) \Pr\left[S_t, S_{t-1} \mid \psi_{t-1}\right]}{f(x_t \mid \psi_{t-1})}$$

$$\Pr\left[S_t \mid \psi_t\right] = \sum_{S_{t-1}} \Pr\left[S_t, S_{t-1} \mid \psi_t\right].$$

Approximations

$$\beta_{t|t}^{j}=\frac{\sum_{i=1}^{2}\Pr\left[S_{t-1}=i,S_{t}=j\mid\psi_{t}\right]\beta_{t|t}^{(i,j)}}{\Pr\left[S_{t}=j\mid\psi_{t}\right]};$$

$$\mathbf{P}_{t|t}^{j}=\frac{\sum_{i=1}^{2}\Pr\left[S_{t-1}=i,S_{t}=j\mid\psi_{t}\right]\left[\mathbf{P}_{t|t}^{(i,j)}+\left(\beta_{t|t}^{j}-\beta_{t|t}^{(i,j)}\right)\left(\beta_{t|t}^{j}-\beta_{t|t}^{(i,j)}\right)'\right]}{\Pr\left[S_{t}=j\mid\psi_{t}\right]}.$$

Log-likelihood function

$$l(\theta)=\sum_{t=1}^{T}\ln\left[f\left(x_{t}\mid\psi_{t-1}\right)\right].$$

Appendix C
Kalman Filter with Endogenous Regressors

Kim (2006) uses the following series of equations to describe the Kalman filter with endogenous regressors.

$$\boldsymbol{\beta}_{t|t-1} = E\left(\boldsymbol{\beta}_t \mid \mathbf{w}_t, \mathbf{v}_t, \xi_t^*, \psi_{t-1}\right) = \boldsymbol{\beta}_{t-1|t-1};$$

$$\mathbf{P}_{t|t-1} = \operatorname{var}\left(\boldsymbol{\beta}_t \mid \mathbf{w}_t, \mathbf{v}_t, \xi_t^*, \psi_{t-1}\right) = \mathbf{P}_{t-1|t-1} + \mathbf{Q}_\eta;$$

$$\tau_{t|t-1} = x_t - E\left(x_t \mid \mathbf{w}_t, \mathbf{v}_t, \xi_t^*, \psi_{t-1}\right) = x_t - (\mathbf{w}_t, \mathbf{v}_t)' \boldsymbol{\beta}_{t-1|t-1} - \xi_t^{*\prime} \rho \sigma_\varepsilon;$$

$$\mathbf{H}_{t|t-1} = \operatorname{var}\left(x_t \mid \mathbf{w}_t, \mathbf{v}_t, \xi_t^*, \psi_{t-1}\right) = (\mathbf{w}_t, \mathbf{v}_t)' \mathbf{P}_{t|t-1} (\mathbf{w}_t, \mathbf{v}_t) + (1 - \rho'\rho) \sigma_\varepsilon^2;$$

$$\boldsymbol{\beta}_{t|t} = E\left(\boldsymbol{\beta}_t \mid \mathbf{w}_t, \mathbf{v}_t, \xi_t^*, \psi_{t-1}\right) = \boldsymbol{\beta}_{t|t-1} + \mathbf{P}_{t|t-1} (\mathbf{w}_t, \mathbf{v}_t) \mathbf{H}_{t|t-1}^{-1} \tau_{t|t-1};$$

$$\mathbf{P}_{t|t} = \operatorname{var}\left(\boldsymbol{\beta}_t \mid \mathbf{w}_t, \mathbf{v}_t, \xi_t^*, \psi_{t-1}\right) = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1} (\mathbf{w}_t, \mathbf{v}_t) \mathbf{H}_{t|t-1}^{-1} (\mathbf{w}_t, \mathbf{v}_t)' \mathbf{P}_{t|t-1}.$$

Appendix D
Log-Likelihood Function

The log-likelihood function defined by Kim and Nelson (1999) is as follows:

$$f(x_t \mid \psi_{t-1}) = \sum_{i=1}^{2}\sum_{j=1}^{2} f(x_t, S_t = i, S_{t-1} = j \mid \psi_{t-1})$$
$$= \sum_{i=1}^{2}\sum_{j=1}^{2} f(x_t \mid S_t = i, S_{t-1} = j \mid \psi_{t-1}) \Pr\left[S_t = i, S_{t-1} = j \mid \psi_{t-1}\right],$$

where

$$f(x_t \mid S_t = i, S_{t-1} = j, \psi_{t-1}) = (2\pi)^{-\frac{N}{2}} \mid f_{t|t-1}^{(i,j)} \mid^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\tau_{t|t-1}^{(i,j)}{}' f_{t|t-1}^{(i,j)-1}\tau_{t|t-1}^{(i,j)}\right\}.$$

Appendix E

Bootstrap of the Decomposition of the Conditional Variance of the Forecast Error

Table E1 presents the results obtained from the bootstrap of the decomposition of the conditional variance of the forecast error for the three models (mean estimation and 95 percent confidence intervals). The table also shows, for comparison purposes, the previous results found under the assumption of Gaussian errors in the estimation. The bootstrap resampling followed the methodologies of Stoffer and Wall (1991) for state-space models that use the Kalman filter and Psaradakis (1998) for the sampling of errors with Markov regime changes.

Table E1. Bootstrap Decomposition of the Conditional Variance of the Forecast Error

	Gaussian maximum likelihood			*Bootstrap*					
				Time-varying parameters		*Markov*		*Total*	
Variable and period	*Time-varying parameters*	*Markov*	*Total*	*Mean*	*0.95 confidence interval*	*Mean*	*0.95 confidence interval*	*Mean*	*0.95 confidence interval*
A. Output gap									
1993–95	0.00424	0.02566	0.02990	0.00585	[0.00572, 0.00598]	0.05667	[0.05542, 0.05790]	0.06251	[0.06119, 0.06384]
1996–98	0.00353	0.01881	0.02234	0.00548	[0.00533, 0.00564]	0.02330	[0.02264, 0.02401]	0.02878	[0.02796, 0.02961]
1999–2006	0.00208	0.01616	0.01824	0.00193	[0.00188, 0.00197]	0.01807	[0.01749, 0.01870]	0.02000	[0.01938, 0.02066]
Full sample	0.00279	0.01842	0.02121	0.00342	[0.00334, 0.00351]	0.02596	[0.02524, 0.02671]	0.02938	[0.02860, 0.03020]
B. Inflation rate									
1993–95	0.01172	0.01428	0.02599	0.01204	[0.01199, 0.01207]	0.06555	[0.02062, 0.15871]	0.07758	[0.03267, 0.18638]
1996–98	0.00612	0.01099	0.01711	0.00588	[0.00586, 0.00590]	0.04010	[0.01541, 0.09914]	0.04598	[0.02129, 0.09725]
1999–2006	0.00337	0.00869	0.01205	0.00289	[0.00288, 0.00291]	0.02381	[0.01130, 0.04950]	0.02670	[0.01420, 0.05276]
Full sample	0.00545	0.01019	0.01563	0.00516	[0.00514, 0.00518]	0.03479	[0.01386, 0.07986]	0.03996	[0.01903, 0.08616]
C. Real exchange rate									
1993–95	0.54577	10.54444	11.09022	0.81944	[0.77777, 0.86192]	9.21296	[9.02628, 9.40670]	10.03222	[9.83458, 10.24011]
1996–98	0.37324	10.46275	10.83625	0.74982	[0.70561, 0.79519]	9.20454	[9.01600, 9.39305]	9.95438	[9.75733, 10.16217]
1999–2006	0.56164	10.58610	11.14777	0.91215	[0.86287, 0.96336]	9.19028	[8.99863, 9.38625]	10.10250	[9.89786, 10.31116]
Full sample	0.51542	10.55042	11.06592	0.85864	[0.81185, 0.90699]	9.19750	[9.00743, 9.39136]	10.05616	[9.85448, 10.26448]

Source: Authors' calculations.

References

Al-Eyd, A.J. and M. Karasulu. 2008. "Ambition versus Gradualism in Disinflation Horizons under Bounded Rationality: The Case of Chile." Discussion paper 308. London: National Institute of Economic and Social Research.

Balvers, R. and T. Cosimano. 1994. "Inflation Variability and Gradualist Monetary Policy." *Review of Economic Studies* 61(4): 721–38.

Basar, T. and M. Salomon. 1990. "Credibility and the Value of Information Transmission in a Model of Monetary Policy and Inflation." *Journal of Economic Dynamics and Control* 14(1): 97–116.

Baxter, M. and R.G. King. 1999. "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series." *Review of Economics and Statistics* 81(4): 575–93.

Beck, G.W. and V. Wieland. 2002. "Learning and Control in a Changing Economic Environment." *Journal of Economic Dynamics and Control* 26(9–10): 1359–77.

Bernhardsen, T., O. Eitrheim, A.S. Jore and O. Roisland. 2005. "Real-Time Data for Norway: Challenges for Monetary Policy." *North American Journal of Economics and Finance* 16(3): 333–49.

Bertocchi, G. and M. Spagat. 1993. "Learning, Experimentation, and Monetary Policy." *Journal of Monetary Economics* 32(1): 169–83.

Blinder, A.S. 1998. *Central Banking in Theory and Practice*. MIT Press.

Brainard, W.C. 1967. "Uncertainty and the Effectiveness of Policy." *American Economic Review* 57(2): 411–25.

Calvo, G. 1983. "Staggered Prices in a Utility Maximization Framework." *Journal of Monetary Economics* 12(4): 383–98.

Caputo, R., F. Liendo and J.P. Medina. 2006. "New Keynesian Models for Chile during the Inflation Targeting Regime: A Structural Approach." Working paper 402. Santiago: Central Bank of Chile.

Céspedes, L.F., M. Ochoa and C. Soto. 2005. "The New Keynesian Phillips Curve in an Emerging Market Economy: The Case of Chile." Working Paper 355. Santiago: Central Bank of Chile.

Christiano, L.J. and T.J. Fitzgerald. 2003. "The Band Pass Filter." *International Economic Review* 44(2): 435–65.

Clarida, R., J. Galí and M. Gertler. 1998. “Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory.” Working paper 6442. Cambridge, Mass.: National Bureau of Economic Research.

————. 1999. “The Science of Monetary Policy: A New Keynesian Perspective.” *Journal of Economic Literature* 37(4): 1661–707.

Clark, P.K. 1987. “The Cyclical Component of U.S. Economic Activity.” *Quarterly Journal of Economics* 102(4): 797–814.

Cogley, T. and T.J. Sargent. 2002. “Evolving Post-World War II U.S. Inflation Dynamics.” In *NBER Macroeconomics Annual 2001*, edited by B. Bernanke and K. Rogoff, 331–73. MIT Press.

————. 2005. “Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US.” *Review of Economic Dynamics* 8(2): 262–302.

Corbo, V. 2002. “Monetary Policy in Latin America in the 90s.” In *Monetary Policy: Rules and Transmission Mechanisms,* edited by N. Loayza and K. Schmidt-Hebbel, 117–66. Santiago: Central Bank of Chile.

Craine, R. 1979. “Optimal Monetary Policy with Uncertainty.” *Journal of Economic Dynamics and Control* 1(1): 59–83.

Croushore, D. and T. Stark. 2001. “A Real-Time Data Set for Macroeconomists.” *Journal of Econometrics* 105(1): 111–30.

De Grauwe, P. and M.A. Senegas. 2006. “Monetary Policy Design and Transmission Asymmetry in EMU: Does Uncertainty Matter?” *European Journal of Political Economy* 22(4): 787–808.

Feldstein, M. 2003. “Monetary Policy in an Uncertain Environment.” Working paper 9969. Cambridge, Mass.: National Bureau of Economic Research.

Fuentes, R., F. Gredig and M. Larraín. 2007. “Estimating the Output Gap for Chile.” Working paper 455. Santiago: Central Bank of Chile.

Galí, J. and M. Gertler. 1999. “Inflation Dynamics: A Structural Econometric Analysis.” *Journal of Monetary Economics* 44(2): 195–222.

Galí, J., M. Gertler and J.D. López-Salido. 2005. “Robustness of the Estimates of the Hybrid New Keynesian Phillips Curve.” *Journal of Monetary Economics* 52(6): 1107–18.

Gallego, F. and C. Johnson. 2001. “Teorías y métodos de medición del producto de tendencia: una aplicación al caso de Chile “ *Economía Chilena* 4(2): 27–58.

Ghysels, E., N.R. Swanson and M. Callan. 2002. "Monetary Policy Rules with Model and Data Uncertainty." *Southern Economic Journal* 69(2): 239–65.

Giannoni, M.P. 2002. "Does Model Uncertainty Justify Caution? Robust Optimal Monetary Policy in a Forward-Looking Model." *Macroeconomic Dynamics* 6(1): 111–44.

Gonzalez, F. and A. Rodriguez. 2004. "Robust Control: A Note on the Response of the Control to Changes in the 'Free' Parameter Conditional on the Character of Nature." *Computational Economics* 24(3): 223–38.

Gredig, F. 2007. "Índices podados como medidas de tendencia para el IMACEC." Working paper 414. Santiago: Central Bank of Chile.

Hall, S., C. Salmon, T. Yates and N. Batini. 1999. "Uncertainty and Simple Monetary Policy Rules: An Illustration for the United Kingdom." Working papers 96. London: Bank of England.

Hamilton, J.D. 1989. "A New Approach to the Economic Analysis of Non-Stationary Time Series and The Bussines Cycles." *Econometrica* 57(2): 357–84.

————. 1990. "Analysis of Time Series Subject to Changes in Regime." *Journal of Econometrics* 45(1–2): 39–70.

Heckman, J. 1976. "The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models." *Annals of Economic and Social Measurement* 5(4): 475–92.

Hodrick, R.J. and E.C. Prescott. 1997. "Postwar U.S. Business Cycles: An Empirical Investigation." *Journal of Money, Credit, and Banking* 29(1): 1–16.

Holly, S. and A. Hughes Hallet. 1989. *Optimal Control, Expectations, and Uncertainty*. Cambridge University Press.

Holt, C.C. 1962. "Linear Decision Rules for Economic Stabilization and Growth." *Quarterly Journal of Economics* 76(1): 20–45.

Isard, P., D. Laxton and A.C. Eliasson. 1999. "Simple Monetary Policy Rules under Model Uncertainty." *International Tax and Public Finance* 6(4): 537–77.

Kim, C.J. 1993. "Sources of Monetary Growth Uncertainty and Economic Activity: The Time-Varying-Parameter Model with Heteroskedastic Disturbances." *Review of Economics and Statistics* 75(3): 483–92.

————. 2004. "Markov-Switching Models with Endogenous Explanatory Variables." *Journal of Econometrics* 122(1): 127–36.

————. 2006. "Time-Varying Parameter Models with Endogenous Regressors." *Economics Letters* 91(1): 21–26.

Kim, C.J. and C.R. Nelson. 1999. *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. MIT Press.

————. 2006. "Estimation of a Forward-Looking Monetary Policy Rule: A Time-Varying Parameter Model Using ex post Data." *Journal of Monetary Economics* 53(8): 1949–66.

Kozicki, S. 2004. "How Do Data Revisions Affect the Evaluation and Conduct of Monetary Policy?" *Economic Review* (first quarter): 5–38. Federal Reserve Bank of Kansas.

Kuttner, K.N. 1994. "Estimating Potential Output as a Latent Variable." *Journal of Business and Economic Statistics* 12(3): 361–68.

Martin, B. and C. Salmon. 1999. "Should Uncertain Monetary Policy-Makers Do Less?" Working paper 99. London: Bank of England.

Meyer, L.H., E.T. Swanson and V.W. Wieland. 2001. "NAIRU Uncertainty and Nonlinear Policy Rules." *American Economic Review* 92(2): 226–31.

Onatski, A. and J.H. Stock. 2002. "Robust Monetary Policy under Model Uncertainty in a Small Model of the U.S. Economy." *Macroeconomic Dynamics* 6(1): 85–110.

Orphanides, A. 2001. "Monetary Policy Rules Based on Real-Time Data "*American Economic Review* 91(4): 964–85.

Orphanides, A. and S. van Norden. 1999. "The Reliability of Output Gap Estimates in Real Time." Finance and economics discussion paper 1999-38. Washington: Board of Governors of the Federal Reserve System.

Orphanides, A. and J.C. Williams. 2002. "Imperfect Knowledge, Inflation Expectations, and Monetary Policy." Working paper 2002-04. Federal Reserve Bank of San Francisco.

Orphanides, A, R.D. Porter, D. Reifschneider, R.J. Tetlow and F. Finan. 2000. "Errors in the Measurement of the Output Gap and the Design of Monetary Policy." *Journal of Economics and Business* 52(1–2): 117–41.

Phillips, A.W. 1954. "Stabilisation Policy in a Closed Economy." *Economic Journal* 64(254): 290–323.

Psaradakis, Z. 1998. "Bootstrap-Based Evaluation of Markov-Switching Time Series Models." *Econometric Reviews* 17(3): 275–88.

Roberts, J.M. 1995. "New Keynesian Economics and the Phillips Curve." *Journal of Money, Credit, and Banking* 27(4): 975–84.

———. 2001. "How Well Does the New Keynesian Sticky-Price Model Fit the Data?" Finance and economics discussion paper 2001-13. Washington: Board of Governors of the Federal Reserve System.

Rudebusch, G.D. 2001. "Is the Fed Too Timid? Monetary Policy in an Uncertain World." *Review of Economics and Statistics* 83(2): 203–17.

Sack, B. 2000. "Does the Fed Act Gradually? A VAR Analysis." *Journal of Monetary Economics* 46(1): 229–56.

Sargent, T.J. 1998. "Discussion of 'Policy Rules for Open Economies' by Laurence Ball." Presented at the NBER Conference on Monetary Policy Rules. National Bureau of Economic Research, Islamorada, Florida, 15–17 January.

Schellekens, P. 2002. "Caution and Conservatism in the Making of Monetary Policy." *Journal of Money, Credit, and Banking* 34(1): 160–77.

Semmler, W., A. Greiner and W. Zhang. 2005. *Monetary and Fiscal Policies in the Euro-Area: Macro Modeling, Learning, and Empirics*. Amsterdam: Elsevier Science Publisher.

Sims, C.A. 2002. "Evolving Post World War II U.S. Inflation Dynamics: Comment." in *NBER Macroeconomics Annual 2001*, edited by B. Bernanke and K. Rogoff, 373–79. MIT Press.

Sims, C.A. and T. Zha. 2006. "Were There Regime Switches in U.S. Monetary Policy?" *American Economic Review* 96(1): 54–81.

Söderström, U. 2002. "Monetary Policy with Uncertain Parameters." *Scandinavian Journal of Economics* 14(1): 125–45.

St-Amant, P. and S. van Norden. 1998. "Measurement of the Output Gap: A Discussion of Recent Research at the Bank of Canada." Technical report 79. Ottawa: Bank of Canada.

Stoffer, D.S. and K.D. Wall. 1991. "Bootstrapping State-Space Models: Gaussian Maximum Likelihood Estimation and the Kalman Filter." *Journal of the American Statistical Association* 86(416): 1024–33.

Svensson, L.E.O. 1997. "Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets." *European Economic Review* 41(6): 1111–46.

———. 1999. "Inflation Targeting: Some Extensions." *Scandinavian Journal of Economics* 101(3): 337–61.

———. 2000. "Open-Economy Inflation Targeting." *Journal of International Economics* 50(1): 155–83.

Tetlow, R.J. and P. von zur Muehlen. 2001. "Simplicity versus Optimality: The Choice of Monetary Policy Rules when Agents Must Learn." *Journal of Economic Dynamics and Control* 25(1–2): 245–79.

Theil, H. 1964. *Optimal Decision Rules for Government and Industry*. Rand McNally.

Walsh, C.E. 2004. "Implications of a Changing Economic Structure for the Strategy of Monetary Policy." Working paper 1023. University of California at Santa Cruz, Center for International Economics.

Wieland, V. 2000. "Monetary Policy, Parameter Uncertainty and Optimal Learning." *Journal of Monetary Economics* 46(1): 199–228.

Zampolli, F. 2006. "Optimal Monetary Policy in a Regime-Switching Economy: The Response to Abrupt Shifts in Exchange Rate Dynamics." *Journal of Economic Dynamics and Control* 30(9–10): 1527–67.

Zhang, W. and W. Semmler. 2005. "Monetary Policy Rules under Uncertainty: Empirical Evidence, Adaptive Learning, and Robust Control." *Macroeconomic Dynamics* 9(5): 651–81.

Inflation Dynamics in a Small Open Economy Model under Inflation Targeting: Some Evidence from Chile

Marco Del Negro
Federal Reserve Bank of New York

Frank Schorfheide
University of Pennsylvania

Following the influential work of Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2003), many central banks are building and estimating dynamic stochastic general equilibrium (DSGE) models with nominal rigidities and are using them for policy analysis. This new generation of sticky price (and wage) models typically emphasizes that relative price distortions caused by firms' partial inability to respond to changes in the aggregate price level lead to an inefficient use of factor inputs and, in turn, to welfare losses. In such an environment, monetary policy can partially offset these relative price distortions by stabilizing aggregate inflation. The policy problem is more complicated in an open economy environment, because domestic price movements are tied to exchange rate and terms-of-trade movements.

DSGE models can be used at different stages of the policymaking process. If the structure of the theoretical model is enriched to the point at which the model is able to track historical time series, DSGE models can be used as a tool to generate multivariate macroeconomic forecasts. Monetary policy in these models is typically represented by an interest rate feedback rule, and the innovations in the policy rule

We thank the Central Bank of Chile, and Mauricio Calani in particular, for providing us with the data used in the empirical analysis. Schorfheide gratefully acknowledges financial support from the Central Bank of Chile, the Alfred P. Sloan Foundation, and the National Science Foundation. The views expressed in this paper do not necessarily reflect those of the Federal Reserve Bank of New York or the Federal Reserve System.

can be interpreted as modest, unanticipated changes in monetary policy. These impulse responses can then be used to determine, say, the interest rate change necessary to keep inflation rates near a target level over the next year or two. Finally, one can use DSGE models to qualitatively or quantitatively analyze more fundamental changes in monetary policy, such as inflation versus output targeting or fixed versus floating exchange rates.

An important concern in the use of DSGE models is that some of the cross-equation restrictions generated by the economic theory are misspecified. This misspecification potentially distorts forecasts as well as policy predictions. In a series of papers (Del Negro and Schorfheide, 2004, 2008b; Del Negro and others, 2007), we develop an econometric framework that allows us to gradually relax the cross-coefficient restrictions and construct an empirical model that can be regarded as a structural vector autoregression (VAR) and that retains many of the features of the underlying DSGE model, at least to the extent that they are not grossly inconsistent with historical time series. We refer to this empirical model as DSGE-VAR.

Based on a small open economy model developed by Galí and Monacelli (2005) and modified for estimation purposes by Lubik and Schorfheide (2007), we here present estimation results for such a DSGE-VAR model for the Chilean economy, using data on output growth, inflation, interest rates, exchange rates, and terms of trade. Throughout the 1990s, monetary policy was conducted within a partial inflation-targeting regime, since the monetary authorities were targeting the exchange rate in addition to inflation. Moreover, the inflation target was evolving during this period. In September 1999, Chile entered a floating exchange rate regime, thus adopting full-fledged inflation targeting. We therefore choose to use only post-1999 data, which leaves a fairly short sample for the estimation of an empirical model for monetary policy analysis. An important advantage of the DSGE-VAR framework is that it allows us to estimate a vector autoregressive system with a short time series. Roughly speaking, this estimation augments actual observations by hypothetical observations, generated from a DSGE model, to determine the coefficients of the VAR. Over time, as more actual observations become available, our procedure will decrease or increase the fraction of actual observations in the combined sample, depending on whether the data contain evidence of model misspecification.

The empirical analysis is divided in four parts. We begin by estimating both the DSGE model and the DSGE-VAR. The DSGE-VAR produces estimates of the coefficients of the underlying theoretical model along with the VAR coefficients. Our discussion first focuses on the monetary policy rule estimates. Starting from a prior that implies a strong reaction of the Central Bank to inflation movements, we find that since 1999 the Central Bank has not reacted significantly to exchange rate or terms-of-trade movements, which is consistent with the official policy statements. In the second part, we study the fit of our small-scale DSGE model. As in our earlier work, the fit of the empirical vector autoregressive model can be improved by relaxing the theoretical cross-coefficient restrictions. More interestingly, due to the short sample size, the fraction of DSGE-model-generated observations in the mixed sample that is used for the estimation of the VAR is higher than, say, in estimations that we have conducted for the United States. Consequently, the dynamics of the DSGE-VAR closely resemble those of the underlying DSGE model, which is documented in the third part of the empirical analysis. Here, we focus specifically on how the various structural shocks affect inflation movements.

In the final part of the empirical analysis, we study the effect of changes in the monetary policy rule. Conceptually, this type of analysis is very challenging. If one believes that the DSGE model is not misspecified, then one can determine the behavioral responses of firms and households by re-solving the model under alternative policy rules. Empirical evidence of misspecification of cross-equation restrictions, however, raises questions about the reliability of the DSGE model's policy implications. In Del Negro and Schorfheide (2008b), we develop tools that allow us, under particular invariance assumptions, to check for the robustness of the DSGE model conclusions to the presence of misspecification. We apply some of these tools to explore what would happen to the variability of inflation if the Central Bank responded more or less to inflation as well as terms-of-trade movements.

A substantial amount of empirical literature explores the Chilean economy, including many of the issues analyzed in the paper: the specification of the policy rule, the dynamics of inflation, and the responses of domestic variables to external shocks.[1] For most of this literature, the estimation period comprises the 1990s, a period of convergence toward full-fledged inflation targeting (see Mishkin and Schmidt-Hebbel, 2007). Because of concerns about structural change between the early phase of inflation targeting and the current one, we

1. See Chumacero (2005) and Céspedes and Soto (2007); these two papers also provide a survey of the existing literature.

do not use the early period in the estimation. Our results, therefore, are not directly comparable with those of the previous literature. Caputo and Liendo (2005) present a very close paper to ours, in that they also estimate the Galí-Monacelli/Lubik-Schorfheide model on Chilean data. Most of their results include the 1990s, however, which makes comparisons hard. Furthermore, Caputo and Liendo (2005) use an estimate of the output gap as an observable, as opposed to the output growth rate used in this paper. They also perform subsample analysis, and one of their subsamples is close to the one used here. For that subsample, many of their results are similar to ours. Another close paper to ours is by Caputo, Liendo, and Medina (2007), who estimate a more sophisticated small open economy DSGE model using Bayesian methods on Chilean data. Again, their use of 1990s data makes the results not directly comparable. In future work it would be interesting to apply some of the techniques used in our paper to a larger-scale small open economy DSGE model.

The remainder of the paper is organized as follows. Section 1 contains a description of the small open economy model. The DSGE-VAR framework developed in Del Negro and Schorfheide (2004, 2008b) is reviewed in section 2. The data set used for the empirical analysis is discussed in section 3. Empirical results are summarized in section 4, and section 5 concludes. Detailed derivations of the DSGE model are provided in the appendix.

1. A Small Open Economy Model

We now describe a simple small open economy DSGE model for the Chilean economy. The model has been previously estimated with data from Australia, Canada, New Zealand, and the United Kingdom in Lubik and Schorfheide (2007). It is a simplified version of the model developed by Galí and Monacelli (2005). We restrict our exposition to the key equilibrium conditions, represented in log-linearized form.[2] Derivations of these equations are relegated to the appendix. All variables below are measured in percentage

2. We follow Galí and Monacelli (2005) and Lubik and Schorfheide (2007) in solving the detrended model by log-linearization around its steady state. The appendix describes the nonlinear equilibrium conditions and the log-linearization step. In the case of the Chilean economy, it is an open question whether log-linearization provides an accurate solution to the model, given that shocks are larger in size than in developed economies. We are aware of the issue, but at this stage our computational capabilities limit the extent to which we can use alternative solution methods.

deviations from a stochastic balanced growth path, induced by a technology process, Z_t, that follows a first-order autoregressive, or AR(1), process in growth rates:

$$\Delta \ln Z_t = \gamma + \hat{z}_t, \tag{1}$$

where $\hat{z}_t = \rho_z \hat{z}_{t-1} + \sigma_z \varepsilon_{z,t}$. Here, Δ denotes the temporal difference operator.

We begin with a characterization of monetary policy. We assume that monetary policy is described by an interest rate rule. The central bank adjusts its instrument in response to movements in consumer price index (CPI) inflation and output growth. Moreover, we allow for the possibility of including nominal exchange rate depreciation or terms-of-trade changes in the policy rule:

$$\hat{R}_t = \rho_R \hat{R}_{t-1} + (1-\rho_r)\left[\psi_1 \hat{\pi}_t + \psi_2 (\Delta \hat{y}_t + \hat{z}_t) + \psi_3 \Delta x_t\right] + \sigma_r \varepsilon_{r,t}. \tag{2}$$

Since $\hat{y}_t$ measures percentage deviations from the stochastic trend induced by the productivity process Z_t, output growth deviations from the mean, γ, are given by $\Delta \hat{y}_t + \hat{z}_t$. We use Δx_t to represent either exchange rate or terms-of-trade changes. To match the persistence in nominal interest rates, we include a smoothing term in the rule, with $0 \leq \rho_r < 1$. Finally, $\varepsilon_{r,t}$ is an exogenous policy shock that can be interpreted as the nonsystematic component of monetary policy.

The household behavior in the home country is described by a consumption Euler equation in which we use equilibrium conditions to replace domestic consumption and CPI inflation by a function of domestic output, $\hat{y}_t$; output in the absence of nominal rigidities (potential output), $\bar{y}_t$; and inflation associated with domestically produced goods, $\hat{\pi}_{H,t}$:

$$\hat{y}_t - \bar{y}_t = E_t\left[\hat{y}_{t+1} - \bar{y}_{t+1}\right] - (\tau + \lambda)\left(\hat{R}_t - E_t\left[\hat{\pi}_{H,t+1} + \hat{z}_{t+1}\right]\right), \tag{3}$$

where E_t is the expectation operator and where

$$\lambda = \alpha\,(2-\alpha)(1-\tau)$$

and

$$\bar{y}_t = -\frac{\lambda}{\tau}\hat{y}_t^*. \tag{4}$$

The parameter $0 < \alpha < 1$ represents the fraction of imported goods consumed by domestic households, τ is their intertemporal substitution elasticity, and $\hat{y}_t^*$ is an exogenous process that captures foreign output (relative to the level of total factor productivity). Notice that for $\alpha = 0$, equation (3) reduces to its closed economy variant.

Optimal price setting by domestic firms leads to the following Phillips curve relationship:

$$\hat{\pi}_{H,t} = \beta E_t\left[\hat{\pi}_{H,t+1}\right] + \kappa\widehat{mc}_t, \tag{5}$$

where marginal costs can be expressed as

$$\widehat{mc}_t = \frac{1}{\tau + \lambda}(\hat{y}_t - \bar{y}_t). \tag{6}$$

The slope coefficient $\kappa > 0$ reflects the degree of price stickiness in the economy. As $\kappa \to \infty$, the nominal rigidities vanish.

We define the terms of trade, $\hat{Q}_t$, as the relative price of exports in terms of imports and let $\hat{q}_t = \Delta\hat{Q}_t$. The relationship between CPI inflation and $\hat{\pi}_{H,t}$ is given by

$$\hat{\pi}_t = \hat{\pi}_{H,t} - \alpha\hat{q}_t. \tag{7}$$

Assuming that relative purchasing power parity (PPP) holds, we can express the nominal exchange rate depreciation as

$$\hat{e}_t = \hat{\pi}_t - \hat{\pi}_t^* - (1-\alpha)\hat{q}_t, \tag{8}$$

where $\hat{\pi}_t^*$ is a world inflation shock that we treat as an unobservable. An alternative interpretation, as in Lubik and Schorfheide (2006), is that $\hat{\pi}_t^*$ captures misspecification or deviations from PPP. Since the other variables in the exchange rate equation are observed, this relaxes the potentially tight cross-equation restrictions embedded in the model.

Equations (3) and (5) have been derived under the assumption of complete asset markets and perfect risk sharing, which implies that

$$\hat{q}_t = -\frac{1}{\tau + \lambda}(\Delta \hat{y}_t - \Delta \hat{y}_t^*). \tag{9}$$

This equilibrium condition clearly indicates that the terms of trade are endogenous in the model, because domestic producers have market power. Instead of imposing this condition, however, we follow the approach in Lubik and Schorfheide (2007) and specify an exogenous law of motion for the terms-of-trade movements:

$$\hat{q}_t = \rho_q \hat{q}_{t-1} + \sigma_q \varepsilon_{q,t} \tag{10}$$

In the empirical section, we provide evidence on the extent to which this assumption is supported by the data.

Equations (2) to (8) form a rational expectations system that determines the law of motion for domestic output, $\hat{y}_t$; flexible price output, $\bar{y}_t$; marginal costs $\widehat{mc}_t$; CPI inflation, $\hat{\pi}_t$; domestically produced goods inflation, $\hat{\pi}_{H,t}$; interest rates, $\hat{R}_t$; and nominal exchange rate depreciations, $\hat{e}_t$. We treat monetary policy shocks, $\varepsilon_{r,t}$, technology growth, $\hat{z}_t$, and terms-of-trade changes, $\hat{q}_t$, as exogenous. Moreover, we assume that rest-of-the-world output and inflation, $\hat{y}_t^*$ and $\hat{\pi}_t^*$, follow exogenous autoregressive processes:

$$\begin{aligned} \hat{\pi}_t^* &= \rho_{\pi^*} \hat{\pi}_{t-1}^* + \sigma_{\pi^*} \varepsilon_{\pi^*,t}; \\ \hat{y}_t^* &= \rho_{y^*} \hat{y}_{t-1}^* + \sigma_{y^*} \varepsilon_{y^*,t}. \end{aligned} \tag{11}$$

The rational expectations model described by equations (1) to (11) can be solved with standard techniques, such as Sims (2002). We collect the DSGE model parameter in the vector $\boldsymbol{\theta}$, defined as

$$\boldsymbol{\theta} = [\psi_1, \psi_2, \psi_3, \rho_r, \alpha, \beta, \tau, \rho_z, \rho_q, \rho_{\pi^*}, \rho_{y^*}, \sigma_r, \sigma_z, \sigma_q, \sigma_{\pi^*}, \sigma_{y^*}].$$

Finally, we assume that the innovations $\varepsilon_{r,t}$, $\varepsilon_{z,t}$, $\varepsilon_{q,t}$, $\varepsilon_{\pi^*,t}$, and $\varepsilon_{y^*,t}$ are independent standard normal random variables. We stack these innovations in the vector ε_t.

2. The DSGE-VAR Approach

To capture potential misspecification of the stylized small open economy model described in the previous section, we embed it into a vector autoregressive specification that allows us to relax cross-coefficient restrictions. We refer to the resulting empirical model as DSGE-VAR. We have developed this DSGE-VAR framework in a series of papers, including Del Negro and Schorfheide (2004, 2008b) and Del Negro and others (2007). The remainder of this section reviews the setup in Del Negro and Schorfheide (2008b), which is used in the subsequent empirical analysis.

2.1 A VAR with Hierarchical Prior

Equation (2), which describes the policymaker's behavior, can be written in more general form as

$$y_{1,t} = \mathbf{x}_t'\boldsymbol{\beta}_1(\boldsymbol{\theta}) + \mathbf{y}_{2,t}'\boldsymbol{\beta}_2(\boldsymbol{\theta}) + \varepsilon_{1,t}\,\sigma_r\,, \tag{12}$$

where $\mathbf{y}_t = [y_{1,t}, \mathbf{y}'_{2,t}]'$ and the $k \times 1$ vector $\mathbf{x}_t = [\mathbf{y}'_{t-1},\ldots, \mathbf{y}'_{t-p}, 1]'$ is composed of the first p lags of $\mathbf{y}_t$ and an intercept. Here $y_{1,t}$ corresponds to the nominal interest rate, $\tilde{R}_t$, while the subvector $\mathbf{y}_{2,t}$ is composed of output growth, inflation, exchange rate depreciation, and terms-of-trade changes:

$$\mathbf{y}_{2,t} = \left[(\Delta\hat{y}_t + \hat{z}_t), \hat{\pi}_t, \hat{e}_t, \hat{q}_t\right].$$

The vector-valued functions $\boldsymbol{\beta}_1(\boldsymbol{\theta})$ and $\boldsymbol{\beta}_2(\boldsymbol{\theta})$ interact with $\mathbf{x}_t$ and $\mathbf{y}_{2,t}$ to reproduce the policy rule.

The solution of the linearized DSGE model presented in section 1 generates a moving-average representation of $\mathbf{y}_{2,t}$ in terms of ε_t. We proceed by approximating this moving-average representation with a pth-order autoregression, which we write as

$$\mathbf{y}_{2,t} = \mathbf{x}_t'\boldsymbol{\Psi}^*(\boldsymbol{\theta}) + \mathbf{u}_{2,t}'. \tag{13}$$

If we ignore the approximation error for a moment, the one-step-ahead forecast errors, $\mathbf{u}_{2,t}$, are functions of structural innovations ε_t.

Assuming that, under the DSGE model, the law of motion for $\mathbf{y}_{2,t}$ is covariance stationary for every $\boldsymbol{\theta}$, we define the moment matrices

$$\boldsymbol{\Gamma}_{XX}(\boldsymbol{\theta}) = E_{\theta}^{D}[\mathbf{x}_t \mathbf{x}_t']$$

and

$$\boldsymbol{\Gamma}_{XY_2}(\boldsymbol{\theta}) = E_{\theta}^{D}[\mathbf{x}_t \mathbf{y}_{2,t}'].$$

In our notation, $E_{\theta}^{D}[\bullet]$ denotes an expectation taken under the probability distribution for $\mathbf{y}_t$ and $\mathbf{x}_t$ generated by the DSGE model conditional on the parameter vector $\boldsymbol{\theta}$. We define the VAR approximation of $\mathbf{y}_{2,t}$ through

$$\boldsymbol{\Psi}^*(\boldsymbol{\theta}) = \boldsymbol{\Gamma}_{XX}^{-1}(\boldsymbol{\theta})\, \boldsymbol{\Gamma}_{XY_2}(\boldsymbol{\theta}). \tag{14}$$

The equation for the policy instrument (12) can be rewritten by replacing $\mathbf{y}_{2,t}$ with expression (13):

$$y_{1,t} = \mathbf{x}_t' \boldsymbol{\beta}_1(\boldsymbol{\theta}) + \mathbf{x}_t' \boldsymbol{\Psi}^*(\boldsymbol{\theta})\, \boldsymbol{\beta}_2(\boldsymbol{\theta}) + u_{1,t}. \tag{15}$$

Let $\mathbf{u}_t' = [u_{1,t}, \mathbf{u}_{2,t}']$ and define

$$\boldsymbol{\Sigma}^*(\boldsymbol{\theta}) = \boldsymbol{\Gamma}_{YY}(\boldsymbol{\theta}) - \boldsymbol{\Gamma}_{YX}(\boldsymbol{\theta})\, \boldsymbol{\Gamma}_{XX}^{-1}(\boldsymbol{\theta})\, \boldsymbol{\Gamma}_{XY}(\boldsymbol{\theta}). \tag{16}$$

If we assume that the $\mathbf{u}_t$ variables are normally distributed, denoted by $\mathbf{u}_t \sim \mathrm{N}(0, \boldsymbol{\Sigma}_*(\boldsymbol{\theta}))$, then equations (13) to (16) define a restricted VAR(p) for the vector $\mathbf{y}_t$. While the moving-average representation of $\mathbf{y}_t$ under the linearized DSGE model does not, in general, have an exact VAR representation, the restriction functions $\boldsymbol{\Psi}^*(\boldsymbol{\theta})$ and $\boldsymbol{\Sigma}^*(\boldsymbol{\theta})$ are defined such that the covariance matrix of $\mathbf{y}_t$ is preserved. Let $E_{\Psi,\Sigma}^{\mathrm{VAR}}[\cdot]$ denote expectations under the restricted VAR. It can be verified that

$$E_{\Psi^*(\theta),\Sigma^*(\theta)}^{VAR}\left[\mathbf{y}_t \mathbf{y}_t'\right] = E_{\theta}^{D}\left[\mathbf{y}_t \mathbf{y}_t'\right].$$

To account for potential misspecification we now relax the DSGE model restrictions and allow for VAR coefficient matrices $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$

that deviate from the restriction functions $\boldsymbol{\Psi}^*(\boldsymbol{\theta})$ and $\boldsymbol{\Sigma}^*(\boldsymbol{\theta})$. Thus,

$$\begin{aligned} y_{1,t} &= \mathbf{x}_t' \boldsymbol{\beta}_1(\boldsymbol{\theta}) + \mathbf{x}_t' \boldsymbol{\Psi}\boldsymbol{\beta}_2(\boldsymbol{\theta}) + u_{1,t}, \\ \mathbf{y}_{2,t} &= \mathbf{x}_t' \boldsymbol{\Psi} + \mathbf{u}_{2,t}', \end{aligned} \tag{17}$$

and $\mathbf{u}_t \sim N(0, \boldsymbol{\Sigma})$. Our analysis is cast in a Bayesian framework in which initial beliefs about the DSGE-model parameter $\boldsymbol{\theta}$ and the VAR parameters $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$ are summarized in a prior distribution. Our prior distribution for $\boldsymbol{\Psi}$ and $\boldsymbol{\Sigma}$ is chosen such that, conditional on a DSGE-model parameter $\boldsymbol{\theta}$,

$$\begin{aligned} \boldsymbol{\Sigma} \mid \boldsymbol{\theta} &\sim IW(T^* \boldsymbol{\Sigma}^*(\boldsymbol{\theta}), T^* - k), \\ \boldsymbol{\Psi} \mid \boldsymbol{\Sigma}, \boldsymbol{\theta} &\sim N\left(\boldsymbol{\Psi}^*(\boldsymbol{\theta}), \frac{1}{T^*}\left[\left(\mathbf{B}_2(\boldsymbol{\theta})\boldsymbol{\Sigma}^{-1}\mathbf{B}_2(\boldsymbol{\theta})'\right) \otimes \boldsymbol{\Gamma}_{XX}(\boldsymbol{\theta})\right]^{-1}\right), \end{aligned} \tag{18}$$

where IW denotes the inverted Wishart distribution, N is a multivariate normal distribution, $\mathbf{B}_1(\boldsymbol{\theta}) = [\boldsymbol{\beta}_1(\boldsymbol{\theta}), \mathbf{0}_{k\times(n-1)}]$, and $\mathbf{B}_2(\boldsymbol{\theta}) = [\boldsymbol{\beta}_2(\boldsymbol{\theta}), \mathbf{I}_{(n-1)\times(n-1)}]$.

Our hierarchical prior is computationally convenient. We use Markov-chain-Monte-Carlo (MCMC) methods (described in Del Negro and Schorfheide, 2008b) to generate draws from the joint posterior distribution of $\boldsymbol{\Psi}$, $\boldsymbol{\Sigma}$, and $\boldsymbol{\theta}$. We refer to the empirical model comprising the likelihood function associated with the restricted VAR in equation and the prior distributions $p_\lambda(\boldsymbol{\Psi}, \boldsymbol{\Sigma}|\boldsymbol{\theta})$, given in equation (18), and $p(\boldsymbol{\theta})$ as DSGE-VAR(λ).

2.2 Selecting the Tightness of the Prior

The distribution of prior mass around the restriction functions $\boldsymbol{\Psi}^*(\boldsymbol{\theta})$ and $\boldsymbol{\Sigma}^*(\boldsymbol{\theta})$ is controlled by the hyperparameter T^*, which we reparameterize in terms of multiples of the actual sample size T, that is, $T^* = \lambda T$. Large values of λ imply that large discrepancies are unlikely to occur and the prior concentrates near the restriction functions. We consider values of λ on a finite grid, Λ, and use a data-driven procedure to determine an appropriate value for this hyperparameter. A natural criterion to select λ in a Bayesian framework is the marginal data density:

$$p_\lambda(Y) = \int p(Y \mid \boldsymbol{\Psi}, \boldsymbol{\Sigma}, \boldsymbol{\theta}) p_\lambda(\boldsymbol{\Psi}, \boldsymbol{\Sigma}, \boldsymbol{\theta}) d(\boldsymbol{\Psi}, \boldsymbol{\Sigma}, \boldsymbol{\theta}). \tag{19}$$

Here $p_\lambda(\mathbf{\Psi},\mathbf{\Sigma},\boldsymbol{\theta})$ is a joint prior distribution for the VAR coefficient matrices and the DSGE model parameters. This prior is obtained by combining the prior in equation (18) with a prior density for $\boldsymbol{\theta}$, denoted by $p(\boldsymbol{\theta})$:

$$p_\lambda(\mathbf{\Psi},\mathbf{\Sigma},\boldsymbol{\theta}) = p(\boldsymbol{\theta})\, p_\lambda(\mathbf{\Sigma} \mid \boldsymbol{\theta})\, p_\lambda(\mathbf{\Phi} \mid \mathbf{\Sigma},\boldsymbol{\theta}). \tag{20}$$

Suppose that Λ consists of only two values, λ_1 and λ_2. Moreover, suppose that the econometrician places equal prior probability on these two values. The posterior odds of λ_1 versus λ_2 are given by the marginal likelihood ratio, $p_{\lambda_1}(Y)/p_{\lambda_2}(Y)$. More generally, if the grid consists of J values that have equal prior probability, then the posterior probability of λ_j is proportional to $p_{\lambda_j}(Y)$, $j = 1, \ldots, J$. Rather than averaging our conclusions with respect to the posterior distribution of the lambdas, we condition on the value λ_j that has the highest posterior probability. Such an approach is often called empirical Bayes analysis in the literature. In particular, we define

$$\hat{\lambda} = \operatorname{argmax}_{\lambda \in \Lambda}\; p_\lambda(Y). \tag{21}$$

As discussed in Del Negro and others (2007), the marginal likelihood ratio, $p_{\lambda=\hat{\lambda}}(Y)/p_{\lambda=\infty}(Y)$, provides an overall measure of fit for the DSGE model. If the data are not at odds with the restrictions implied by the DSGE model—that is, there exists a parameterization, $\tilde{\boldsymbol{\theta}}$, of the DSGE model for which the model-implied autocovariances are similar to the sample autocovariances—then $\hat{\lambda}$ will be large and $p_{\lambda=\hat{\lambda}}(Y)/p_{\lambda=\infty}(Y)$ will be small. If the data turn out to be at odds with the DSGE model implications, $\hat{\lambda}$ will be fairly small and $p_{\lambda=\hat{\lambda}}(Y)/p_{\lambda=\infty}(Y)$ will be large. We come back to the interpretation of these marginal likelihood ratios when we discuss the empirical results.

2.3 Identification of Structural Shocks

Up to this point, we have expressed the VAR in terms of one-step-ahead forecast errors, $\mathbf{u}_t$. It is more useful, however, to express the VAR as a function of the structural shocks, ε_t, both for understanding the dynamics of the DSGE-VAR and for the purpose of policy analysis.

In our setup, the monetary policy shock is identified through exclusion restrictions:

$$y_{1,t} = \mathbf{x}'_t\boldsymbol{\beta}_1(\boldsymbol{\theta}) + (\mathbf{x}'_t\boldsymbol{\Psi} + \mathbf{u}'_{2,t})\,\boldsymbol{\beta}_2(\boldsymbol{\theta}) + \varepsilon_{1,t}\,\sigma_R;$$

$$\mathbf{y}'_{2,t} = \mathbf{x}'_t\boldsymbol{\Psi} + \mathbf{u}'_{2,t}.$$

According to the underlying DSGE model, $\mathbf{u}_{2,t}$ is a function of the monetary policy shock, $\varepsilon_{1,t}$, and other structural shocks, $\varepsilon_{2,t}$. We assume that the shocks $\varepsilon_{2,t}$ have unit variance and are uncorrelated with each other or with the monetary policy shock. We express $\mathbf{u}_{2,t}$ as

$$\mathbf{u}'_{2,t} = \varepsilon_{1,t}\mathbf{A}_1 + \varepsilon'_{2,t}\,\mathbf{A}_2. \tag{22}$$

Straightforward matrix algebra leads to the following formulas for the effect of the structural shocks on $\mathbf{u}'_{2,t}$:

$$\mathbf{A}_1 = \left[\boldsymbol{\Sigma}_{11} - \boldsymbol{\beta}'_2\,\boldsymbol{\Sigma}_{22}\,\boldsymbol{\beta}_2 - 2(\boldsymbol{\Sigma}_{12} - \boldsymbol{\beta}'_2\,\boldsymbol{\Sigma}_{22})\,\boldsymbol{\beta}_2\right]^{-1}(\boldsymbol{\Sigma}_{12} - \boldsymbol{\beta}'_2\,\boldsymbol{\Sigma}_{22}); \tag{23}$$

$$\mathbf{A}'_2\,\mathbf{A}_2 = \boldsymbol{\Sigma}_{22} - \mathbf{A}'_1\left[\boldsymbol{\Sigma}_{11} - \boldsymbol{\beta}'_2\,\boldsymbol{\Sigma}_{22}\,\boldsymbol{\beta}_2 - 2(\boldsymbol{\Sigma}_{12} - \boldsymbol{\beta}'_2\,\boldsymbol{\Sigma}_{22})\,\boldsymbol{\beta}_2\right]\mathbf{A}_1. \tag{24}$$

While the above decomposition of the forecast error covariance matrix identifies $\mathbf{A}_1$, it does not uniquely determine the matrix $\mathbf{A}_2$. To do so, we follow the approach taken in Del Negro and Schorfheide (2004). Let $\mathbf{A}'_{2,tr}\mathbf{A}_{2,tr} = \mathbf{A}'_2\mathbf{A}_2$ be the Cholesky decomposition of $\mathbf{A}'_2\mathbf{A}_2$. The relationship between $\mathbf{A}_{2,tr}$ and $\mathbf{A}_2$ is given by $\mathbf{A}'_2 = \mathbf{A}'_{2,tr}\boldsymbol{\Omega}$, where $\boldsymbol{\Omega}$ is an orthonormal matrix that is not identifiable based on the estimates of $\boldsymbol{\beta}(\boldsymbol{\theta})$, $\boldsymbol{\Psi}$, and $\boldsymbol{\Sigma}$. However, we are able to calculate an initial effect of $\varepsilon_{2,t}$ on $\mathbf{y}_{2,t}$ based on the DSGE model, denoted by $\mathbf{A}_2^D(\boldsymbol{\theta})$. This matrix can be uniquely decomposed into a lower triangular matrix and an orthonormal matrix:

$$\mathbf{A}_2^{D'}(\boldsymbol{\theta}) = \mathbf{A}_{2,tr}^{D'}(\boldsymbol{\theta})\boldsymbol{\Omega}^*(\boldsymbol{\theta}). \tag{25}$$

To identify $\mathbf{A}_2$ above, we combine $\mathbf{A}'_{2,tr}$ with $\boldsymbol{\Omega}^*(\boldsymbol{\theta})$. Loosely speaking, the rotation matrix is constructed such that in the absence of misspecification, the DSGE model's and the DSGE-VAR's impulse responses to $\varepsilon_{2,t}$ coincide. To the extent that misspecification is mainly in the dynamics as opposed to the covariance matrix of innovations, the identification procedure can be interpreted as matching, at least qualitatively, the short-run responses of the VAR with those from the DSGE model. Since the matrix $\mathbf{\Omega}$ does not affect the likelihood function, we can express the joint distribution of data and parameters as follows:

$$p_{\lambda}(Y,\mathbf{\Psi},\mathbf{\Sigma},\mathbf{\Omega},\theta) = p\ (Y|\mathbf{\Psi},\mathbf{\Sigma})p_{\lambda}(\mathbf{\Psi},\mathbf{\Sigma}|\theta)\ p\ (\mathbf{\Omega}|\theta)\ p\ (\theta).$$

Here $p(\mathbf{\Omega}|\theta)$ is a point mass centered at $\mathbf{\Omega}^{*}(\theta)$. The presence of $\mathbf{\Omega}$ does not affect the MCMC algorithm. We can first draw the triplet $\mathbf{\Psi}$, $\mathbf{\Sigma}$, θ from the posterior distribution associated with the reduced-form DSGE-VAR, and then calculate $\mathbf{\Omega}$ according to $\mathbf{\Omega}^{*}(\theta)$. Details are provided in Del Negro and Schorfheide (2008b).

3. Data

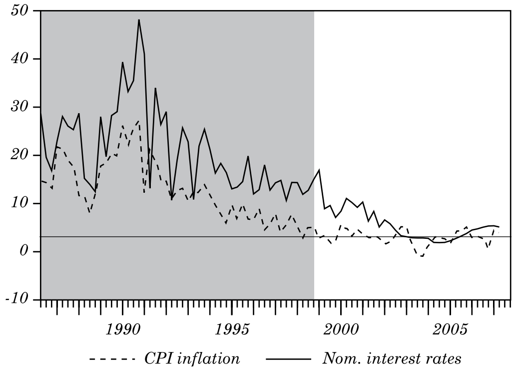
For our empirical analysis, we compiled a data set of observations on output growth, inflation, interest rates, exchange rates, and the terms of trade. Unless otherwise noted, the raw data are taken from the online database maintained by the Central Bank of Chile and seasonally adjusted. Output growth is defined as the log difference of real gross domestic product (GDP), scaled by 400 to convert it into annualized percentages. To construct the inflation series, we pass the consumer price index extracted from the Central Bank database through the X12 filter (using the default settings in EViews) to obtain a seasonally adjusted series; we then compute log differences, scaled by 400. The monetary policy rate (MPR) serves as our measure of nominal interest rates.[3] Annualized depreciation rates are computed from log differences of the Chilean peso / U.S. dollar exchange rate series. Finally, annualized quarter-to-quarter percentage changes in the terms of trade are computed from the export and import price indexes maintained by the Central Bank.

While we compile a data set that contains quarterly observations from 1986 to 2007, we restrict the estimation sample to the the period from 1999:1 to 2006:4 and hence to the most recent monetary policy regime. Between 1991 and 1999, the Central Bank applied a partial inflation-targeting approach that involved two nominal anchors: an exchange rate band and an inflation target. In 1999

3. To construct the MPR before 2001, we follow the approach in Chumacero (2005).

the Central Bank implemented a floating exchange rate and the institutional arrangements for full inflation targeting.[4] Official bank publications state that the operating objective of monetary policy is to keep annual inflation projections around 3.0 percent annually over a horizon of about two years. Indeed, the average inflation rate in our estimation sample is 2.8 percent. We plot the path of the inflation rate and the nominal interest rate in figure 1 for the period 1986 to 2007. Chile experienced a decade-long disinflation process throughout the 1990s, and with the adoption of the 3 percent target inflation rate in 1999, inflation and nominal interest rates stabilized at a low level.

Figure 1. Interest Rates and Inflation in Chile



Sources: Central Bank of Chile and authors' calculations.

The average growth rate of real output (4.4 percent during our sample period) provides an estimate of γ in equation (1). The average inflation rate can be viewed as an estimate of the target inflation rate π_*, and the average nominal interest rate can be linked to the discount factor β, because our model implies that $R_* = \gamma/\beta + \pi_*$. It turns out that the sum of average inflation and output growth is 7.2 percent and exceeds the average nominal interest rate, which is about 5.6 percent. The sample averages are thus inconsistent with the model's steady-state implications. Rather than estimating the steady-state

4. Since 2000, the Central Bank of Chile has provided an inflation report with public inflation and growth forecasts. The inflation target has been stable at 3 percent since 2001.

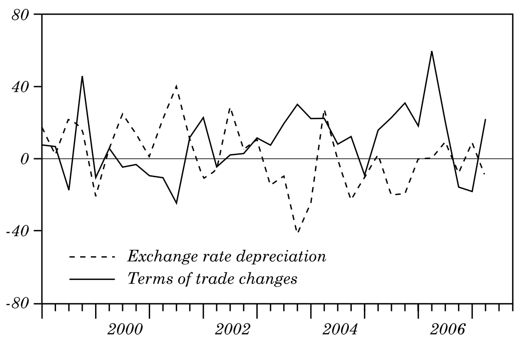
parameters jointly with the remaining DSGE model parameters and imposing the steady-state restrictions, we decided to demean our observations and fit the DSGE model and the DSGE-VAR to demeaned data.

To provide further details on the features of our data set, we plot the peso-dollar exchange rate in figure 2, together with percentage changes in the terms of trade. Both series exhibit very little autocorrelation and are very volatile. According to our DSGE model, the exchange rate fluctuations are a function of inflation differentials and terms-of-trade movements:

$$\hat{e}_t = \hat{\pi}_t - \hat{\pi}_t^* - (1-\alpha)\hat{q}_t$$

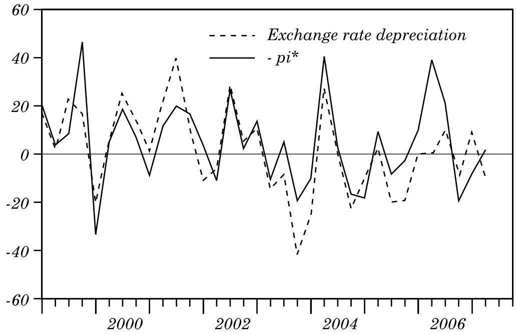
The rest of the world's inflation rate, $\hat{\pi}_t^*$, is treated as a latent variable. In figure 3 we plot the exchange rate depreciation and the implicit inflation in the rest of the world, $-\hat{\pi}_t^* = \hat{e}_t - \hat{\pi}_t + (1-\alpha)\hat{q}_t$ for $\alpha = 0.3$. The figure illustrates the well-known exchange rate disconnect: most of the fluctuations in the nominal exchange rate are generated by the exogenous process $\hat{\pi}_t^*$.

Figure 2. Exchange Rate and Terms of Trade Dynamics



Sources: Central Bank of Chile and authors' calculations.

Figure 3. Exchange Rate Movements and PPP



Sources: Central Bank of Chile and authors' calculations.

4. Empirical Results

The empirical analysis has four parts. In section 4.1 we estimate a monetary policy rule for Chile and examine the extent to which the Central Bank responds to exchange rate and terms-of-trade movements. We proceed in section 4.2 by studying the degree of misspecification of the DSGE model. More specifically, we compare the marginal likelihood of the DSGE model with that of the DSGE-VAR for various choices of the hyperparameter λ. Section 4.3 examines whether the Central Bank managed to insulate the Chilean economy, in particular inflation, from external shocks. Finally, section 4.4 explores the effect of changing the response to inflation in the feedback rule on the variance of inflation.

4.1 Estimating the Policy Rule

This section investigates the feedback rule followed by the Central Bank in the recent period. As discussed before, Chile witnessed significant movements in the nominal exchange rate after it entered the freely floating regime in 1999. Moreover, it was subject to large swings in the terms of trade. Did the Central Bank respond to these movements in order to pursue the inflation target? Table 1 addresses this question. The table estimates the coefficients of the policy rule (equation 2) under three different specifications. Under the first specification, which we refer to as the baseline, policy only responds to inflation and real output growth, in addition to the lagged interest rate. Under the second and third specifications (response to the exchange rate and response to the terms of trade, respectively), policy also responds to the exchange rate depreciation. Finally, the terms of trade also enter the feedback rule in the third specification, as do real output growth, inflation, and the nominal exchange rate. The posterior means of the policy rule coefficients estimated using the DSGE model under the three specifications are reported in columns (1a), (2a), and (3a) of the table, with the associated posterior standard deviations in parenthesis. For each specification we also compute the marginal likelihood, which measures model fit in a Bayesian framework, as well as the posterior odds relative to the baseline specification. This latter figure is computed under the assumption that we assign equal prior weights to all specifications.

Both posterior estimates of the parameters and model comparison results are, in a finite sample, sensitive to the choice of prior.[5] Since the sample considered here is fairly short, we want to examine the robustness of our conclusions to the choice of prior. Table 1 thus presents our results for two priors, which differ in terms of the marginal distribution for two key parameters of interest: the policy responses to fluctuations in the exchange rate (ψ_3) and in the terms of trade (ψ_4). For the first prior, the marginal distribution for ψ_3 is centered at 0.25 with a standard deviation of 0.12. This prior embodies the belief that the response to the exchange rate depreciation is on average substantial, but also quite diffuse. That is, it allows for the possibility that the response can either be small or very large. Likewise, the first prior is agnostic as to the response to the terms-of-trade depreciation. The prior is symmetric around zero, as we do not have a priori views on the sign of the response, and the standard deviation is quite large (0.50). This prior therefore allows for large positive or negative responses. The second is far less agnostic. Here the marginal distribution for ψ_3 is centered at 0.10 with a standard deviation of 0.05. This prior embodies a relatively sharp belief that policy responds to depreciation, albeit not too strongly. The center of the marginal distribution on ψ_4 is still zero, but the standard deviation is 0.10, five times smaller than in the case of the first prior.

5. See Del Negro and Schorfheide (2008a) for a discussion of prior elicitation and robustness in the context of DSGE models.

Table 1. Which Policy Rule?[a]

		DSGE			*DSGE-VAR* (λ = *1.5*)		
		Baseline	*Response to FX*	*Response to ToT*	*Baseline*	*Response to FX*	*Response to ToT*
Parameter	*Prior*	*(1a)*	*(2a)*	*(3a)*	*(1b)*	*(2b)*	*(3b)*
				Prior 1			
ψ_1	2.50	2.23	2.36	2.28	2.86	2.82	2.87
	(0.50)	(0.47)	(0.52)	(0.53)	(0.47)	(0.49)	(0.51)
ψ_2	0.25	0.33	0.29	0.32	0.16	0.16	0.16
	(0.13)	(0.14)	(0.14)	(0.14)	(0.08)	(0.08)	(0.08)
ψ_3	0.25		0.07	0.07		0.08	0.09
	(0.12)		(0.03)	(0.03)		(0.03)	(0.04)
ψ_4	0.00			−0.02			0.05
	(0.50)			(0.05)			(0.08)
ρ_r	0.50	0.38	0.41	0.40	0.47	0.48	0.49
	(0.20)	(0.11)	(0.11)	(0.12)	(0.10)	(0.11)	(0.11)
Marginal likelihood		−571.02	−575.17	−577.42	−558.38	−562.12	−563.78
Posterior odds		(1.000)	(0.016)	(0.002)	(1.000)	(0.024)	(0.005)

Table 1. (continued)

		DSGE			*DSGE-VAR* (λ = *1.5*)		
Parameter	*Prior*	*Baseline* (1a)	*Response to FX* (2a)	*Response to ToT* (3a)	*Baseline* (1b)	*Response to FX* (2b)	*Response to ToT* (3b)
				Prior 2			
ψ_1	2.50	2.23	2.36	2.28	2.86	2.82	2.84
	(0.50)	(0.47)	(0.53)	(0.51)	(0.47)	(0.48)	(0.47)
ψ_2	0.25	0.33	0.29	0.29	0.16	0.16	0.16
	(0.13)	(0.14)	(0.14)	(0.13)	(0.08)	(0.07)	(0.08)
ψ_3	0.10		0.05	0.05		0.06	0.06
	(0.05)		(0.02)	(0.02)		(0.03)	(0.03)
ψ_4	0.00			–0.02			0.02
	(0.10)			(0.04)			(0.06)
ρ_r	0.50	0.38	0.40	0.41	0.47	0.47	0.47
	(0.20)	(0.11)	(0.11)	(0.11)	(0.10)	(0.11)	(0.11)
Marginal likelihood		–571.02	–572.99	–573.66	–558.38	–560.13	–560.76
Posterior odds		(1.000)	(0.139)	(0.071)	(1.000)	(0.174)	(0.093)

Source: Authors' calculations.
a. We report means and standard deviations (in parentheses).

The distributions for the remaining policy parameters, ψ_1, ψ_2, and ρ_r, are the same across the two priors. Since Chile entered the full-fledged inflation-targeting regime in 1999 and it had acquired a reputation as an inflation fighter in the previous decade, we posit a fairly large prior mean on ψ_1, the response to inflation. The prior is centered at 2.50 with a standard deviation of 0.50. The priors on ψ_2, the response to real output growth, and ρ_r, the persistence parameter, are similar to those used in Lubik and Schorfheide (2006). These priors are also similar to that used in the estimation of DSGE models for the United States. The prior distribution on the remaining DSGE model parameters are again the same across the two priors, and they are discussed in detail in the next section.

We now discuss the posterior estimates of the policy parameters for the three specifications, which are shown in columns (1a), (2a), and (3a) of table 1. The estimates of ψ_1 are consistent across specifications, ranging from 2.23 to 2.36, with a standard deviation of about 0.5. Our prior was that the Central Bank responds strongly to inflation, and there is little updating from the prior to the posterior. [6] The estimates of ψ_2 range from 0.29 to 0.33 and imply only modest updating relative to the prior. The estimates for ψ_1 and ψ_2 are also roughly the same for both priors.

The main focus of the section lies in the responses to nominal depreciation and to the terms of trade. In these dimensions the data are quite informative. Column (2a) reveals that under the first prior, the posterior mean for ψ_3, the response to nominal depreciation, is 0.07, much lower than the prior mean. Moreover, the posterior is much more concentrated than the prior. The data strongly indicate that the response to exchange rate depreciations, if at all nonzero, is much smaller than the response to CPI inflation. To put this estimate into perspective, define the target nominal interest rate as

$$\hat{R}_t^* = 2.36\hat{\pi}_t + 0.29(\Delta\hat{y}_t + \hat{z}_t) + 0.07\hat{e}_t.$$

6. Since the official inflation target is stated in terms of year-over-year inflation, we also consider a fourth specification in which we replace quarter-to-quarter inflation. We find that this specification is strongly rejected by the data using our posterior odds criterion. This result should not be interpreted as contradicting the statement that the Central Bank target is year-over-year inflation, but simply as providing information on the rule the Central Bank follows to achieve this target. Caputo and Liendo (2005) consider a rule in which the policymaker responds to expected inflation and find it does not improve fit.

Here we replaced the policy rule coefficients by their posterior mean estimates. The sample standard deviations of inflation, output growth, and nominal exchange rate depreciations are 1.57, 3.77, and 18.10, respectively. We can therefore rewrite the target interest rate as

$$\hat{R}_t^* = 3.17\frac{\hat{\pi}_t}{\hat{\sigma}(\hat{\pi})} + 1.09\frac{\Delta\hat{y}_t + \hat{z}_t}{\hat{\sigma}(\Delta\hat{y} + \hat{z})} + 1.27\frac{\hat{e}_t}{\hat{\sigma}(\hat{e})}.$$

After the standardization, the coefficient on the exchange rate depreciation is 1.27, whereas the coefficient on CPI inflation is 3.17.

Column (3a) shows that when we further add the response to the terms of trade to the feedback rule, the estimated coefficient ψ_4 under the first prior is negative but small. Our standardized posterior mean estimate for the terms-of-trade coefficient is –0.37. The posterior standard deviation is also relatively small, indicating that the data rule out a large response. The marginal likelihood and posterior odds show that under the first prior, the alternative specifications are rejected by the data. The posterior odds relative to the baseline are 1.6 and 0.2 percent, respectively.

It is conceivable that the response to the exchange rate or terms of trade, while not as important as that of inflation, is still significant. We embody this belief in the second prior. The posterior mean of ψ_3 under the second prior is 0.05, which is smaller than under the first prior. However, ψ_3 is now more precisely estimated. The posterior mean of ψ_4 is the same under both priors, while the posterior standard deviation decreases by 0.01 under the second prior. Even under the tighter prior, the posterior odds favor the baseline specification. To summarize, based on the DSGE model estimation, we conclude that responding to inflation is much more important for the Central Bank than responding to the exchange rate or the terms of trade.

Full-information estimation has pros and cons if one is interested in the parameters of a particular equation in the system, in this case the policy rule. On the one hand, if the cross-equation restrictions imposed by the model are correct, full-information estimation is more efficient than single-equation instrumental variable estimation. On the other hand, to the extent that these cross-equation restrictions are invalid, the full-information estimates are potentially biased, and limited-information methods may be preferable. In this context, DSGE-VAR strikes a compromise between full- and limited-information estimation, as it allows for deviations from the cross-equation restrictions. In the case at hand, such a compromise may be necessary, since the sample size is small and estimators that completely ignore the restrictions ($\lambda = 0$) tend to produce poor estimates in a mean-squared-error sense. At the same time, our DSGE model generates strong cross-equation restrictions (such as exogeneity of the terms of trade), so we may not want to impose them dogmatically ($\lambda = \infty$). For these reasons, columns (1b), (2b), and (3b) of table 1 show the estimates of ψ_1, ψ_2, ψ_3, and ρ_r according to the three specifications of interest using a DSGE-VAR with two lags and $\lambda = 1.5$. We justify the choice of lag length and hyperparameter in section 4.2. For now, notice that the marginal likelihood of the DSGE-VAR for all specifications is substantially higher than that of the corresponding DSGE model, validating some of the concerns about the cross-equation restrictions.

The DSGE-VAR estimates imply a stronger response to inflation and a weaker response to output growth than the DSGE model estimates, with a posterior mean of ψ_1 between 2.8 and 2.9 and ψ_2 at 0.16. The DSGE-VAR estimation confirms our previous findings regarding the response to exchange rate and terms-of-trade movements. Under the first prior the posterior means of ψ_3 are 0.08 and 0.09 for the second and third specifications, respectively. The posterior mean for ψ_4, the response to terms-of-trade changes, has the opposite sign than under the DSGE estimation, but it is still relatively small. Most importantly, the posterior odds suggest that the richer specifications are rejected relative to the baseline. Under the second prior, the estimates for ψ_3 are also in line with those obtained under the DSGE estimation. The estimates for ψ_4 again have the opposite sign, but are close to zero.

As emphasized by Galí and Monacelli (2005), optimal monetary policy in our DSGE model would consist of stabilizing domestic inflation, $\hat{\pi}_{H,t}$, and the gap between actual and flexible price output, $\hat{y}_t - \bar{y}_t$. Since according to our model $\hat{\pi}_{H,t} = \hat{\pi}_t + \alpha\hat{q}_t$, and the estimated import share α is between 25 and 30 percent, our posterior estimates in columns (3a) and (3b) of table 1 suggest that the Central Bank does not try to stabilize $\hat{\pi}_{H,t}$.

In summary, we have robust empirical evidence that the Central Bank responded only very mildly to movements in the nominal exchange rate or the terms of trade in the recent period, if it responded at all. Rather, CPI inflation is the driving force behind changes in interest rates. Our post-1999 findings are consistent with the official policy statements of the Central Bank of Chile.

4.2 The Fit of the Small Open Economy DSGE Model

This section discusses the fit of the small open economy DSGE model and the estimates of the nonpolicy parameters. More specifically, we examine how the fit of the DSGE-VAR changes as we relax the cross-equation restrictions implied by the DSGE model. From a policy perspective, this analysis is useful for assessing whether forecasting should be conducted with a tightly parameterized empirical specification that closely resembles the DSGE model, or with a densely parameterized VAR that uses little a priori information.

Table 2 shows the log marginal likelihood for the DSGE model and the DSGE-VAR, where λ varies in a grid from 0.75 to 5.0. As discussed in section 2, high values of λ correspond to tightly imposed cross-equation restrictions, while low values imply a relatively flat prior on the VAR parameters. The table also shows the posterior odds relative to the best-fitting model, which are computed under the assumption that all specifications have equal prior probabilities.

In previous studies that employ the DSGE-VAR methodology (Del Negro and Schorfheide, 2004; Del Negro and others, 2007), we use a VAR specification with four lags, which we denote VAR(4). Four lags are fairly standard in applications with twenty to forty years of quarterly data. Since we have fewer than nine years of data in the present application, an unrestricted estimation of a VAR(4) would imply that we are using only thirty-four observations to determine twenty parameters per equation. Consequently, a DSGE-VAR with four lags would require high values of λ, not because the DSGE model is a particularly good description of the data, but because only a very tight prior is able to reduce the variability of the estimates. We proceed by reducing the number of lags in the VAR. Columns (1) and (2) and columns (3) and (4) of the table show the log marginal likelihood and posterior odds results for DSGE-VARs with two and three lags, respectively.

Four features emerge from table 2. First, for any value of λ, the log marginal likelihood for two lags (column 1) is always greater than that for three lags (column 3), indicating that reducing the number of lags, and hence the number of free parameters, increases the fit of the empirical model. If we raise the number of lags to four, the log marginal likelihood decreases even further. Second, the gap in marginal likelihoods between columns (1) and (3) tends to decrease with λ: increasing the weight of the DSGE model's restrictions implicitly decreases the number of free parameters and hence makes the difference between VARs with two and three lags less stark.

Table 2. The Fit of the Small Open Economy DSGE Model[a]

Specification	*Lambda* (λ)	*Log marginal likelihood* *(1)*	*Posterior odds* *(2)*	*Log marginal likelihood* *(3)*	*Posterior odds* *(4)*
DSGE	—	–571.02	(3e–06)	–571.02	(8e–05)
		Two lags		*Three lags*	
DSGE-VAR	5.0	–562.89	(0.011)	–563.79	(0.110)
	3.0	–560.69	(0.099)	–562.83	(0.286)
	2.5	–559.83	(0.235)	–561.58 ($\hat{\lambda}$)	(1.000)
	2.0	–559.11	(0.482)	–561.86	(0.756)
	1.5	–558.38 ($\hat{\lambda}$)	(1.000)	–561.76	(0.835)
	1.0	–559.19	(0. 445)	–564.48	(0.055)
	0.75	–561.46	(0. 046)	–570.93	(9e–05)

Source: Authors' calculations.

a. The difference of log marginal data intensities can be interpreted as log posterior odds under the assumption that the two specifications have equal prior probabilities. We report odds relative to the DSGE-VAR($\hat{\lambda}$).

Third, the best fit is achieved for a value of λ that is lower for the VAR(2) than the VAR(3) specification. Using the notation of section 2, $\hat{\lambda}$ takes the values 1.5 and 2.5, respectively. The DSGE model restrictions help in part because they reduce the number of free parameters, and this reduction becomes more valuable the larger the lag length.[7] Finally, the fit of the DSGE model is considerably worse than that of the DSGE-VAR($\hat{\lambda}$), regardless of the number of lags. Columns (2) and (4) of table 2 show that the posterior odds of the DSGE model relative to DSGE-VAR($\hat{\lambda}$) are 1 percent and 10 percent, respectively, indicating that from a statistical point of view there is evidence that the cross-equation restrictions are violated in the data. We investigate in section 4.3 whether this statistical evidence is economically important, that is, whether it translates into sizeable differences with respect to the dynamic response of the endogenous variables to different shocks.

Table 3 provides the estimates of the DSGE model's nonpolicy parameters. We focus on the estimates obtained with the the two-lag DSGE-VAR($\hat{\lambda}$). Results for the VAR(3) are quantitatively similar. The first column of the table shows the prior mean and standard deviations. The parameter α measures the fraction of foreign-produced goods in the domestic consumption basket. In 2006 imported goods as a share of total domestic demand in Chile was about 30 percent. Restricted to consumer goods, this share was 10 percent. We decided to center our prior at the 30 percent value, allowing for substantial variation. The parameter r^* can be interpreted as the growth-adjusted real interest rate. While our observations on average GDP growth, inflation, and nominal interest rates between 1999 and 2007 suggest that this value is negative, we view this as a temporary phenomenon and center our prior for r^* at 2.5 percent. The parameter κ corresponds to the slope of the Phillips curve, which captures the degree of price stickiness. According to our prior, κ falls with high probability in the interval 0 to 1, which encompasses large nominal rigidities as well as the case of near flexible prices. The parameter τ captures the inverse of the relative risk aversion. We center our prior at 2, which implies that consumers are slightly more risk averse than consumers with a log utility function. Finally, the priors for the parameters of the exogenous processes were chosen with presample evidence in mind.

7. Using the dummy observation interpretation of Del Negro and Schorfheide (2004), $\lambda = 1.5$ implies that the actual data are augmented by $1.5 \times T$ artificial observations from the DSGE model.

Table 3. DSGE Model Parameters[a]

Parameter	*Prior*		*DSGE*		*DSGE-VAR* 2 lags, $\hat{\lambda} = 1.5$	
α	0.30	(0.10)	0.28	(0.07)	0.26	(0.07)
r^*	2.50	(1.00)	2.49	(0.97)	2.47	(0.94)
κ	0.50	(0.25)	0.65	(0.17)	0.89	(0.26)
τ	0.50	(0.20)	0.31	(0.07)	0.40	(0.09)
ρ_z	0.20	(0.10)	0.61	(0.07)	0.53	(0.06)
ρ_q	0.50	(0.10)	0.41	(0.05)	0.40	(0.06)
ρ_{y^*}	0.85	(0.05)	0.88	(0.05)	0.87	(0.05)
ρ_{π^*}	0.70	(0.15)	0.35	(0.11)	0.38	(0.12)
σ_z	1.88	(0.99)	0.94	(0.20)	0.85	(0.13)
σ_q	4.39	(2.29)	4.60	(0.53)	3.51	(0.53)
σ_{y^*}	1.88	(0.99)	1.95	(0.68)	1.78	(0.71)
σ_π	1.88	(0.99)	4.77	(0.64)	3.36	(0.58)
σ_r	0.63	(0.33)	0.77	(0.14)	0.65	(0.14)

Source: Authors' calculations.
a. We report means and standard deviations (in parentheses).

The second column shows the posterior mean and standard deviations obtained from the estimation of the DSGE model. In light of the DSGE model misspecification discussed above, it is important to ask whether accounting for deviations from the cross-equation restrictions affects the inference about the DSGE parameters. Therefore, the third column shows the estimates obtained using DSGE-VAR($\hat{\lambda}$). The data provide little information on r^*, which enters the log-linear equations through the discount factor β, and the slope of the Phillips curve, κ. The estimated import share is about 25 percent, which again is not very different from the prior. The information from output, inflation, interest rate, exchange rate, and terms-of-trade data is not in contrast with that obtained from import quantities. Finally, the posterior mean of τ decreases compared to its prior, and its standard deviation shrinks from 0.2 to 0.1 or less. The posterior estimates for α, κ, and τ are similar to those obtained by Caputo and Liendo (2005) for the 1999–2005 sample. The estimated standard deviation of the monetary policy shock is around 60 to 80 basis points. Overall, the parameter estimates obtained from the state-space representation of the DSGE model and the DSGE-VAR are very similar.

Since the DSGE model itself exhibits very little endogenous propagation, the dynamics of the data are mostly captured by the estimated autocorrelation parameters of the exogenous shock processes. The terms of trade are purely exogenous in the DSGE model, and thus the posterior means of ρ_q and σ_q measure the autocorrelation and innovation standard deviation in our terms-of-trade series. The foreign inflation process, π_t^*, is plotted in figure 3, and the estimates of ρ_{π^*} and σ_{π^*} capture its persistence of volatility. The remaining sources of cyclical fluctuations are a foreign demand shock, $\hat{y}_t^*$, and a technology growth shock, $\hat{z}_t$. The estimated autocorrelations of these shocks are 0.88 and 0.61 (DSGE) and 0.87 and 0.53 (DSGE-VAR), respectively. In general, we observe that the shock-standard-deviation and autocorrelation estimates obtained with the DSGE-VAR are slightly smaller. The reason is that the DSGE-VAR can capture model misspecification by deviating from cross-equation restrictions, whereas the directly estimated DSGE model has to absorb this misspecification in the exogenous shock processes.

4.3 The Determinants of Inflation

This section discusses the impulse responses of the endogenous variables to internal and external shocks. Given that the Central Bank is in an inflation-targeting regime, we focus the discussion on the determinants of inflation dynamics. In section 4.4 we showed that the Central Bank seemingly does not respond to exchange rate or terms-of-trade movements. Did this policy manage to insulate the economy, and inflation in particular, from external shocks?

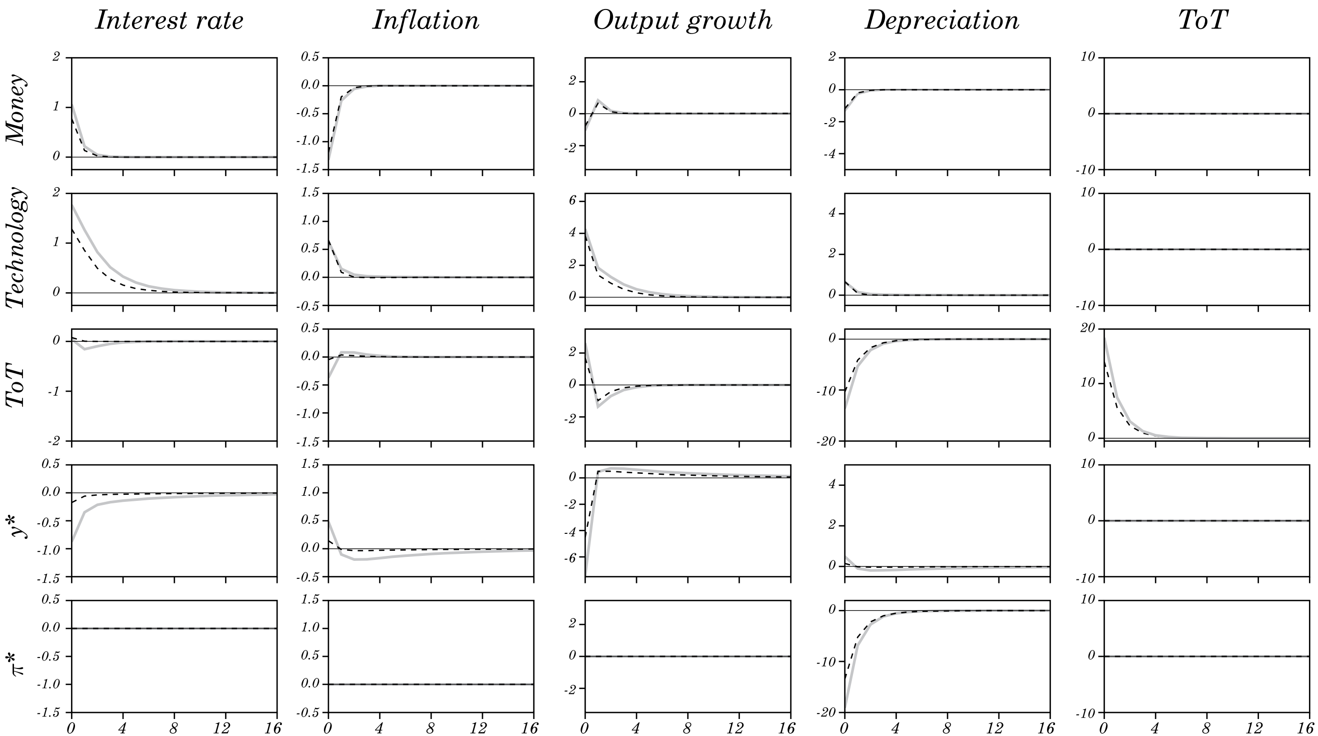
Figure 4 shows the impulse response functions to the five shocks described in section 1: monetary policy, technology, terms of trade, foreign output, and foreign inflation shocks. We overlay two impulse response functions, both of which are computed using the DSGE model. The difference between the two consists in the underlying estimates of the DSGE model parameter: one set of responses is based on the DSGE model estimates, whereas the other reflects the DSGE-VAR estimates. From a qualitative standpoint, the shape of the two response functions is the same. The main difference between them is that the DSGE-VAR responses are more pronounced, reflecting the larger estimated standard deviation of shocks documented in table 3.

Monetary policy shocks are contractionary shocks to the feedback rule (equation 2). As the interest rate increases, inflation and output decrease, and the exchange rate appreciates. Notably, the small estimated amount of nominal rigidities implies that the output response is very modest. Positive technology shocks raise output. As in Lubik and Schorfheide (2007), these shocks also raise marginal costs and thereby increase inflation and interest rates.[8]

Improvements in the terms of trade lead to an increase of output and a depreciation of the exchange rate, but they have only a moderate effect on inflation. To understand these responses, it is helpful to substitute the definition of CPI inflation (equation 7) into the policy rule (equation 2). We now have a three-equation system in $\hat{R}_t$, $\hat{\pi}_{H,t}$, and $\hat{y}_t$. In this system, shocks to the terms of trade, which are assumed to be exogenous, play essentially the same role as policy shocks. They thus have a similar impact on domestic inflation, $\hat{\pi}_{H,t}$ and $\hat{y}_t$, as the monetary policy shocks, but of the opposite sign. An appreciation of the terms of trade therefore leads to an increase in output and domestic inflation. The latter roughly compensates the impact of the appreciation, so that overall inflation, $\hat{\pi}_t$, does not move much in the end. Output does move, however, indicating that when a central bank responds to overall rather than domestic inflation, it fails to insulate the real side of the economy from external shocks (see Galí and Monacelli, 2005).

8. Equation (6) shows that marginal costs and detrended output, $\hat{y}_t$, move one to one, for given flexible price output. The latter is an exogenous function of foreign output, y_t^*; see expression (4).

Figure 4. DSGE Model Impulse Responses: DSGE vs DSGE-VAR($\hat{\lambda}$) Parameter Estimates[a]



Source: Authors' calculations.

a. The figure depicts DSGE model impulse responses based on the DSGE (gray solid line) and DSGE-VAR($\hat{\lambda}$) (black dashed line) respective posterior estimates summarized in tables 1 and 3.

Shocks to foreign output have a negative impact on domestic output, again as in Lubik and Schorfheide (2007). The other variables are not particularly affected. Recall from expression that flexible price output, $\bar{y}_t$, depends negatively on foreign output. According to the estimated parameters, the degree of stickiness in this economy is limited, so actual output pretty much behaves as flexible price output. Consequently, inflation is unaffected since marginal costs barely move. Finally, in the baseline specification the Central Bank does not respond to movements in the exchange rate; this isolates the economy from shocks to foreign inflation, which only lead to an appreciation of the currency.

In terms of the determinants of inflation, the interesting feature of figure 4 is that the shocks that move the terms of trade and the nominal exchange rate depreciation—namely, terms-of-trade and foreign inflation shocks—barely affect CPI inflation. According to the DSGE model identification, the shocks that have the largest impact on inflation are largely domestic, namely technology and monetary policy shocks. Notably, these shocks have little effect on the exchange rate depreciation (or on the terms of trade, which are exogenous by construction). In summary, the impulse responses indicate absence of strong comovements between inflation and the external variables. These findings suggest that the monetary authorities have been successful in isolating inflation from foreign disturbances.

It is somewhat surprising that monetary policy shocks have a significant effect on inflation, given that these shocks are avoidable. One possibility is that the Central Bank, in the attempt to respond to future rather than current inflation, makes errors in forecasting inflation. From the model's perspective, these errors appear as policy shocks. Another possible explanation is that the policy reaction function is misspecified, and policy responds to some other variable not included in the reaction function. While this is certainly a possibility, we know from section 4.4 that the missing variable cannot be the exchange rate or the terms of trade.

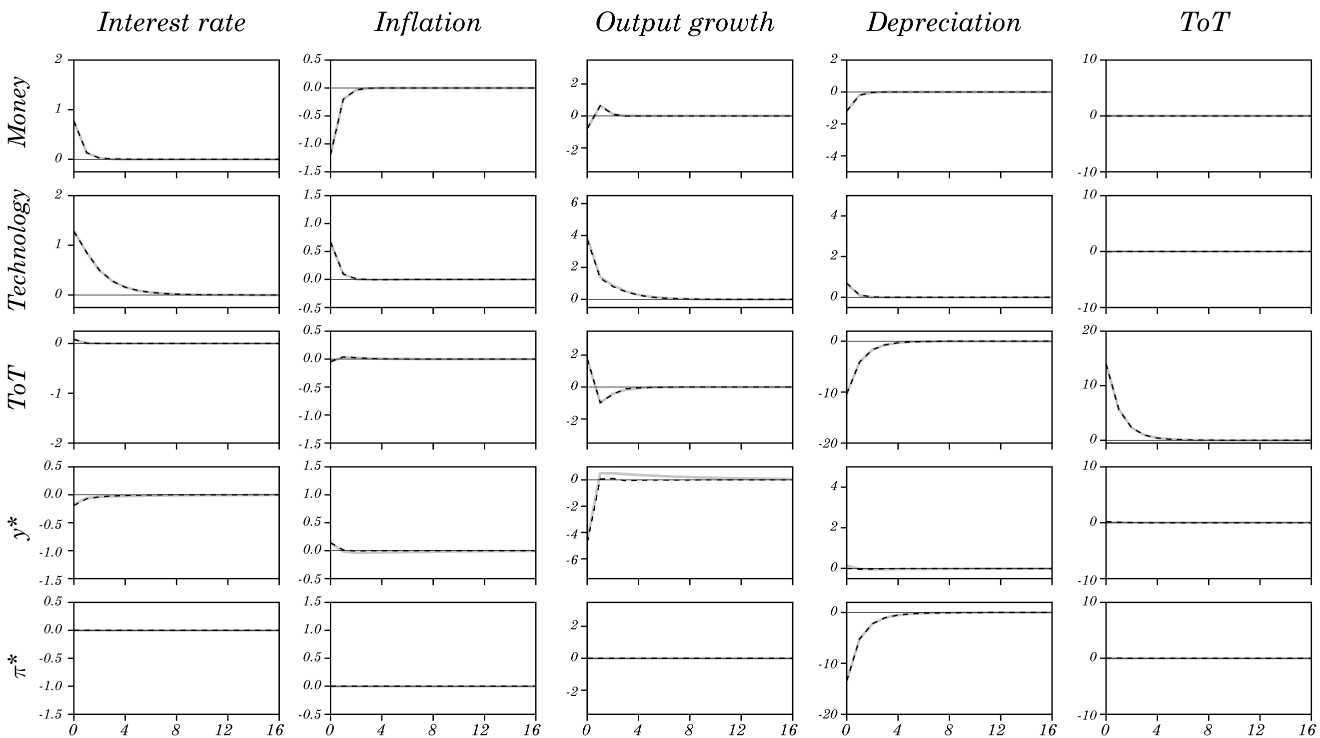
Figure 4 shows that the impulse responses are generally not very persistent, reflecting the fact that the DSGE model does not generate much internal propagation. Moreover, the DSGE impulse responses are computed under stark identification assumptions, such as exogeneity of the terms of trade. These limitations, as well as the evidence of misspecification discussed in the previous section, suggest that we may want to compare the DSGE model impulse responses to those from the DSGE-VAR and check whether relaxing the cross-equation restrictions alters the dynamics substantially. In comparing the DSGE model impulse responses with those from the DSGE-VAR, one should bear in mind that in principle some differences may arise from the fact that the DSGE model does not have an exact finite VAR representation (see Ravenna, 2007, among others). Figure 5 shows that in the case considered here, this is not a quantitatively important issue. The figure compares the DSGE impulse responses with those obtained from the finite order VAR approximation of the DSGE model, that is, DSGE-VAR($\lambda = \infty$). The two are virtually identical. This implies that if the data were generated by the DSGE model at hand, the DSGE-VAR would recover the "true" impulse response functions.

Figure 6 compares the impulse responses computed from the DSGE-VAR($\lambda = \infty$), which are identical to the DSGE-VAR responses in figure 4, to those from DSGE-VAR($\hat{\lambda}$).[9] The figure shows that by and large the differences between the DSGE-VAR($\lambda = \infty$) and the DSGE-VAR($\hat{\lambda}$) impulse responses lie in the dynamics of the nominal exchange rate, which is somewhat more volatile and persistent in the DSGE-VAR than in the DSGE model. In discussing the DSGE model's impulse responses, we remarked that shocks that move inflation do not affect the terms of trade or the exchange rate. This is less the case for the DSGE-VAR($\hat{\lambda}$), where technology shocks have a substantial impact on inflation and a prolonged effect on the exchange rate. However, compared with the response of exchange rates to terms-of-trade or foreign inflation shocks, the response to technology shocks is small. Hence, the conclusion that inflation has by and large been isolated from external shocks seems to be robust to the presence of misspecification.

The terms-of-trade impulse responses are not very different either. Note that the assumption of exogeneity of the terms of trade is not strictly imposed on the DSGE-VAR. If the data were substantially at odds with this assumption, we would see differences between the two sets of impulse responses in the last column. While we see some differences, these are small relative to the magnitude of movements in the terms of trade. Thus, the short-cut of treating the terms of trade as exogenous in the DSGE model is supported by our empirical analysis. In discussing figure 4, we noted that according to the DSGE model, the economy is isolated from foreign inflation shocks since the Central Bank does not respond to the exchange

9. We do not show the posterior bands for simplicity of exposition; they are available on request.

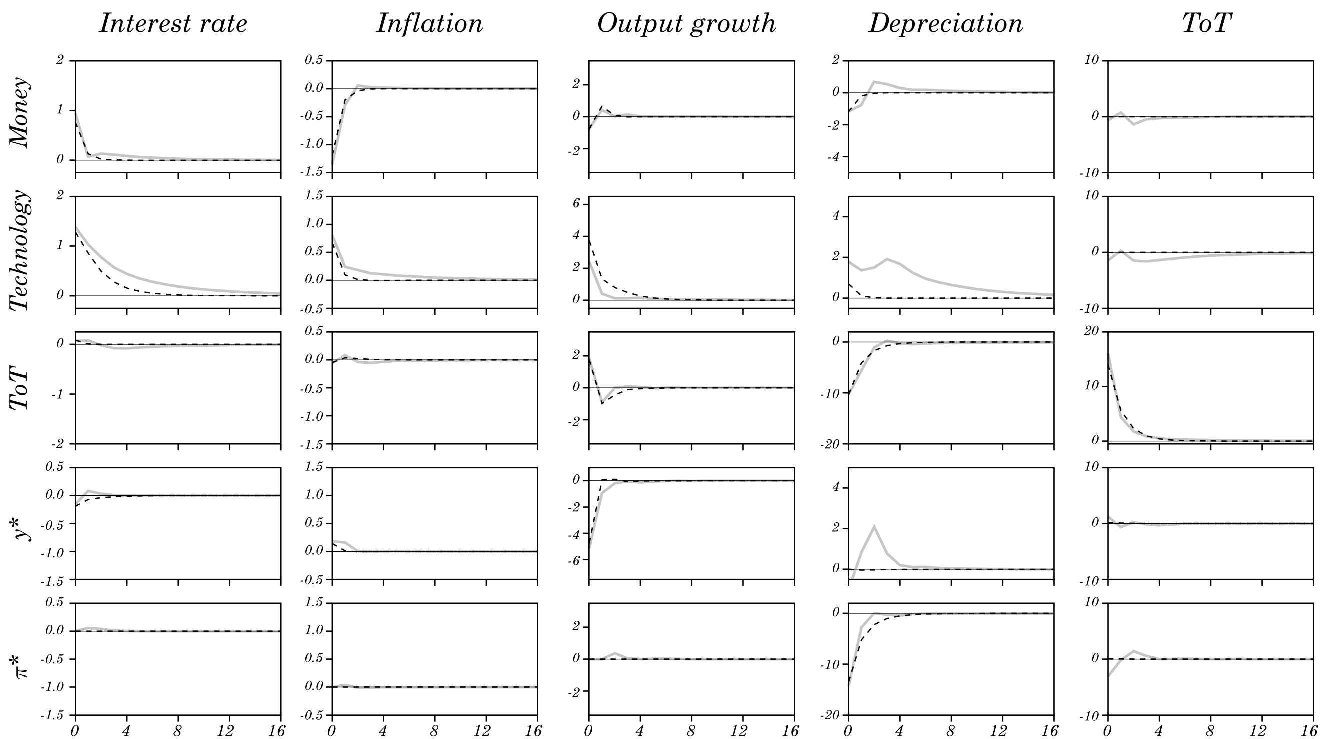
Figure 5. VAR Approximation Error: DSGE vs DSGE-VAR($\lambda = \infty$) Impulse Responses[a]



Source: Authors' calculations.

a. The figure depicts impulse responses from the DSGE model (black dashed line) and the DSGE-VAR(∞) (gray solid line) based on the DSGE-VAR($\hat{\lambda}$) posterior estimates summarized in tables 1 and 3. The DSGE-VAR is estimated with two lags.

Figure 6. Impulse Responses: DSGE-VAR($\lambda = \infty$) versus DSGE-VAR($\hat{\lambda}$)[a]



Source: Authors' calculations.

a. The figure depicts impulse responses from the DSGE model (black dashed line) and the DSGE-VAR(∞) (gray solid line) based on the DSGE-VAR($\hat{\lambda}$) = 1.5 posterior estimates summarized in tables 1 and 3.

rate. This is still the case according to the DSGE-VAR in figure 6, even though the cross-equation restrictions that deliver this result are not dogmatically imposed.

In summary, figure 6 suggests that the misspecification found in section 4.2 is not very important from an economic standpoint. This result must be interpreted with caution, however. The identification in the DSGE-VAR is, by construction, linked to that in the DSGE model. While this may be a virtue, as it ties the DSGE-VAR impulse responses to those of the underlying DSGE model, it can also be a drawback. There may be other DSGE models, and other identification schemes, that are equally capable of describing the data. By construction, DSGE-VAR is not going to be able to uncover such models. Finally, because of the short sample, the data may simply not be informative enough to point out the deficiencies of this model.

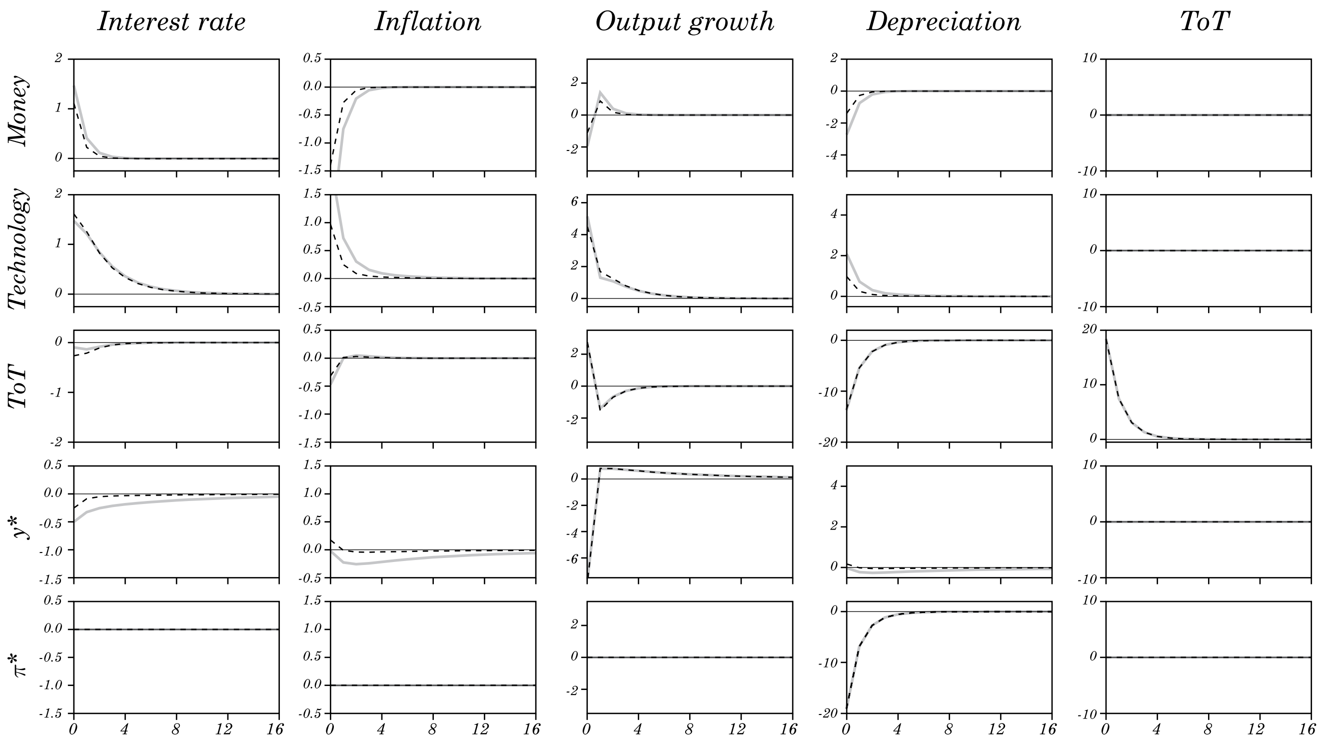
4.4 A Look at Alternative Policy Rules

In this section, we examine the effect on macroeconomic volatility of responding more or less aggressively to inflation. Conducting this policy analysis with the DSGE model is straightforward. We simply re-solve the model under the new policy rule. Using the DSGE-VAR to assess the effect of changes in the monetary policy rule is conceptually more difficult. We apply the approaches recently proposed in Del Negro and Schorfheide (2008b) to use the DSGE-VAR to check the robustness of the DSGE model analysis in view of the misspecification of the structural model that we documented in the previous subsections.

Figure 7 describes how the DSGE model impulse responses change as the parameter ψ_1 in the policy reaction function varies from 1.25 to 2.75 and to 3.50. Although each plot has three lines, it appears to have only two because raising the reaction to inflation from its estimated value of 2.75 to 3.50 has virtually no impact on the dynamics. Hence, responding more aggressively to inflation would not have any effect on the Chilean economy, at least according to this estimated model. Conversely, a much weaker response to inflation ($\psi_1 = 1.25$) would have serious effects, especially on inflation. The response to both technology and monetary policy shocks would be much more pronounced.

Figure 8 shows DSGE-model-based variance differentials with respect to the historical policy rule, $\psi_1 = 2.75$, as we vary ψ_1 on a grid ranging from 1.00 to 3.50. The figure graphs the posterior mean (90 percent posterior bands) differentials under the DSGE

Figure 7. DSGE Model Impulse Responses as Function of ψ_1[a]

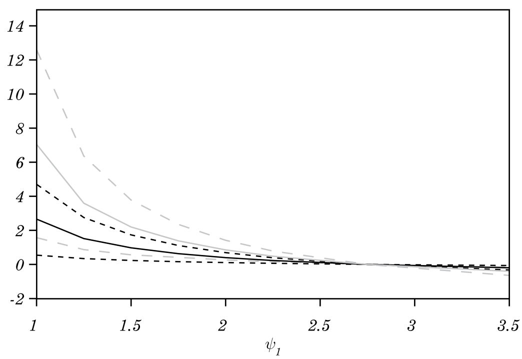


Source: Authors' calculations.

a. The figure plots the posterior mean of the DSGE model impulse responses computed for three different values of the response to inflation in the policy rule, ψ_1: 3.50 (black), 2.75 (dark gray), and 1.25 (light gray). The remaining policy parameters, ψ_2, ψ_3, ψ_4, and ρ_r, are kept at the baseline values of 0.125, 0, 0, and 0.5, respectively. For all impulse responses, we use the DSGE model posterior estimates of the nonpolicy parameters, $\theta_{(np)}$, summarized in table 3.

model estimates of parameters (second column of table 3), as well as the posterior mean (90 percent posterior bands) differentials under the DSGE-VAR estimates of parameters (third column of table 3). Consistent with figure 7, under both sets of estimates the variance of inflation increases substantially as ψ_1 decreases below 1.50, while not much happens as ψ_1 increases from 2.75 to 3.50. The magnitude of the increase in the variance differential differs substantially under the two sets of estimates. The shocks are estimated to be more persistent and more variable under the DSGE than under DSGE-VAR, so the effect of changes in policy on the variability of inflation is larger. One can view the higher persistence and variability of the exogenous shocks under the DSGE model estimates as a consequence of the model's misspecification, as discussed in section 4.2, and therefore not trust the outcomes of the policy analysis exercise under these estimates. In any case, these results highlight the sensitivity of the policy exercises to the estimates of the processes followed by the exogenous shocks, a point made in Del Negro and Schorfheide (2008b).

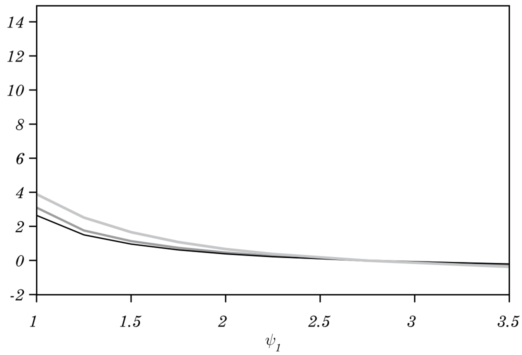
Figure 8. Comparative Performance of Policy Rules: Benchmark DSGE versus DSGE-VAR(λ = 2) Parameter Estimates[a]



Source: Authors' calculations.

a. Posterior expected variance differentials as a function of ψ_1 relative to the baseline policy rule, ψ_1 = 2.75. The remaining policy parameters, ψ_2, ψ_3, ψ_4, and ρ_r, are kept at the baseline values of 0.125, 0, 0, and 0.5, respectively. Negative differentials signify a variance reduction relative to the baseline rule. Differentials are computed using DSGE-VAR posterior (gray) and DSGE model (black) posterior estimates of the nonpolicy parameters, $\theta_{(np)}$, summarized in table 3. The solid (dashed) gray lines represent the posterior mean (90 percent posterior bands) differentials under the DSGE model estimates of parameters (second column of table 3). The solid (dotted) black lines represent the posterior mean (90% posterior bands) differentials under the DSGE-VAR estimates of parameters (third column of table 3).

Figure 9. Comparative Performance of Policy Rules: DSGE versus DSGE-VAR Policy-Invariant Misspecification and DSGE-VAR Backward-Looking Analysis[a]



Source: Authors' calculations.

a. Posterior expected variance differentials as a function of ψ_1 relative to the baseline policy rule, ψ_1 = 2.75. The remaining policy parameters, ψ_2, ψ_3, ψ_4, and ρ_r, are kept at the baseline values of 0.125, 0, 0, and 0.5, respectively. Negative differentials signify a variance reduction relative to baseline rule. Differentials are computed using the DSGE-VAR backward-looking analysis (light gray), the DSGE-VAR policy-invariant misspecification scenario (dark gray), and the DSGE model (black), where the latter uses the DSGE-VAR(λ = 2) posterior estimates of the nonpolicy parameters, $\theta_{(np)}$, summarized in table 3.

Figure 9 shows the expected changes in the variability of inflation under three different approaches to performing the policy experiment. Under all three approaches, the experiment is the one just described—that is, varying ψ_1 in a grid ranging from 1.00 to 3.50. The first approach is the same as in the previous paragraph: it amounts to performing the experiment using the DSGE model under the DSGE-VAR estimates of the nonpolicy parameters. The second approach is called the DSGE-VAR policy-invariant misspecification and is described in detail in Del Negro and Schorfheide (2008b). This approach to policy assumes that while the cross-equation restrictions change with policy, the deviations from the cross-equation restrictions outlined in figure 6 are policy invariant. More specifically in terms of the DSGE-VAR notation, the matrices that embody the cross-equation restriction—namely, $\boldsymbol{\Psi}^*(\boldsymbol{\theta})$ and $\boldsymbol{\Sigma}^*(\boldsymbol{\theta})$—change with ψ_1, but the deviations—namely, $\boldsymbol{\Psi}^\Delta = \boldsymbol{\Psi} - \boldsymbol{\Psi}^*(\boldsymbol{\theta})$ and $\boldsymbol{\Sigma}^\Delta = \boldsymbol{\Sigma} - \boldsymbol{\Sigma}^*(\boldsymbol{\theta})$—do not.[10]

10. We work with the moving average rather than the VAR representations, as discussed in Del Negro and Schorfheide (2008b). Thus we literally treat the deviations from the DSGE-VAR(∞) impulse responses in figure 5 as policy invariant.

This approach may be appealing if one thinks that these deviations capture low or high frequency movements in the data that are not going to be affected by policy. The variance differential under this alternative approach is about the same as under the DSGE model (as are the bands, which we do not show to avoid cluttering the figure). This is not surprising given that the deviations from the cross-equation restrictions are small, particularly for inflation.

The second approach is called the DSGE-VAR backward-looking analysis and is again described in detail in Del Negro and Schorfheide (2008b). Under this approach, the DSGE-VAR is treated as an identified VAR: the change in ψ_1 only affects the policy rule (Sims, 1999), and it does not affect the remaining equation of the system. Under this approach, the cross-equation restriction are completely ignored. Although the rationale for ignoring the cross-equation restrictions when the deviations are small is questionable, the line in question is not very different from the other two. In the end, we find that in this application the treatment of misspecification leads to rather small differences relative to those shown in figure 8. As in Del Negro and Schorfheide (2008b), inference about the nonpolicy parameters, and in particular about the persistence and standard deviation of the shocks, is key in evaluating the outcomes of different policy rules.

5. Conclusion

We estimate the small open economy DSGE model used in Lubik and Schorfheide (2007) on Chilean data for the inflation-targeting period, 1999–2007, using data on the policy rate, inflation, real output growth, nominal exchange depreciation, and log differences in the terms of trade. We also estimate a Bayesian VAR with a prior generated from the small open economy DSGE model, following the DSGE-VAR methodology proposed in Del Negro and Schorfheide (2004, 2008b). The purpose of the DSGE-VAR is to check whether the answers provided by the DSGE model are robust to the presence of misspecification, where misspecification is defined as deviations from the cross-equation restrictions imposed by the model.

Our empirical results can be summarized as follows. First, our estimates of a monetary reaction function indicate that the Central Bank of Chile did not respond significantly to movements in the exchange rate and terms of trade. Second, our DSGE-VAR analysis suggests that, in part because of a short estimation sample, it is helpful to tilt the VAR estimates toward the restrictions generated by our small open economy DSGE model. A VAR that is estimated without a tight prior is unlikely to produce good forecasts or sharp policy advice. Third, both our estimated DSGE model and the DSGE-VAR indicate that the observed inflation variability is mostly due to domestic shocks. Moreover, despite the statistical evidence of DSGE model misspecification, the DSGE-VAR's implied dynamic responses to structural shocks closely mimic the DSGE model impulse response functions.

Finally, we find that a stronger Central Bank response to inflation movements would produce little change in inflation volatility, whereas a substantial decrease would lead to a spike in volatility. We obtain a quantitatively similar result if we conduct the policy analysis with the DSGE model. An important caveat to the policy analysis exercise is that the DSGE model used here has many restrictive assumptions, so it may not capture some important policy trade-offs. Nevertheless, we believe that a few lessons can be learned from this exercise, which are likely to carry over to more sophisticated models. First, the outcome of policy experiments is very sensitive to the estimates for the parameters describing the law of motion of the exogenous shocks. Second, the presence of misspecification—that is, the fact that the DSGE model is rejected relative to a more loosely parameterized model—does not necessarily imply that the answers to policy exercises obtained from the DSGE model are not robust. The DSGE-VAR methodology provides ways of checking the robustness of the policy advice under different assumptions about misspecification, and we hope this can be useful in applied work at central banks.

APPENDIX

Model Derivation

Time t decisions are made after observing all current shocks. Variables with the subscript $t-s$, where $s \geq 0$, are known at time t. We assume that asset markets are complete. For each state of nature, there is a security that pays one Chilean peso or one U.S. dollar.

Households

Domestic households solve the following decision problem

$$\max_{\{C_{h_t}, N_{h_t}, D_{h_t}, D^*_{h_t}\}_{t=1}^{\infty}} \sum_{t=0}^{\infty} \sum_{h_t} \beta^t \mathrm{P}(h_t \mid h_0) \left[\frac{(C_t/Z_t)^{1-\sigma} - 1}{1-\sigma} - \frac{N_t^{1+\varphi}}{1+\varphi} \right], \tag{A1}$$

subject to ,

$$P_{h_t} C_{h_t} + \sum_{h_{t+1}} \left(\mathrm{Q}_{h_{t+1}|h_t} D_{h_{t+1}} + \varepsilon_{h_t} \mathrm{Q}^*_{h_{t+1}|h_t} D^*_{h_{t+1}} \right) \leq W_{h_t} N_{h_t}$$

$$+ D_{h_t} + \varepsilon_{h_t} D^*_{h_t} + \int \Omega_{h_t}(i) di$$

where $h_{t+\tau}$ denotes the history of events up to time $t+\tau$, $\mathrm{P}(h_{t+\tau}|h_t)$ is its probability conditional on time t information, C_{h_t} is consumption of a composite good in state h_t, N_{h_t} is hours worked, P_{h_t} is the nominal price level of the composite good, $D_{h_{t+\tau}}$ ($D^*_{h_{t+\tau}}$) is holdings of a security that pays one unit of the domestic currency (foreign currency) in state $h_{t+\tau}$, $\mathrm{Q}_{h_{t+\tau}|h_t}$ ($\mathrm{Q}^*_{h_{t+\tau}|h_t}$) is its current price in pesos (in state h_t), ε_{h_t} is the nominal exchange rate (domestic currency/foreign currency), T_{h_t} are nominal transfers, and $\Omega_{h_t}(i)$ are (nominal) dividends earned from domestic firm i. Note that $\Sigma_{h_{t+1}} \mathrm{Q}_{h_{t+1}|h_t} = 1/R_{h_t}$ is the inverse of the one-period gross nominal risk-free interest rate in pesos, and $\Sigma_{h_{t+1}} \mathrm{Q}^*_{h_{t+1}|h_t} = 1/R^*_{h_t}$ is the inverse of the one-period gross nominal risk-free interest rate in dollars. Finally, Z_t is a world technology process, which is assumed to follow a random walk with drift. From now on, we use X_t to denote a variable X_{h_t}, and E_t to denote $\Sigma_{h_{t+\tau}} P(h_{t+\tau}|h_t)$.

After detrending consumption and nominal wages according to $c_t = C_t / Z_t$ and $w_t = W_t / (P_t Z_t)$, the first-order conditions can be written as

$$N_t^{\varphi} = c_t^{-\sigma} w_t, \tag{A2}$$

$$c_t^{-\sigma} = \beta E_t \left[R_t c_{t+1}^{-\sigma} \left(z_{t+1} \pi_{t+1} \right)^{-1} \right], \tag{A3}$$

and

$$0 = E_t \left[(R_t - R_t^* e_{t+1}) \frac{c_{t+1}^{-\sigma}}{c_t^{-\sigma}} \left(z_{t+1} \pi_{t+1} \right)^{-1} \right], \tag{A4}$$

where $z_t = Z_t / Z_{t-1}$, $\pi_t = P_t / P_{t-1}$ is the gross inflation rate, and $e_t = \varepsilon_t / \varepsilon_{t-1}$ is the gross depreciation rate.

Terms of Trade and the Real Exchange Rate

Let $P_{H,t}$ and $P_{F,t}$ be the domestic price of home- and foreign-produced goods, respectively. Define the terms of trade as follows:

$$Q_t = \frac{P_{H,t}}{P_{F,t}}. \tag{A5}$$

We assume that the law of one price for foreign goods holds:

$$P_{F,t} = \varepsilon_t P_{F,t}^*. \tag{A6}$$

Here $P_{F,t}^*$ is the price of the foreign-produced good in the foreign country, measured in foreign currency. We also assume that domestically produced goods have a negligible weight in foreign consumption. Specifically, let ϑ be the relative size of the domestic economy (defined more precisely below). We define $\alpha_* = \vartheta\alpha$ and we let $\vartheta \to 0$. Hence, $P_{F,t}^*$ will be approximately equal to the foreign CPI, P_t^*, and we can express the terms of trade as

$$Q_t = \frac{P_{H,t}}{(\varepsilon_t P_t^*)}. \tag{A7}$$

Both an exchange rate depreciation and foreign inflation reduce the terms of trade—that is, they make imports more expensive. Let P_t be the domestic CPI. The real exchange rate is defined as

$$S_t = \frac{\varepsilon_t P_t^*}{P_t}. \tag{A8}$$

Thus, the relative price $P_{H,t} / P_t$ can be expressed as

$$\frac{P_{H,t}}{P_t} = Q_t S_t \tag{A9}$$

Composite Goods

There are firms that buy quantities $C_{H,t}$ and $C_{F,t}$ of domestically-produced and foreign-produced goods and package them into a composite good that is used for consumption by the households. These firms maximize profits in a perfectly competitive environment:

$$\max_{C_t, C_{H,t}, C_{F,t}} P_t C_t - P_{H,t} C_{H,t} - P_{F,t} C_{F,t}, \tag{A10}$$

subject to $C_t = \left[(1-\alpha)^{1/\eta} C_{H,t}^{\frac{\eta-1}{\eta}} + \alpha^{1/\eta} C_{F,t}^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}}$.

We deduce from the first-order conditions and a zero-profit condition that

$$C_{H,t} = (1-\alpha)\left(\frac{P_{H,t}}{P_t}\right)^{-\eta} C_t; \quad C_{F,t} = \alpha\left(\frac{P_{F,t}}{P_t}\right)^{-\eta} C_t; \tag{A11}$$

$$P_t = \left[(1-\alpha) P_{H,t}^{1-\eta} + \alpha P_{F,t}^{1-\eta} \right]^{\frac{1}{1-\eta}}. \tag{A12}$$

Dividing equation (A12) by P_t and rearranging terms leads to the following relationship between the real exchange rate and the terms of trade:

$$S_t = \left[(1-\alpha) Q_t^{1-\eta} + \alpha \right]^{\frac{1}{\eta-1}}. \tag{A13}$$

The domestically-produced good, supplied in overall quantity Y_t, is itself a composite made up of a continuum of domestic intermediate goods, $Y_t(i)$:

$$Y_t = \left[\int_0^1 Y_t(i)^{\frac{\mu-1}{\mu}} di \right]^{\frac{\mu}{\mu-1}}. \tag{A14}$$

We assume that there are perfectly competitive firms that buy the domestic intermediate goods, package them, and resell the composite good to the firms that aggregate $C_{H,t}$ and $C_{F,t}$. These firms solve the following problem:

$$\max_{Y_t, Y_t(i)} P_{H,t} Y_t - \int_0^1 P_{H,t}(i) Y_t(i) di, \tag{A15}$$

subject to $Y_t = \left[\int_0^1 Y_t(i)^{\frac{\mu-1}{\mu}} di \right]^{\frac{\mu}{\mu-1}}$.

The first-order conditions and a zero-profit condition lead to

$$Y_t(i) = \left(\frac{P_{H,t}(i)}{P_{H,t}} \right)^{-\mu} Y_t, \quad P_{H,t} = \left[\int_0^1 P_{H,t}(i)^{1-\mu} di \right]^{\frac{1}{1-\mu}}. \tag{A16}$$

Domestic Intermediate Goods

The producers of the domestic intermediate goods, $Y_t(i)$, are monopolistic competitors. Firms can reoptimize prices in each period with probability $1 - \theta$. We assume that firms that are unable to reoptimize their price, $P_{H,t}(i)$, will increase according to the steady-state inflation rate, $\pi_{H,*}$. The firms use today's prices of state-contingent securities to discount future nominal profits. The firms' production function is linear in labor:

$$Y_t(i) = Z_t N_t(i), \tag{A17}$$

where productivity, Z_t, is not firm specific and its growth rate, $z_t = Z_t / Z_{t-1}$, follows an AR(1) process:

$$(\ln z_t - \gamma) = \rho_z(\ln z_{t-1} - \gamma) + \varepsilon_t^z, \tag{A18}$$

where γ is the steady-state growth rate of productivity. The firms' problem is given by

$$\max_{\tilde{P}_{H,t}(i), \{Y_{t+\tau}(i)\}_{\tau=0}^{\infty}} E_t \left[\sum_{\tau=0}^{\infty} \theta^{\tau} Q_{t+\tau|t} Y_{t+\tau}(i) (\tilde{P}_{H,t}(i) \pi_{H,*}^{\tau} - MC_{t+\tau}^n) \right], \tag{A19}$$

Subject to $Y_{t+\tau}(i) \leq \left(\frac{\tilde{P}_{H,t}(i)\pi_{H,*}^{\tau}}{P_{H,t+\tau}} \right)^{-\mu} Y_{t+\tau}$,

where $MC_{t+\tau}^{n} = W_{t+\tau} / Z_{t+\tau}$ is the nominal marginal cost and $Q_{t+\tau|t}$ is the time t price of a security that pays one unit of domestic currency in period $t + \tau$.

We are considering a symmetric equilibrium in which all firms solve the same problem and we eliminate the index i. The firms' first-order condition can then be written as:

$$E_t \left[\sum_{\tau=0}^{\infty} \theta^{\tau} Q_{t+\tau|t} \left(\frac{\tilde{P}_{H,t}\pi_{H,*}^{\tau}}{P_{H,t+\tau}} \right)^{-\mu} Y_{t+\tau} \left[(\mu - 1)\tilde{P}_{H,t}\pi_{H,*}^{\tau} - \mu MC_{t+\tau}^{n} \right] \right] = 0. \tag{A20}$$

The fraction of firms that are allowed to reoptimize their price is $1 - \theta$, while all others update their price by the steady-state inflation rate. Hence,

$$P_{H,t} = \left[\theta \tilde{P}_{H,t}^{1-\mu} + (1-\theta)(\pi_{H,*}P_{H,t-1})^{1-\mu} \right]^{\frac{1}{1-\mu}}. \tag{A21}$$

We now express both the nominal marginal costs and the price chosen by firms that are able to reoptimize in terms of the price of the domestic good:

$$mc_t = \frac{MC_t^n}{P_t^H} = \frac{W_t}{Z_t P_t} \frac{P_t}{P_t^H} = w_t Q_t^{-1} S_t^{-1} \tag{A22}$$

and

$$\tilde{p}_{H,t} = \frac{\tilde{P}_{H,t}}{P_{H,t}}.$$

Moreover, in equilibrium

$$Q_{t+\tau|t} = \beta^{\tau} \frac{\lambda_{t+\tau}}{\lambda_t} = \beta^{\tau} \frac{c_{t+\tau}^{-\sigma} P_t Z_t}{c_t^{-\sigma} P_{t+\tau} Z_{t+\tau}}. \tag{A23}$$

Thus, the optimal pricing rule can be restated as

$$\tilde{p}_{H,t} = \frac{\mu}{\mu-1} \frac{E_t\left[\sum_{\tau=0}^{\infty}(\beta\theta^{\tau})\frac{c_{t+\tau}^{-\sigma}}{c_t^{-\sigma}}\left(\frac{\pi_{H,*}^{\tau}}{\prod_{s=1}^{\tau}\pi_{H,t+s}}\right)^{-\mu} S_{t+\tau}Q_{t+\tau}y_{t+\tau}mc_{t+\tau}\prod_{s=1}^{\tau}\pi_{H,t+s}\right]}{E_t\left[\sum_{\tau=0}^{\infty}(\beta\theta^{\tau})\frac{c_{t+\tau}^{-\sigma}}{c_t^{-\sigma}}\left(\frac{\pi_{H,*}^{\tau}}{\prod_{s=1}^{\tau}\pi_{H,t+s}}\right)^{-\mu} S_{t+\tau}Q_{t+\tau}y_{t+\tau}\pi_{H,*}^{\tau}\right]}. \quad \text{(A24)}$$

Here we used the fact that $P_{H,t} / P_t = S_t Q_t$. We can re-write equation (A21) as

$$1 = \left[\theta\tilde{p}_{H,t}^{1-\mu} + (1-\theta)(\pi_{H,*}/\pi_{H,t-1})^{1-\mu}\right]^{\frac{1}{1-\mu}}. \quad \text{(A25)}$$

Domestic Market Clearing and Aggregate Production Function

The market for domestically produced goods clears if the following condition in terms of variables detrended by Z_t is satisfied

$$y_t = c_{H,t} + c^*_{H,t}. \quad \text{(A26)}$$

After substituting equation (A11) into equation (A26), we obtain

$$\begin{aligned} y_t &= c_{H,t} + c^*_{H,t} = (1-\alpha)\left(\frac{P_{H,t}}{P_t}\right)^{-\eta} c_t + \alpha\vartheta\left(\frac{P_{H,t}/\varepsilon_t}{P^*_t}\right)^{-\eta} c^*_t \\ &= (1-\alpha)(S_t Q_t)^{-\eta} c_t + \alpha\vartheta Q_t^{-\eta} c^*_t. \end{aligned} \quad \text{(A27)}$$

Finally, the aggregate production function for the domestic economy is given by

$$y_t = N_t\left[\int\left(\frac{P_{H,t}(i)}{P_{H,t}}\right)^{-\mu} di\right]^{-1}. \quad \text{(A28)}$$

The households receive the profits generated by the monopolistically competitive domestic intermediate goods producers. Firm i generates the following profit:

$$\Omega_t(i) = Y_t(i)P_{H,t}(i) - N_t(i)W_t.$$

Using the demand function (A16), we can write

$$\Omega_t(i) = P_{H,t}^{1-\mu} \frac{Y_t}{P_{H,t}^{\mu}} - N_t(i)W_t.$$

Integrating both sides and using the expression for the price of the composite good, we obtain

$$\int \Omega_t(i)di = P_{H,t}Y_t - W_t N_t. \tag{A29}$$

Finally, we deduce from the budget constraint that

$$P_{h_t}C_{h_t} - P_{H,h_t}Y_{h_t} = D_{h_t} + \varepsilon_{h_t}D^*_{h_t} - \sum_{h_{t+1}} \begin{pmatrix} Q_{h_{t+1}|h_t} D_{h_{t+1}} \\ +\varepsilon_{h_t} Q^*_{h_{t+1}|h_t} D^*_{h_{t+1}} \end{pmatrix}. \tag{A30}$$

The Rest of the World

We begin by exploiting the perfect-risk-sharing assumption to obtain a relationship between domestic and foreign consumption:

$$\left(\frac{c_{t+1}}{c_t}\right)^{\sigma} \pi_{t+1} = \left(\frac{c^*_{t+1}}{c^*_t}\right)^{\sigma} \pi^*_{t+1} e_{t+1}. \tag{A31}$$

Equation (A.31) can be rewritten as follows:

$$\left(\frac{c_{t+1}}{c^*_{t+1}}\right)^{\sigma} \frac{P_{t+1}}{\varepsilon_{t+1}P^*_{t+1}} = \left(\frac{c_t}{c^*_t}\right)^{\sigma} \frac{P_t}{\varepsilon_t P^*_t}.$$

This equation links consumption growth at home and abroad. To obtain implications about the level of consumption in the two countries, we assume that in period $t = 0$, $S_0 = 1$. Moreover, we let $\vartheta = C_0 / C^*_0$. We thus deduce that in period t

$$c_t = \vartheta c_t^* S_t^{1/\sigma}. \tag{A32}$$

We can now rewrite the market-clearing condition for the domestically produced good, recalling that $P_{H,t} / P_t = Q_t S_t$. We substitute equation (A32) into equation (A27) to obtain

$$y_t = \vartheta c_t^* Q_t^{-\eta} \left[(1-\alpha) S_t^{1/\sigma - \eta} + \alpha \right]. \tag{A33}$$

In slight abuse of notation, the foreign analog of equation (A30) is

$$\varepsilon_{h_t} P_{h_t}^* C_{h_t}^* - \varepsilon_{h_t} P_{F,h_t}^* Y_{h_t}^* = D_{F,h_t} + \varepsilon_{h_t} D_{F,h_t}^* - \sum_{h_{t+1}} \begin{pmatrix} \mathrm{Q}_{h_{t+1}|h_t} D_{F,h_{t+1}} \\ + \varepsilon_{h_t} \mathrm{Q}_{h_{t+1}|h_t}^* D_{F,h_{t+1}}^* \end{pmatrix}.$$

Since all state-contingent securities are in zero net supply, we obtain the following global resource constraint from the budget constraints of the domestic and foreign households:

$$c_t + S_t c_t^* = Q_t S_t y_t + S_t y_t^*. \tag{A34}$$

The equilibrium conditions are given by equations (A2), (A3), (A4), (A8), (A9), (A13), (A22), (A24), (A25), (A28), (A32), and (A33). Moreover, we let $\vartheta \to 0$ and use the approximation $c_t^* = y_t^*$. The system can be closed with interest rate feedback rules for the domestic and foreign central banks.

Steady States

The central banks at home and abroad are determining the steady-state inflation rates, π_* and π_*^*. Moreover, we assume that $S_0 = S_* = 1$. The consumption Euler equation implies that the domestic nominal interest rate is $R_* = z_* \pi_* / \beta$. A constant real interest rate implies that the nominal exchange rate depreciation in steady state is $e_* = \pi_* / \pi_*^*$. Uncovered interest rate parity determines the foreign nominal rate: $R_*^* = R_* / e_*$. The terms of trade are

$$Q_* = \left[\frac{1}{1-\alpha} \left(S_*^{\eta-1} - \alpha \right) \right]^{\frac{1}{1-\eta}} = 1.$$

Steady state inflation for the domestic good is $\pi_{H,*} = \pi_*$. According to the small open economy assumption, $c_*^* = y_*$. Clearing of the domestic goods market requires $y_* = \vartheta y_*^*$. Perfect risk sharing implies $c_* = \vartheta c_*^*$. The supply-side condition for the domestic good determines the steady-state labor input $N_* = y_*$. Finally, we can determine y_* from the marginal cost condition:

$$y_* = \left(\frac{\mu - 1}{\mu}\right)^{\frac{1}{\varphi+\sigma}}.$$

Now the global resource constraint (A34) is indeed satisfied for $Q_* = 1$.

Log-linearizations

We use the notation $\hat{X}_t$ to denote deviations of a variable X_t from its steady state, $X_* : \hat{X}_t = \ln X_t / X_*$. The relationship between the real exchange rate and the terms of trade is given by

$$\hat{S}_t = -(1-\alpha)\hat{Q}_t. \tag{A35}$$

Nominal exchange rates evolve according to

$$\hat{e}_t = \Delta\hat{S}_t + \hat{\pi}_t - \hat{\pi}_t^* = -(1-\alpha)\hat{q}_t + \hat{\pi}_t - \hat{\pi}_t^*, \tag{A36}$$

where $\hat{q}_t = \Delta\hat{Q}_t$. Inflation in the relative price of the domestic good is given by

$$\hat{\pi}_{H,t} = \hat{\pi}_t + \alpha\hat{q}_t. \tag{A37}$$

We can use the market-clearing condition for the domestically produced good to determine the level of the terms of trade:

$$\hat{Q}_t = -\frac{1}{\eta + (1-\alpha)^2(1/\sigma - \eta)}(\hat{y}_t - \hat{y}_t^*). \tag{A38}$$

From perfect risk sharing and the market-clearing condition for the foreign good, we have

$$\hat{c}_t = \hat{c}_t^* + \frac{1}{\sigma}\hat{S}_t = \hat{y}_t^* - \frac{(1-\alpha)}{\sigma}\hat{Q}_t,$$

where we substituted for S_t using equation (A35). We can now write the marginal costs as

$$\widehat{mc}_t = \varphi\hat{y}_t + \sigma\hat{y}_t^* + \frac{1}{\eta + (1-\alpha)^2 (1/\sigma - \eta)}(\hat{y}_t - \hat{y}_t^*). \tag{A39}$$

Define $\kappa = (1 - \theta\beta)(1 - \theta)/\theta$. Marginal costs determine the inflation of the domestically produced goods via the following Phillips curve:

$$\hat{\pi}_{H,t} = \beta E_t\left[\hat{\pi}_{H,t+1}\right] + \kappa\widehat{mc}_t. \tag{A40}$$

The consumption Euler equation is of the form

$$\hat{c}_t = E_t\left[\hat{c}_{t+1}\right] - \frac{1}{\sigma}(R_t - E_t\left[\hat{\pi}_{t+1} + \hat{z}_{t+1}\right]). \tag{A41}$$

Moreover, by combining the market-clearing condition for the domestic good with the perfect-risk-sharing condition, we deduce that

$$\hat{y}_t = \hat{c}_t - \alpha\left[\eta + (1-\alpha)\left(\eta - \frac{1}{\sigma}\right)\right]\hat{Q}_t. \tag{A42}$$

Equations (A35) to (A42) determine the evolution of $\hat{Q}_t, \hat{S}_t, \hat{e}_t$, $\hat{\pi}_t, \hat{\pi}_{H,t}$, $\widehat{mc}_t$, $\hat{c}_t$, and $\hat{y}_t$. We treat $\hat{\pi}_t^*$ and $\hat{y}_t^*$ as exogenous and close the system with an interest rate feedback rule that determines $\hat{R}_t$.

In section 1 of the main text, we consider a version of the open economy model in which $\varphi = 0$, $\eta = 1$, and $1/\sigma = \tau$. Notice that

$$\eta + (1-\alpha)^2\left(\frac{1}{\sigma} - \eta\right) = \tau + \alpha(2-\alpha)(1-\tau) = \tau + \lambda,$$

where $\lambda = \alpha(2 - \alpha)(1 - \tau)$.

We begin by determining the level of output $\bar{y}_t$ in the absence of nominal rigidities, which has to satisfy

$$0 = \widehat{mc}_t = \frac{1}{\tau}\hat{y}_t^* + \frac{1}{\tau+\lambda}(\bar{y}_t - \hat{y}_t^*).$$

We deduce that

$$\bar{y}_t = -\frac{\lambda}{\tau}\hat{y}_t^*.$$

The marginal costs satisfy

$$\widehat{mc}_t = \frac{1}{\tau+\lambda}(\hat{y}_t - \bar{y}_t), \tag{A43}$$

which leads to the Phillips curve (equation 5) reported in the main text:

$$\hat{\pi}_{H,t} = \beta E_t\left[\hat{\pi}_{H,t+1}\right] + \frac{\kappa}{\tau+\lambda}(\hat{y}_t - \bar{y}_t). \tag{A44}$$

We now manipulate the Euler equation. Using equation (A42), replacing CPI inflation by $\hat{\pi}_{H,t} - \alpha\Delta\hat{Q}_t$, and plugging in equation (A38), we can rewrite the consumption Euler equation as

$$\hat{y}_t - \bar{y}_t = E_t\left[\hat{y}_{t+1} - \bar{y}_{t+1}\right] - (\tau+\lambda)\left(\hat{R}_t - E_t\left[\hat{\pi}_{H,t+1} + \hat{z}_{t+1}\right]\right), \tag{A45}$$

which is equation in the main text.

References

Caputo, R. and F. Liendo. 2005. "Monetary Policy, Exchange Rate, and Inflation Inertia in Chile: A Structural Approach." Working paper 352. Santiago: Central Bank of Chile.

Caputo, R., F. Liendo and J.P. Medina. 2007. "New Keynesian Models for Chile in the Inflation-Targeting Period: A Structural Investigation." In *Monetary Policy under Inflation Targeting*, edited by F. Mishkin and K. Schmidt-Hebbel. Santiago: Central Bank of Chile.

Céspedes, L.F. and C. Soto. 2007. "Credibility and Inflation Targeting in Chile." In *Monetary Policy under Inflation Targeting*, edited by F. Mishkin and K. Schmidt-Hebbel. Santiago: Central Bank of Chile.

Christiano, L.J., M. Eichenbaum and C. Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113(1): 1–45.

Chumacero, R. 2005. "A Toolkit for Analyzing Alternative Policies in the Chilean Economy." In *General Equilibrium Models for the Chilean Economy*, edited by R. Chumacero and K. Schmidt-Hebbel. Santiago: Central Bank of Chile.

Del Negro, M. and F. Schorfheide. 2004. "Priors from General Equilibrium Models for VARs." *International Economic Review* 45(2): 643–73.

————. 2008a. "Forming Priors for DSGE Models. and How It Affects the Assessment of Nominal Rigidities." Working paper 13741. Cambridge, Mass.: National Bureau of Economic Research.

————. 2008b. "Monetary Policy Analysis with Potentially Misspecified Models." *American Economic Review*. (forthcoming).

Del Negro, M., F. Schorfheide, F. Smets and R. Wouters. 2007. "On the Fit of New Keynesian Models." *Journal of Business and Economics Statistics* 25(2): 123–43.

Galí, J. and T. Monacelli. 2005. "Monetary Policy and Exchange Rate Volatility in a Small Open Economy." *Review of Economic Studies* 72(3): 707–34.

Lubik, T. and F. Schorfheide. 2006. "A Bayesian Look at New Open Economy Macroeconomics." In *NBER Macroeconomics Annual 2005*, edited by M. Gertler and K. Rogoff. MIT Press.

————. 2007. "Do Central Banks Respond to Exchange Rate Movements? A Structural Investigation." *Journal of Monetary Economics* 54(4): 1069–87.

Mishkin, F. and K. Schmidt-Hebbel, eds. 2007. *Monetary Policy under Inflation Targeting*. Santiago: Central Bank of Chile.

Ravenna, F. 2007. "Vector Autoregressions and Reduced-Form Representations of DSGE Models." *Journal of Monetary Economics* 54(7): 2048–64.

Sims, C.A. 1999. "The Role of Interest Rate Policy in the Generation and Propagation of Business Cycles: What Has Changed since the 30s?" In *Beyond Shocks: What Causes Business Cycles?,* edited by Jeffrey C. Fuhrer and Scott Schuh. Federal Reserve Bank of Boston.

————. 2002. "Solving Rational Expectation Models." *Computational Economics* 20(1–2): 1–20.

Smets, F. and R. Wouters. 2003. "An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area." *Journal of the European Economic Association* 1(5): 1123–75.

Overoptimism, Boom-Bust Cycles, and Monetary Policy in Small Open Economies

Manuel Marfán
Central Bank of Chile

Juan Pablo Medina
Central Bank of Chile

Claudio Soto
Central Bank of Chile

In the 1990s, several emerging market economies, such as Chile, Mexico, and a number of southeast Asian countries, displayed episodes of peaking growth rates combined with increasing current account deficits and appreciating currencies, which ended with abrupt reversions in capital flows and recessions.[1] In all cases, optimism about future prospects was strong prior to the recessions. Mexico was negotiating both its entrance into the North America Free Trade Agreement (NAFTA) and its membership in the Organization for Economic Cooperation and Development (OECD). Chile had undergone a smooth transition to democracy. Investors were increasingly enthusiastic about the prospects of harvesting the benefits of the market reforms introduced both in the previous period and under the new democracy. The southeast Asian economies, in turn, had their own reasons for optimism based on their impressive growth record of previous years. In all cases, optimism was grounded on reasonable arguments, but the prospects of future economic growth could not be estimated accurately.

We thank Jorge Selaive for his contributions to a previous version of the paper and Emerson Melo for research assistance. Comments by Rodrigo Cerda, José De Gregorio, Pablo García, and Klaus Schmidt-Hebbel are also greatly appreciated.

1. A similar pattern can also be observed in industrial economies, such as the United States at the end of the 1990s, and in emerging markets in the late 1970s.

Monetary Policy under Uncertainty and Learning, edited by Klaus Schmidt-Hebbel and Carl E. Walsh, Santiago, Chile.

In this paper, we show that overoptimistic perceptions of the future by domestic private agents—that is, domestic "exuberance"—could have been a cause of the boom-bust cycles observed in some emerging economies in the 1990s. To that end, we develop a multi-sector dynamic stochastic general equilibrium (DSGE) model for a small economy with short-run stickiness in prices and wages and with expectations-driven boom-bust cycles. We show that under standard parameterization, the model is able to closely match most of the stylized facts observed in the boom-bust episodes in emerging markets. In the model, private agents are rational and forward looking, so their current decisions rely on their assessment of future productivity prospects. An overoptimistic assessment of future productivity makes them accumulate excess capital and over-increase their consumption, leading to a boom that is accompanied by a current account deficit. When agents realize that productivity will grow by less than expected, they must readjust their investment and consumption profiles, generating a current account reversal and a recession.

Our analytical approach closely follows Christiano and others (2007). We diverge from their work, however, in arguing that overoptimism about productivity trends, rather than productivity level changes, is the source of boom-bust cycles in open economies, as occurred in the 1990s. We show that if productivity levels follow a stationary process, then news about future productivity improvements are not able to replicate the real currency appreciation and the current account deterioration along the boom, as observed in the data. This result is related to the work of Aguiar and Gopinath (2007), who show that the observed countercyclicality of the current account in emerging economies can be explained by productivity trend shocks in a standard real business cycle model.

According to our model, a boom-bust cycle generated by domestic agents' overoptimism is observationally equivalent to a cycle driven by exogenous fluctuations in foreign financial conditions. Several authors claim that swings in external financial conditions were significant factors behind the observed patterns of macroeconomic variables in the 1990s in many emerging markets (Neumeyer and Perri, 2005; Uribe and Yue, 2006; Valdés, 2007). In this sense, our results can be interpreted as a plausible complementary explanation for the episodes of abrupt current account deterioration in emerging markets in the 1990s.[2] Among the policy implications, our model

2. Our results do not provide a formal test in favor of overoptimism as an explanation of boom-bust cycles in emerging economies against other theories based on fluctuations in the fundamentals.

shows that the trade-offs faced by monetary policy in a boom-bust cycle driven by expectations are not trivial. If the central bank tries to stabilize output, the result will be a large fall in inflation and a contraction in output in the tradable goods sector. On the other hand, if the central bank targets inflation strictly, then the boom in activity, the current account deterioration, and the exchange rate appreciation will be larger, and the subsequent recession more severe. Finally, if the monetary authority adjusts the interest rate to reduce exchange rate fluctuations, then the perverse effects on the domestic tradable goods sector are only prevented in the short run, while the boom-bust cycle is amplified in other variables.

The idea of expectations-driven macroeconomic fluctuations goes back at least to Pigou (1926). Recently, this hypothesis has received renewed attention in modern macroeconomics. Marfán (2005) analyzes boom-bust cycles provoked by excess optimism and concentrates mainly on the role of fiscal policy in an extended Mundell-Fleming context. The optimist-pessimist mood of the private sector in his model is completely exogenous. Beaudry and Portier (2004, 2007), Jaimovich and Rebelo (2006, 2007), Mertens (2007), and Christiano and others (2007) present different unique-equilibrium rational expectation models in which business cycles are generated by changes in expectations regarding productivity prospects. Jaimovich and Rebelo (2006, 2007), in particular, analyze the comovements of a set of variables generated in response to unmaterialized productivity shocks. They show that in a closed economy, adjustment costs in investment or labor (or both), variable capital utilization, and weak wealth effects on labor supply are key to replicating the comovements observed in the data. In an open economy, variable capital utilization turns out to be unimportant. Christiano and others (2007) emphasize the role played by the monetary policy at generating expectation-driven boom-bust cycles. Using a sticky-price, sticky-wages model they show that to generate a sizeable output expansion and a boom in stock prices in response to news about increased future productivity, monetary policy has to respond aggressively to the fall in inflation. The boom generated by overoptimistic perceptions about future productivity is thus amplified by a loose monetary policy. Mertens (2007) shows that an expectations-driven real business cycle (RBC) model is able to replicate relevant stylized facts of Korea's sudden stop in the late 1990s. Some studies on the Chilean crisis of 1982 also assign a responsibility to this boom-bust episode to an erroneous perception by private agents regarding their wealth (Barandiarán 1983, Schmidt-Hebbel, 1988).

The expectations-driven business cycle approach in this literature is related to the literature on multiple equilibria and sunspot cycles (Farmer, 1993). It can also be viewed as complementary to the literature on rational herding and information cascade cycles, which emphasizes how improper aggregation of information may occasionally result in cycles led by nonfundamentals (Banerjee, 1992; Chamley and Gale, 1994; Caplin and Leahy, 1993; Zeira, 1994). In this paper, we examine whether the quantitative implications of (rational/nonsystematic) aggregate forecast errors can explain the observed pattern of boom-bust cycles in small open economies within a fully specified dynamic model that features a unique equilibrium.

The remainder of the paper is divided into four sections. Section 1 provides a motivation on the effects of economic reforms and innovations on the expected path for productivity, and describes some stylized facts for economies that went through boom-bust cycle episodes in the 1990s: namely, Chile, Korea, and Mexico. Section 2 presents a detailed description of the theoretical model used to evaluate the effects of overoptimism in small open economies. Section 3 analyzes the dynamics of the model and discusses the tradeoffs faced by monetary policy. The final section summarizes our main findings.

1. Structural Reforms and Boom-bust Cycles in Emerging Markets

Several emerging market economies engaged in reforms in the 1980s and the 1990s. Moreover, with the fall of the Berlin wall, at the beginning of the 1990s, a generalized stimulus for accelerating and expanding market globalization was perceived. At the same time, emerging economies had resumed access to voluntary financial flows under favorable conditions, and trade markets were mutating toward increasing levels of regional integration. In this context, the international forums increasingly concentrated on the new international financial architecture, and the expansion of market institutions. While this macroeconomic context was prone to boost productivity, the actual effect of the reforms was hard to evaluate, given that the scenario was without precedent. It is possible, therefore, that private agents would have overestimated the effects of the reforms on future productivity.

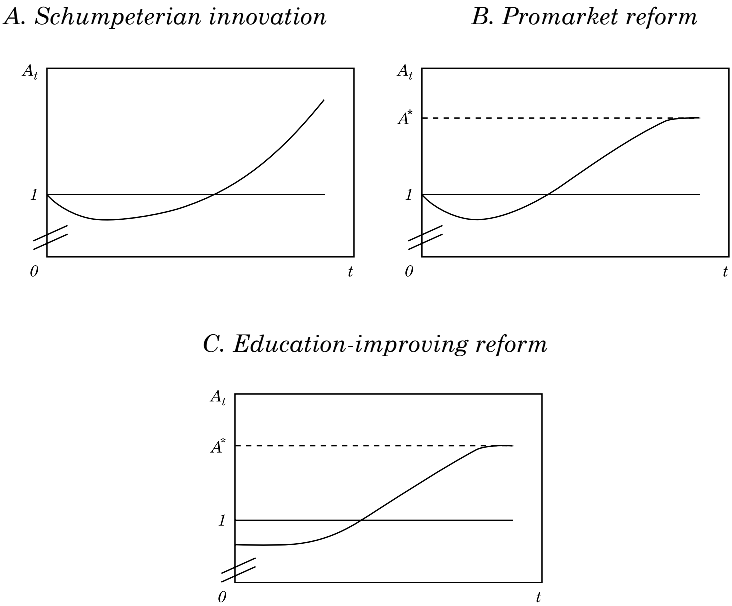
1.1 Structural Reforms, Innovations, and Productivity

Both, structural reforms and innovations give rise to delayed changes in productivity. For example, if $F_{i,t}$ denotes the production function of a generic firm i at time t, the concomitant production function after a reform or a systemic innovation that stimulates productivity would be $F^{'}_{i,t} = F_{i,t}A_t$, where A_t measures the impact of the reform or innovation on productivity at time t. Figure 1 presents examples of the effect on productivity of different types of innovations and reforms initiated at $t = 0$. First, we illustrate the effect on productivity of a Schumpeterian innovation –i.e. the steam machine, electricity, information and communication technology, and so forth. Initially, the destruction of capital, jobs, skills, and public goods related to the old technology dominates the creation process of the blossoming innovation. This would reduce measured productivity. At longer horizons, the benefits of the new technology outpace the costs of destroying the old one and measured productivity rises (the A_t curve could potentially turn concave at a very long horizon, showing decreasing returns).

Second, we present the case of a promarket reform (such as a trade-opening reform). Initially, measured productivity may fall as costly reallocation of resources from different sectors lead to temporary decreases in output. As time goes by, measured productivity increases and converges to a long-term productivity gain, A^*, once the reform is completely internalized. A similar pattern would follow from a reform intended to improve human capital (education-improving reform). There is an initial period in which significant resources are diverted from other activities to implement the reform, with no immediate productive effects. The benefits of the reform start to be harvested when the new well-educated generations graduate, and the reform is completed once the labor force is entirely educated.

In all the innovations or reforms described, there is no prior history to provide economic agents with the basis for accurately predicting its impact through time. Agents may know the functional form followed by A_t through time, but the values of certain parameters such as A^* are initially uncertain. In this context, agents react first by setting notional values for A^*, which may differ from their actual values. In all cases, it takes time for the reforms to materialize into actual productivity gains, making it hard to evaluate ex ante their real impact.

Figure 1. Reforms and Their Impact on Productivity



Source: Authors' drawings.

1.2 Some Stylized Facts

There is a set of stylized facts that characterizes the boom-bust episodes in emerging market that engaged in reforms. In this subsection we describe some of them for three cases: Chile, Korea, and Mexico in the 1990s.

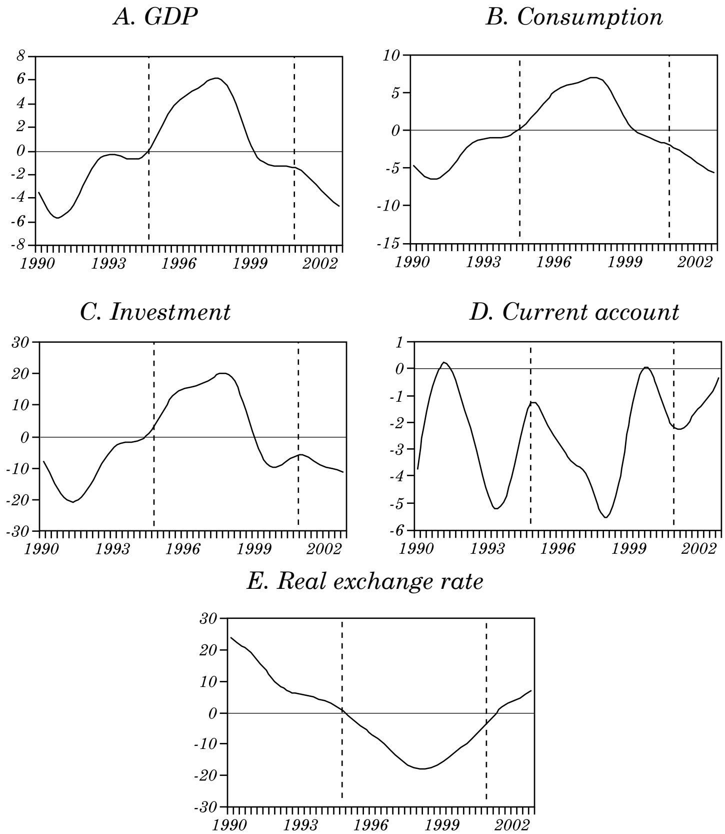
Chile introduced a number of reforms in the 1970s and 1980s. The democratic administrations that started in 1990 reinforced and deepened the structural reforms and gave a high priority to overall macroeconomic equilibrium. The signal to economic agents was that a strong stimulus to productivity growth was coming. Jadresic and Zahler (2000) claim based on time-series modeling, that key factors underlying the rapid productivity growth in the 1990s were precisely the deepening of democracy and the introduction of new structural reforms. Mexico implemented a privatization plan in the late 1980s, followed by a trade liberalization policy in the 1990s that involved the future opening of its economy to trade and capital flows with the United States and Canada. Korea experienced a long period of rapid growth, low inflation, and a sustained improvement in living standards before being hit by the financial crisis of 1997. High domestic savings and investment contributed to Korea's rapid transformation. The government had begun an economic reform program, which gained momentum in 1993–96, to gradually liberalize financial markets and the capital account.

Figures 2, 3, and 4 present some stylized facts for the three economies for the period 1990–2002.[3] In all three cases, we identify a phase in which output rises above trend together with an increase in investment and consumption. During the boom phase, we also observe a real currency appreciation and current account deterioration in the three countries. For Mexico, the expansion in output was less dramatic than in Korea and Chile, but the consumption boom was comparable to those countries. All three cases experienced an abrupt reversion of the boom, with a fall in output, consumption, and investment and a steep reversion of the current account deficit. In Mexico and Korea, the bust coincided with a depreciation of the currency of almost 40 percent. In Chile, the depreciation of the currency during the bust was slower than in the other two countries.

The boom-bust cycle in these three countries involved swings in output and consumption of about 10 percent in a brief period of time. The swings were much larger in the case of investment, with differences of more than 20 percent from peak to trough. In Mexico and Chile, the contraction of the current account deficit did not lead to a surplus in this variable. For Korea, the current account deficit of almost 6 percent of gross domestic product (GDP) was followed by a similar surplus a couple of years after the peak of the boom. Unlike Chile and Mexico, Korea had a stunning recovery from the crisis and output regained its precrisis level. In the case of Chile, growth has not recovered the 1990s rate.

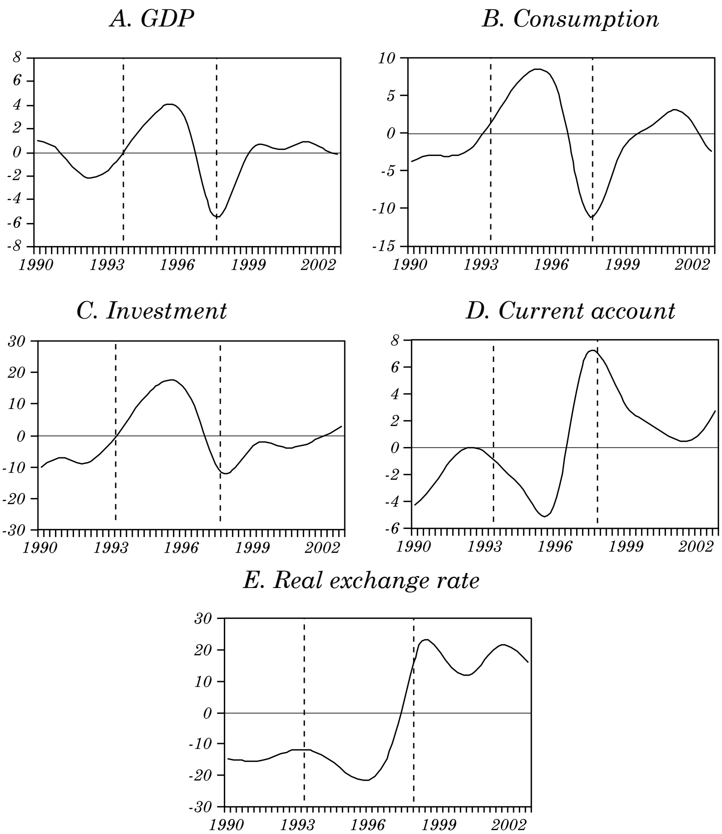
3. To build the stylized facts, we use Chilean quarterly data for the period 1990:1 to 2002:4 from the Central Bank of Chile and the National Institute of Statistics (INE). For Mexico and Korea, the source is the International Monetary Fund's *International Financial Statistics* (IFS). For all series, we applied a Hodrick-Prescott filter with a large smoothing parameter ($\lambda = 3 \times 10^6$) to obtain an almost lineal trend. Once we filtered the series, we computed the respective cycles. We then proceeded to filter these series again to obtain a smoother pattern.

Figure 2. Stylized Facts: Chile



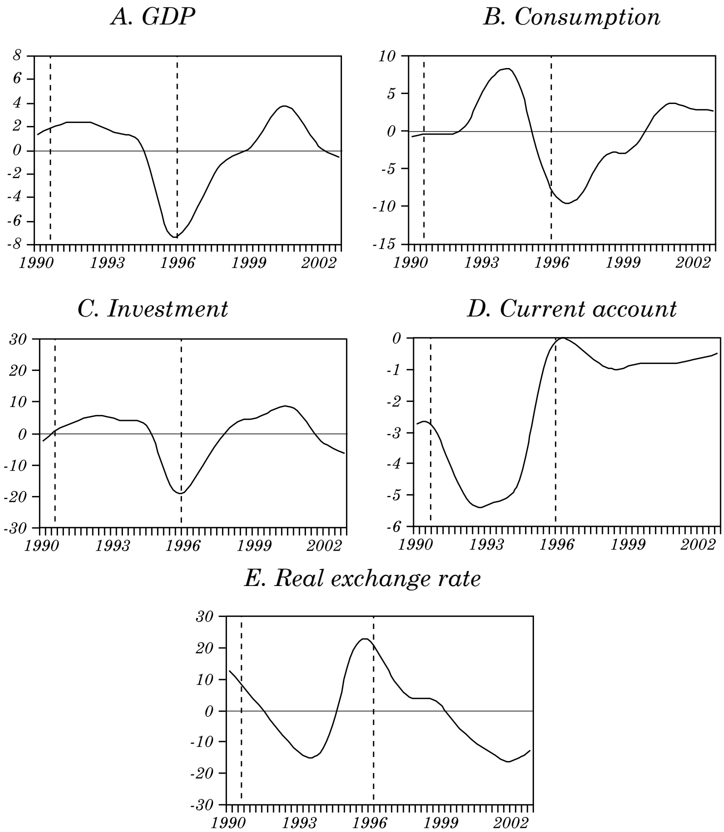
Source: Authors' calculations.

Figure 3. Stylized Facts: Korea



Source: Authors' calculations.

Figure 4. Stylized Facts: Mexico



Source: Authors' calculations.

2. Model Economy

In this section, we present a multi-sector small open economy model with short-run nominal and real rigidities. The model is aimed at replicating prominent features of the business cycles of emerging market economies. There are two domestic productive sectors: one that produces tradable goods (H) and another that produces nontradable goods (N). Domestic agents also import foreign goods (F). Prices and wages are sticky in the short run, and the exchange rate pass-through to imported goods price is incomplete in the short run. Households exhibit habits in their preferences, investment is subject to incremental adjustment costs, and the capital utilization rate is variable. The introduction of nominal and real rigidities is meant to generate richer and more realistic propagation mechanisms.

2.1 Households

The domestic economy is inhabited by a continuum of households indexed by $j \in [0,1]$. At time t, household j maximizes the expected present value of its utility, which is given by

$$U_t(j) = E_t \left\{ \sum_{i=0}^{\infty} \beta^i \left[\begin{array}{l} \log\left(C_{t+i}(j) - hC_{t+i-1}\right) \\ + \dfrac{\zeta_M}{\mu}\left(\dfrac{M_{t+i}(j)}{P_{C,t+i}}\right)^{\mu} - \zeta_L \dfrac{l_{t+i}(j)^{1+\sigma_L}}{1+\sigma_L} \end{array} \right] \right\}, \tag{1}$$

where $l_t(j)$ is labor effort, $C_t(j)$ is the household's total consumption, and $M_t(j)$ corresponds to nominal balances held at the beginning of period t. Parameter σ_L is the inverse real-wage elasticity of labor supply. Habit formation in preferences is determined by parameter h. Household j consumes a basket composed of tradable goods, C_T, and nontradable goods, C_N:

$$C_t(j) = \left[\alpha_C^{1/\eta_C} \left(C_{T,t}(j)\right)^{\frac{\eta_C - 1}{\eta_C}} + (1-\alpha_C)^{1/\eta_C} \left(C_{N,t}(j)\right)^{\frac{\eta_C - 1}{\eta_C}} \right]^{\frac{\eta_C}{\eta_C - 1}}.$$

Traded goods are a composite of domestically produce tradable goods (H) and imported goods (F),

$$C_{T,t}(j)=\left[\gamma_C^{1/\omega_C}(C_{H,t}(j))^{\frac{\omega_C-1}{\omega_C}}+(1-\gamma_C)^{1/\omega_C}(C_{F,t}(j))^{\frac{\omega_C-1}{\omega_C}}\right]^{\frac{\omega_C}{\omega_C-1}}.$$

Parameters α_C and γ_C determine the share of each type of goods in the consumption basket, while η_C and ω_C are the associated price elasticities. By minimizing the cost of the consumption basket and aggregating all households, we obtain the aggregate demands for the three types of goods. The consumer price index (CPI) is given by

$$P_{C,t}=(\alpha_C P_{T,t}^{1-\eta_C}+(1-\alpha_C)P_{N,t}^{1-\eta_C})^{\frac{1}{1-\omega_C}},$$

where $P_{T,t}$ is the price index of the tradable consumption basket (which includes imported and domestic tradable goods), and $P_{N,t}$ is the price index of nontradable goods.

2.1.1 Consumption-savings decisions

Households have access to three types of assets: money, $M_t(j)$; one-period noncontingent foreign bonds (denominated in foreign currency), $B_t^*(j)$; and one-period domestic contingent bonds, $D_{t+1}(j)$, which pays out one unit of domestic currency in a particular state (that is, state-contingent securities). The budget constraint of household j is given by

$$P_{C,t}C_t(j)+E_t\{d_{t,t+1}D_{t+1}(j)\}+\frac{e_tB_t^*(j)}{(1+i_t^*)\Theta(b_t)}+M_t(j)=$$
$$W_t(j)l_t(j)+\Pi_t(j)-\tau_t+D_t(j)+e_tB_{t-1}^*(j)+M_{t-1}(j),$$

where $\Pi_t(j)$ are profits received from domestic firms, $W_t(j)$ is the nominal wage set by the household, τ_t is per capita lump-sum net taxes from the government, and e_t is the nominal exchange rate (expressed as units of domestic currency per one unit of foreign currency). Variable $d_{t,t+1}$ is the period t price of one-period domestic contingent bonds normalized by the probability of the occurrence of the state.

Assuming the existence of a full set of contingent bonds ensures that the consumption of all households is the same, independently of the labor income they receive each period. Variable i_t^* is the interest rate on foreign bonds denominated in foreign currency, and $\Theta(.)$ is a premium domestic households have to pay when borrowing abroad. This premium is a function of the net foreign asset positions relative to GDP, $b_t = e_t B_t^*/P_{Y,t}Y_t$ where $P_{Y,t}Y_t$ is nominal GDP and B_t^* is the aggregate net asset position of the economy.[4]

Each household chooses a consumption path and the composition of its portfolio by maximizing equation (1) subject to its budget constraint. The first-order conditions on different contingent claims over all possible states define the following Euler equation for consumption:

$$\beta E_t \left\{ (1+i_t) \frac{P_{C,t}}{P_{C,t+1}} \left(\frac{C_{t+1}(j) - hC_t}{C_t(j) - hC_{t-1}} \right) \right\} = 1, \tag{2}$$

where i_t is the domestic risk-free interest rate. From this expression and the first-order condition with respect to foreign bonds denominated in foreign currency, we obtain the following expression for the uncovered interest parity (UIP) condition:

$$\frac{1+i_t}{(1+i_t^*)\Theta(b_t)} = E_t \frac{e_{t+1}}{e_t} + \mathrm{cov}_t, \tag{3}$$

where cov_t is a covariance term that disappears in the log-linear version of the model.

2.1.2 Labor supply and wage setting

Each household j is a monopolistic supplier of a differentiated labor service. There is a set of perfectly competitive labor service assemblers that hire labor from each household and combine it into an aggregate labor service unit. This labor unit is then used as an input in production in the domestic tradables (*H*) and nontradables

4. We assume that $\Theta(.) = \Theta$ and $\Theta' b/\Theta = \theta$ in the steady state. When the country is a net debtor, θ corresponds to the elasticity of the upward-slopping supply of international funds. This premium is introduced mainly as a technical device to ensure stationarity (see Schmitt-Grohé and Uribe, 2003).

(N) sectors. Cost minimization by labor unit assemblers gives rise to demands for each type of labor service, which are a function of the corresponding relative wages.

Following Erceg, Henderson, and Levin (2000), we assume that wage setting is subject to a nominal rigidity à la Calvo (1983). In each period, each type of household faces a probability $1 - \phi_L$ of being able to reoptimize its nominal wage. In this setup, the parameter ϕ_L determines the degree of nominal rigidity in wages. We assume that all those households that cannot reoptimize their wages follow an updating rule considering a geometric weighted average of past CPI inflation and the inflation target set by the authority, $\bar{\pi}$. Once a household has set its wage, it must supply any quantity of labor service demanded at that wage. A particular household j that is able to reoptimize its wage at t must solve the following problem:

$$\max_{W_t(j)} = E_t \left\{ \sum_{i=0}^{\infty} \phi_L^i \Lambda_{t,t+i} \left[\begin{array}{l} \dfrac{\Gamma_{W,t}^i W_t(j)}{P_{C,t+i}} l_{t+i}(j) \\ -\zeta_{L,t} \dfrac{l_{t+i}(j)^{1+\sigma_L}}{1+\sigma_L} (C_{t+i} - hC_{t+i-1}) \end{array} \right] \right\},$$

subject to labor demand and the updating rule for the nominal wage of agents who do not optimize, defined by function

$$\Gamma_{W,t}^i = \Gamma_{W,t}^{i-1} \left(1 + \pi_{t+i-1}\right)^{\chi_L} \left(1 + \bar{\pi}\right)^{1-\chi_L}$$

Variable $\Lambda_{t,t+i}$ is the relevant discount factor between periods t and $t + i$.[5] These elements give rise to a Phillips curve for nominal wages that has backward- and forward-looking components.

2.2 Investment and Capital Goods

A representative firm owns and rents capital to firms producing in the domestic tradables (H) and nontradables (N) sectors. We assume that capital is specific to the sector that rents it. Hence, the representative firm decides how much of each type of capital to

5. Since utility exhibits habit formation in consumption, the relevant discount factor is given by $\Lambda_{t,t+i} = \beta^i[(C_t(j) - hC_{t-1}) / (C_{t+i}(j) - hC_{t+i-1})]$.

accumulate over time. The flow of investment devoted to produce new capital goods for sector J, $I_t(J)$, is assembled using the following technology:

$$I_t(J) = \left[\alpha_I^{1/\eta_I} I_{T,t}(J)^{\frac{\eta_I - 1}{\eta_I}} + (1-\alpha_I)^{1/\eta_I} I_{N,t}(J)^{\frac{\eta_I - 1}{\eta_I}}\right]^{\frac{\eta_I}{\eta_I - 1}},$$

for $J = H, N$, where

$$I_{T,t}(J) = \left[\gamma_I^{1/\omega_I} I_{H,t}(J)^{\frac{\omega_I - 1}{\omega_I}} + (1-\gamma_I)^{1/\omega_I} I_{F,t}(J)^{\frac{\omega_I - 1}{\omega_I}}\right]^{\frac{\omega_I}{\omega_I - 1}}$$

is a composite of tradable goods devoted to investment in sector J. Variable $I_{D,t}(J)$ corresponds to the amount of good $D = H, F, N$ used in assembling new capital goods for sector J.

The representative firm may adjust investment each period, but changing the flow of investment is costly. This assumption is introduced as a way to obtain more inertia in the demand for investment (see Christiano, Eichenbaum, and Evans, 2005).[6] Let $Z_t(J)$ and $u_t(J)$ be the rental price and the utilization rate of capital in sector J, respectively. The representative firm must solve the following problem for each type of capital:

$$V_t(J) = \max_{K_{t+i}(J), I_{t+i}(J), u_{t+i}(J)} E_t \left\{ \sum_{i=0}^{\infty} \Lambda_{t,t+i} \frac{\begin{array}{c} u_{t+i}(J) Z_{t+i}(J) K_{t+i}(J) \\ -P_{I,t+i} I_{t+i}(J) \end{array}}{P_{C,t+i}} \right\},$$

subject to the law of motion of the capital stock for sector J,

$$K_{t+1}(J) = \left[1 - \delta(u_t(J))\right] K_t(J) + S\left(\frac{I_t(J)}{I_{t-1}(J)}\right) I_t(J), \tag{4}$$

6. This assumption is a shortcut to more cumbersome approaches to modeling investment inertia, such as time-to-build models.

where $\delta(u_t)$ is the depreciation rate, which is a function of the capital utilization rate. We assume that $\delta(u_t)$ is an increasing function, which implies that a higher utilization rate depreciates physical capital faster than a lower rate. Function $S(.)$ characterizes the adjustment cost for investment. This adjustment cost function satisfies the following conditions: $S(1+g_y)=1$, $S'(1+g_y)=0$, $S''(1+g_y)=-\mu_S<0$, where g_y is the per capita growth rate of the economy in the steady state.

The optimality conditions for the above problem are as follows:

$$\begin{aligned}\frac{P_{I,t}}{P_{C,t}} = {} & \frac{Q_t(J)}{P_{C,t}}\left[S\left(\frac{I_t(J)}{I_{t-1}(J)}\right)+S'\left(\frac{I_t(J)}{I_{t-1}(J)}\right)\frac{I_t(J)}{I_{t-1}(J)}\right] \\ & -E_t\left\{\Lambda_{t,t+1}\frac{Q_{t+1}(J)}{P_{C,t+1}}\left[S'\left(\frac{I_{t+1}(J)}{I_t(J)}\right)\left(\frac{I_{t+1}(J)}{I_t(J)}\right)^2\right]\right\},\end{aligned} \tag{5}$$

$$\frac{Q_t(J)}{P_{C,t}} = E_t\left\langle\Lambda_{t,t+1}\left\{\frac{Z_{t+1}(J)}{P_{C,t+1}}+\frac{Q_{t+1}(J)}{P_{C,t+1}}\left[1-\delta\left(u_t(J)\right)\right]\right\}\right\rangle, \tag{6}$$

and

$$\frac{Z_t(J)}{P_{C,t}} = \delta'\left(u_t(J)\right)\frac{Q_t(J)}{P_{C,t}}. \tag{7}$$

The ratio $P_{I,t}/P_{C,t}$ is the real cost of producing new capital goods (that is, the price of the investment bundle deflated by the CPI), where

$$P_{I,t} = \left[\alpha_I P_{I_T,t}^{1-\eta_I}+(1-\alpha_I)P_{N,t}^{1-\eta_I}\right]^{\frac{1}{(1-\eta_I)}}$$

and

$$P_{I_T,t} = \left[\gamma_I P_{H,t}^{1-\omega_I}+(1-\gamma_I)P_{F,t}^{1-\omega_I}\right]^{\frac{1}{(1-\omega_I)}}.$$

Equations (5), (6), and (7) simultaneously determine the evolution of the shadow price of capital, $Q_t(J)$, real investment expenditure, and the capital utilization rate for each sector.

2.3 Domestic Production

There is a large set of firms that use a constant elasticity of substitution (CES) technology to assemble intermediate varieties into home goods sold to households, to firms producing new capital goods, and to foreign agents. There is also a large set of firms that use a similar CES technology to assemble intermediate varieties into nontradable goods sold to households and to firms producing new capital goods.

Let $Y_{N,t}$ be the total quantity of nontradable goods sold to domestic agents (households and the representative firm assembling new capital goods). The demand for a generic variety z_N to assemble nontradable goods is given by

$$Y_{N,t}(z_N) = \left(\frac{P_{N,t}(z_N)}{P_{N,t}} \right)^{-\varepsilon_N} Y_{N,t}, \tag{8}$$

where $P_{N,t}(z_N)$ is the price of variety z_N. Analogously, let $Y_{H,t}$ be quantity of home goods sold domestically, and $Y^*_{H,t}$ the quantity sold abroad. The demands for a particular variety z_H to assemble these goods are given by

$$Y^*_{H,t}(z_H) = \left(\frac{P^*_{H,t}(z_H)}{P^*_{H,t}} \right)^{-\varepsilon_H} Y^*_{H,t},$$

and (9)

$$Y_{H,t}(z_H) = \left(\frac{P_{H,t}(z_H)}{P_{H,t}} \right)^{-\varepsilon_H} Y_{H,t},$$

where $P_H(z_H)$ is the price of the variety z_H when used to assemble home goods sold in the domestic market, and $P^*_{H,t}(z_H)$ is the foreign-currency price of this variety when used to assemble home goods sold abroad. Variables $P_{H,t}$ and $P^*_{H,t}$ are the corresponding aggregate price indexes. The foreign demand for home goods, $Y^*_{H,t}$ is given by

$$Y^*_{H,t} = \zeta^* \left(\frac{P^*_{H,t}}{P^*_t} \right)^{-\eta^*} Y^*_t,$$

where Y_t^* is foreign output, ζ^* corresponds to the share of domestic intermediate goods in the consumption basket of foreign agents, and η^* is the price elasticity of demand.

Intermediate varieties in the tradables and nontradables sectors are produced by monopolistically competitive firms. These firms maximize profits by choosing the prices of their differentiated variety subject to the corresponding demands and the available technology. Let $Y_{J,t}(z_J)$ be the total quantity produced of a particular variety z_J in sector $J = H, N$. The available technology is given by

$$Y_{J,t}(z_J) = A_{J,t}\left[T_t l_t(z_J)\right]^{\eta_J}\left[u_t(J)K_t(z_J)\right]^{1-\eta_J}, \tag{10}$$

for $J = H, N$, where $l_t(z_J)$ is the amount of labor and $K_t(z_J)$ is the amount of physical capital used in production. Parameter η_J defines the shares of the different factors in production. The variable $A_{J,t}$ represents a stationary productivity shock common to all firms in sector J, while T_t is a stochastic trend in labor productivity that is common to both domestic sectors (H and N). Below we discuss the process followed by these shocks.

We assume that the price adjustment of the domestic varieties faces nominal rigidities à la Calvo. In every period, the probability that a firm producing home goods receives a signal for adjusting its price for the domestic market is $1-\phi_{H_D}$, and the probability of adjusting its price for the foreign market is $1-\phi_{H_F}$. Analogously, the probability that a firm producing nontradable varieties receives a signal for adjusting its price is $1-\phi_N$. These probabilities are the same for all firms, independent of their history. If a firm does not receive a signal, it updates its price following a simple rule that weights past inflation and the inflation target set by the central bank. Thus, when a firm receives a signal to adjust its price, it maximizes the discounted expected value of its profits, conditional on having to passively update its price for a number of periods and subject to equation (9) or (8). Given this pricing structure, the paths for inflation of domestic tradable (H) and nontradable (N) goods are given by new Keynesian Philips curves with indexation. In its log-linear form, inflation in sector J depends on both last period's inflation, expected inflation next period, and marginal cost in sector J.

2.4 Import Goods Retailers

We introduce local-currency price stickiness to allow for incomplete exchange rate pass-through into import prices in the short run. This feature of the model mitigates the expenditure-switching effect of exchange rate movements for a given degree of substitution between foreign and home goods.

There is a set of competitive assemblers that use a CES technology to combine a continuum of differentiated imported varieties to produce a final foreign good, Y_F. This good is consumed by households and used for assembling new capital goods. The optimal mix of imported varieties in the final foreign good defines the demands for each of them. In particular, the demand for variety z_F is given by

$$Y_{F,t}(z_F) = \left(\frac{P_{F,t}(z_F)}{P_{F,t}} \right)^{-\varepsilon_F} Y_{F,t}, \tag{11}$$

where ε_F is the elasticity of substitution among imported varieties, $P_{F,t}(z_F)$ is the domestic-currency price of imported variety z_F in the domestic market, and $P_{F,t}$ is the aggregate price of import goods in this market.

Importing firms buy varieties abroad and resell them domestically to assemblers. Each importing firm has monopoly power in the domestic retailing of a particular variety. They adjust the domestic price of their varieties infrequently, only when they receive a signal. The signal arrives with probability $1 - \phi_F$ each period. As in the case of domestically produced varieties, if a firm does not receive a signal, it updates its price following a passive rule that weights past inflation and the inflation target set by the central bank. Therefore, when a generic importing firm z_F receives a signal, it chooses a new price by maximizing the discounted sum of expected profits subject to the domestic demand for variety z_F (equation 11) and the updating rule.

Under this specification, changes in the nominal exchange rate will not immediately be passed through to prices of imported good sold domestically. Therefore, exchange rate pass-through will be incomplete in the short run. In the long run, firms freely adjust their prices, so the law of one price for foreign goods holds up to a constant.[7]

7. Formally, in the long run, $P_F = [\varepsilon_F / (\varepsilon_F - 1)]\, eP_F^*$.

2.5 Monetary Policy Rule

Monetary policy is modeled as a simple feedback rule for the interest rate. Under the baseline specification of the model, we assume that the central bank adjusts the policy rate in response to contemporaneous deviations of CPI inflation from the target and to deviations of total output from its balanced growth trend:

$$\frac{1+i_t}{1+i} = \left(\frac{1+i_{t-1}}{1+i}\right)^{\psi_i} \left(\frac{Y_t}{\overline{Y}_t}\right)^{(1-\psi_i)\psi_y} \left(\frac{1+\pi_t}{1+\overline{\pi}}\right)^{(1-\psi_i)\psi_\pi},$$

where $\pi_t = P_{C,t} / P_{C,t-1} - 1$ is consumer price inflation, i is the steady-state value of the nominal interest rate, $\overline{\pi}$ is the inflation target, and $\overline{Y}_t$ is the output trend.

2.6 Aggregate Equilibrium

Once firms producing domestic varieties set their prices, they must supply any quantity demanded at those given prices. Therefore, the market clearing condition for each variety implies that

$$Y_{N,t}(z_N) = \left(\frac{P_{N,t}(z_N)}{P_{N,t}}\right)^{-\varepsilon_N} Y_{N,t}$$

and

$$Y_{H,t}(z_H) = \left(\frac{P_{H,t}(z_H)}{P_{H,t}}\right)^{-\varepsilon_H} Y_{H,t} + \left(\frac{P^*_{H,t}(z_H)}{P^*_{H,t}}\right)^{-\varepsilon_H} Y^*_{H,t},$$

where $Y_{N,t} = C_{N,t} + I_{N,t}(H) + I_{N,t}(N)$ and $Y_{H,t} = C_{H,t} + I_{H,t}(H) + I_{H,t}(N)$ and where $Y^*_{H_F,t}$ was defined above. The equilibrium requires that total labor demanded by intermediate varieties producers must be equal to labor supply:

$$\int_0^1 l_t(z_H)dz_H + \int_0^1 l_t(z_N)dz_N = l_t,$$

where l_t is aggregate labor service. Also, the demand for physical capital in sector J has to be equal to the available amount:

$$\int_0^1 K_t(z_J)dz_J = K_t(J),$$

for $J = H, N$.

Using the equilibrium conditions in the goods and labor markets and the budget constraint of households and the government, we obtain the following expression for the evolution of the net foreign asset position:

$$\frac{b_t}{(1+i_t^*)\Theta(b_t)} = b_{t-1}\frac{P_{Y,t-1}Y_{t-1}}{P_{Y,t}Y_t} + \frac{P_{X,t}X_t}{P_{Y,t}Y_t} - \frac{P_{M,t}M_t}{P_{Y,t}Y_t},$$

where b_t is the aggregate net (liquid) asset position of the economy vis-à-vis the rest of the world relative to nominal GDP, and $P_{Y,t}Y_t = P_{C,t}C_t + P_{I,t}I_t + P_{X,t}X_t - P_{M,t}M_t$ is nominal GDP measured from the demand side. Nominal imports and exports are given by $P_{M,t}M_t = e_t P^*_{F,t}Y^*_{F,t}$ and $P_{X,t}X_t = e_t P^*_{H,t}Y^*_{H,t}$, respectively. The total quantity of imported goods is $Y_{F,t} = C_{F,t} + I_{F,t}(H) + I_{F,t}(N)$.

2.7 Model Calibration and Solution

To solve the model, we first tackle the nonstochastic steady state using numerical methods. We then solve the log-linearized decision rules from the behavioral equations and the equilibrium conditions of the model. To that end, we use the *QZ* factorization described in Uhlig (1997). Table 1 presents the value chosen for the structural parameters of the model. The calibration is meant to characterize quarterly data for the Chilean economy. Many of the parameters were taken directly from the literature; others were chosen to match long-run features of this economy. In our simulations, productivity shocks are calibrated to match the observed expansion in output during the Chilean boom of 1995–2001, as discussed above.

Table 1. Base Calibration

Parameter	*Description*	*Calibrated value*
β	Subjective discount factor (quarterly)	0.999
σ_L	Inverse of the elasticity of the labor supply	1.0
h	Habit formation coefficient	0.9
α_C	Share of tradable goods in the consumption basket	0.4
γ_C	Share of home goods in the tradables consumption basket	0.5
η_C	Elasticity of substitution between tradable and nontradable goods in the consumption basket	0.5
ω_C	Elasticity of substitution between home and foreign goods in the tradables consumption basket	1.0
ε_L	Elasticity of substitution among labor varieties	11
ϕ_L	Calvo probb in nominal wages	0.9
χ_L	Wage indexation to past inflation	0.9
α_I	Share of tradable goods in the investment basket [in $I(H)$ and $I(N)$]	0.6
γ_I	Share of home goods in the tradable investment basket [in $I(H)$ and $I(N)$]	0.5
η_I	Elasticity of substitution between tradable and nontradable goods in the investment basket [in $I(H)$ and $I(N)$]	0.5
ω_I	Elasticity of substitution between home and foreign goods in the tradable investment basket [in $I(H)$ and $I(N)$]	1.0
$\delta(1)$	Capital depreciation rate (annual) [in $I(H)$ and $I(N)$]	5.0 percent
μ_S	Elasticity of the adjustment cost in the flow of investment [in $I(H)$ and $I(N)$]	15
σ_I	Elasticity of the cost of capital utilization rate [$\delta''(1)/\ \delta'(1)$]	0.05
η_H	Labor share in the domestic tradable goods sector	0.65
η_N	Labor share in the nontradable goods sector	0.65
ε_N	Elasticity of substitution among nontradable varieties	11
ε_H	Elasticity of substitution among domestic tradable varieties	11
ε_F	Elasticity of substitution among imported varieties	11
ϕ_{H_D}	Calvo probb in prices of domestic tradable goods sold domestically	0.75
χ_{H_D}	Indexation to past inflation of domestic tradable goods sold domestically	0.50
ϕ_{H_F}	Calvo probb in foreign currency prices of domestic tradable goods sold abroad	0.75

Table 1. (continued)

Parameter	*Description*	*Calibrated value*
χ_{H_F}	Indexation to past inflation of domestic tradable goods sold abroad	0.50
ϕ_N	Calvo probb in prices of nontradable goods	0.75
χ_N	Indexation to past inflation of nontradable goods	0.50
ϕ_F	Calvo probb in prices of imported goods	0.75
χ_F	Indexation to past inflation of imported goods	0.50
ψ_i	Smoothing coefficient in the Taylor-type rule	0.80
ψ_π	Inflation coefficient in the Taylor-type rule	1.75
ψ_y	Output coefficient in the Taylor-type rule	0.20
η_F	Elasticity of the foreign demand for domestic tradable goods	0.50
θ	Elasticity of the external premium to the debt-to-GDP ratio	0.00001
NX/Y	Steady-state net-exports-to-GDP ratio	2 percent
CA/Y	Steady-state current-account-to-GDP ratio	–2 percent
g_y	Steady state GDP growth rate	5 percent
ρ_{a_H}	Persistence of productivity level shock in sector *H*	0.999
ρ_{a_N}	Persistence of productivity level shock in sector *N*	0.999
ρ_T	Persistence of productivity trend shock	0.999
ρ_{i*}	Persistence of productivity foreign financial conditions shock	0.999

Source: Authors' calculations.

3. Boom-Bust Cycles in Small Open Economies

We take Chile as a reference country and utilize the model described in the previous section to evaluate the qualitative and quantitative implications of boom-bust cycles driven by expectations. Before considering the case of overoptimism about future productivity, we analyze a case of favorable external financial conditions that are abruptly reversed.

In what follows, we define the real exchange rate in the model as the relative price of domestic tradable (*H*) and nontradable (*N*) goods. The implied evolution of measured total factor productivity (TFP) is estimated in the model as an aggregate Solow residual (without adjusting for the capital utilization rate). We construct a similar measure using actual data for Chile.[8] Tobin's Q is identified in the data with the stock market price, which in the case of Chile corresponds to an aggregate price index (IPSA). In the data, labor is measured as the ratio of formal employment to the working age population, and the real wage corresponds to an index of labor costs.[9]

3.1 Foreign Financial Condition Reversal

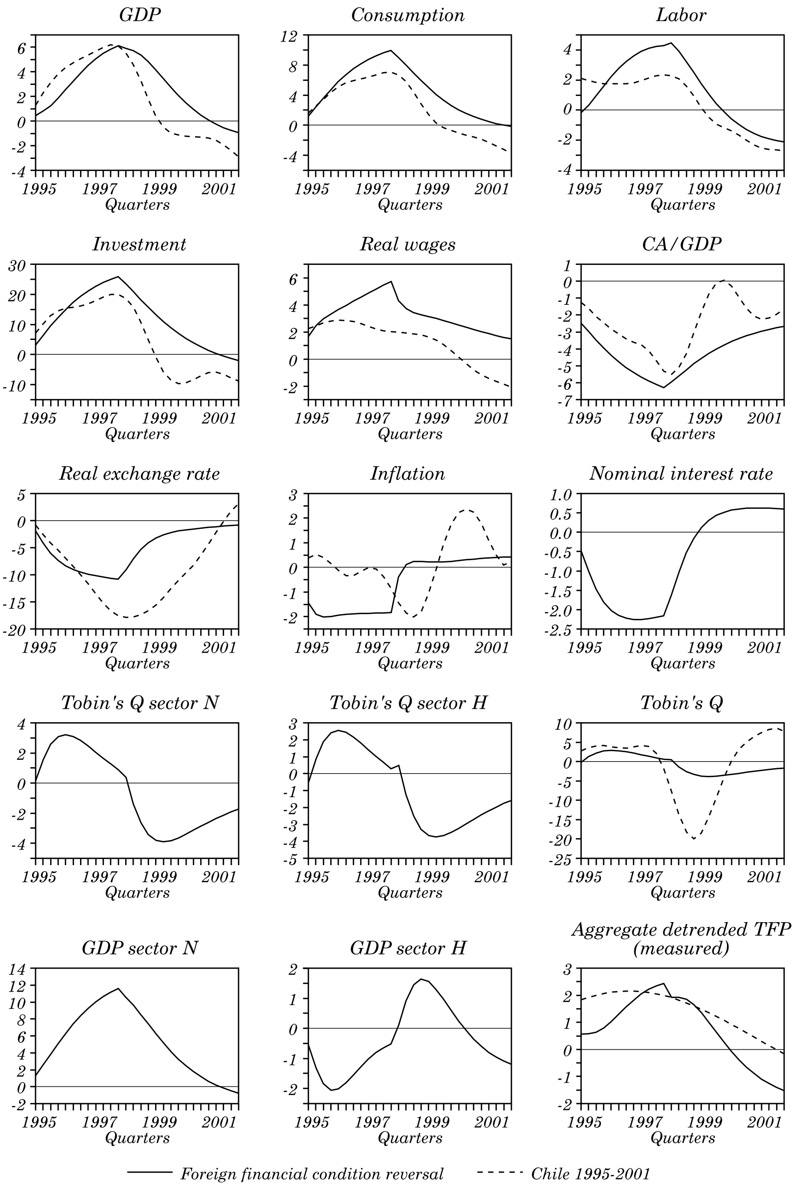
According to several authors, the boom-bust cycle in many emerging market economies in the 1990s was a consequence of changes in external financial conditions. This conclusion is based on the observation that periods of favorable external financial conditions are associated with economic expansions, while depressed economic activity coincides with periods of less beneficial foreign financial conditions (see, for example, Neumeyer and Perri, 2005; Uribe and Yue, 2006; Valdés 2007). Favorable external financial conditions in the early 1990s implied large capital flows to emerging market economies, which produced an economic boom coupled with real exchange rate appreciations and current account deficits. The boom phase was then followed by an abrupt worsening in foreign financial conditions, triggered by the Asian crisis, which would have led to recessions.

Using our model, we analyze the case of an exogenous, highly persistent decrease in the foreign interest rate (i^*). This captures the idea of a relaxation of the foreign financial conditions. Then, we assume that suddenly there is an exogenous increase on the foreign interest rate back to its original level. We calibrate the size of the shock so that the boom in output roughly coincides with the data for Chile. Figure 5 presents the results of this exercise. The model produces expansions in output, labor, consumption, and investment, which are sharply reversed when the foreign interest rate returns back to its original level. During the expansion, the real exchange appreciates by 10 percent, and the current account deficit (as a percentage of GDP) peaks near 6 percent. Contrary to what the data show, the model predicts an initial fall in inflation and a subsequent rise in this variable as the exchange rate depreciates. Despite the muted pass-through from exchange rate to domestic prices, the fall in inflation is due to the

8. Formally, $\ln(\mathrm{TFP}_t) = \ln(Y_t) - \eta\ln(l_t) - (1 - \eta)\ln(K_t)$, where η is the labor share in aggregate output.

9. To construct the cyclical components for these series, we follow the same procedure described in footnote 3.

Figure 5. Foreign Financial Condition Reversal



Source: Authors' calculations.

initial appreciation of the currency. The episode is accompanied by a rise in Tobin's Q for both types of capital. The boom in total output is driven by the evolution of output in the nontradable goods sector. In fact, the real currency appreciation leads to an initial fall in output in the tradable goods sector. Overall, the story of a boom-bust cycle driven by changes in foreign financial conditions is able to satisfactorily account for the stylized facts for Chile.

3.2 Overoptimistic Perceptions

We now explore an alternative—though complementary—explanation for the boom-bust cycle based on the idea that, rather than being caused by external factors, the cycle was triggered by domestic private agents' misperception regarding future productivity prospects. As mentioned above, this idea has recently been formalized by Christiano and others (2007) in a fully specified closed economy model. We build on their approach to model overoptimistic news on future productivity improvements.

3.2.1 Productivity level shocks

We first assume that productivity in sector $J = N, H$ is governed by the following stationary process:

$$a_{J,t} = \rho_{a_J} a_{J,t-1} + \zeta_{a_J,t-p} + \varepsilon_{a_J,t}, \tag{12}$$

for $J = H, N$, where $a_{J,t} = \ln A_{J,t}$ and $\varepsilon_{a_J,t} \sim N(0,\sigma^2_{a_J})$ are independent and identically distributed (i.i.d.) innovations. The varaible $\zeta_{a_J,t-p}$ is a shock to the expected future productivity level p periods ahead and is uncorrelated with $\varepsilon_{a_J,t}$. This shock captures the idea discussed in section 1, that structural reforms lead to expected changes in productivity. Those changes take time to materialize, however, and the agents do not exactly know their effective impact on productivity. Here, we assume that at time t, private agents learn that a set of reforms were carried out and, given equation (12), they expect that productivity p periods ahead will be given by

$$E_t\left[a_{J,t+p}\right] = \rho^p_{a_J} a_{J,t} + \zeta_{a_J,t},$$

where $\zeta_{a_J,t} > 0$. At time $t + p$, agents learn that the productivity level changed by less than expected. To this end, we introduce a shock,

$\varepsilon_{a_J,t+p} < 0$, on productivity at $t + p$. Figure 6 presents the results of this exercise, assuming $p = 12$ and $\rho_{a_J} = 0.999$ together with actual data for Chile.[10] We consider a case in which the news affects the expected productivity levels in both sectors (H and N) equally.[11]

As in Christiano and others (2007), the expected gain in productivity produces a boom in output. In our case, this is mainly due to the boom in the tradable goods sector. In fact, output in the nontradables sector falls in the short run and then increases. Consumption initially falls, but then it slowly expands in response to the expected increase in productivity. Labor rises during the boom phase, in part as a result of sticky wages that contain real wages growth. When wages are flexible in our model, the labor expansion no longer holds.[12] This is consistent with Jaimovich and Rebelo (2007), who show that under flexible wages, in order to generate a boom in labor in response to expected gains in productivity, household preferences should exhibit a weak wealth effect on labor supply. In our case, preferences are standard, but the wealth effect on the labor supply is muted because of sticky wages. Total inflation falls with the output boom. The reason is that expected future productivity gains mean lower future marginal costs. Since inflation is forward looking, firms respond by currently lowering their prices, despite the rise in actual marginal cost associated with the growth of labor and the rise in real wages.

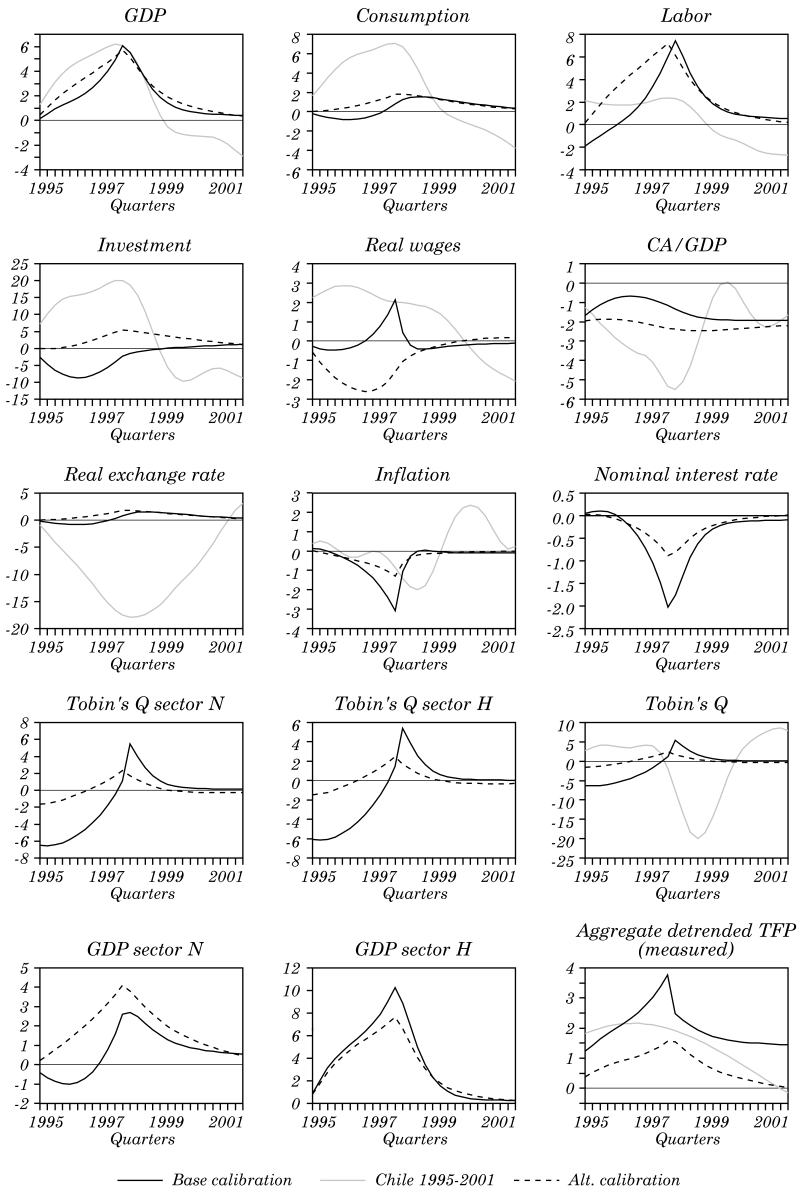
Despite the expected increase in future productivity, investment and Tobin's Q initially fall in both sectors when the signal on future productivity arrives. These variables then increase monotonically until the agents learn that productivity is lower than expected. These predictions on the behavior of investment and Tobin's Q during the news-induced boom-bust cycle are different from the predictions obtained by Christiano and others (2007). In their model, the boom-bust cycle in output coincides with a boom-bust cycle for investment and Tobin's Q. The reason investment responds this way to news about future productivity in Christiano and others (2007) is the presence of low wage indexation to past inflation and an aggressive inflation-targeting policy rule for the central bank. In their case, given the fall of inflation below target, monetary policy follows a

10. These productivity news shocks are highly persistent, but they are still transitory.

11. The real quantities in the figures correspond to the normalized effects of the productivity shock.

12. The simulation under flexible wages is available on request.

Figure 6. Productivity Level Signal



Source: Authors' calculations.

loose stance in response to the news shock. That helps raise Tobin's Q and induces firms to increase investment. Low indexation to past inflation, in turn, helps keep real wages rigid in the short run, amplifying the effects of overoptimistic shocks. In our calibration, we allow for a larger fraction of wages to be indexed to past inflation, and we specify a less hawkish inflation targeting regime –that is more in line with standard parameterization for the monetary policy rule. In figure 6, we also present an alternative calibration of the model, where we reduce the fraction of wages being indexed to past inflation (we set $\chi_L = 0.1$) and increase the reaction of the interest rate to deviations of inflation from target in the policy rule (we set $\psi_\pi = 2.0$). Under this alternative parameterization, the results of our simulation are in line with Christiano and others (2007): output, labor, consumption, investment, and Tobin's Q simultaneously feature a boom-bust cycle.

While the qualitative results of this last exercise resemble some features of the stylized facts discussed in section 1, they fall short in comparison with the observed size of the boom-bust cycle in investment and consumption in Chile. More importantly, the simulation misses two prominent features of the boom-bust cycles in emerging economies in the 1990s, namely, the real appreciation of the exchange rate and the current account deficit. Despite the boom in consumption and investment, which tends to produce a current account deficit, the exchange rate depreciation leads to an improvement in net exports that offsets the detrimental impact on this variable associated with the expansion in consumption and investment. In other words, the expenditure-switching effect induced by the currency depreciation dominates the intertemporal effect of the shock. The counterfactual behavior of the real exchange rate and the current account are even worse under the baseline calibration.

As we mentioned, one of the reasons for the boom after a news shock in the closed economy model of Christiano and others (2007) is the loose monetary policy response to the shock. In our model, a more expansive monetary policy is not enough to generate a sizable boom in expenditure. First, in a closed economy, the policy interest rate determines the equilibrium between domestic investment and savings. In an open economy, investment can differ from domestic saving. Moreover, both the domestic and foreign interest rates affect the cost of financing in an open economy. If the foreign interest rate is constant—and if the country does not face external borrowing constraints—then domestic monetary policy has less of an impact on the relevant cost of financing. As a result, the response of investment to a news shock is less intense. Second, the increase in private consumption in response to a future expected increase in productivity depends on the expected present value of private income. In a closed economy, the sequence of interest rates relevant for discounting future incomes is determined by monetary policy. Thus, if monetary policy is expansive in response to a signal shock, the perceived increase in the present value of income is amplified. In a small open economy facing a constant foreign interest rate, monetary policy alone does not determine the relevant interest rate for discounting expected future incomes. Hence, a loose monetary policy has a limited effect in amplifying the consumption boom.

As mentioned, the model fails at producing a real currency appreciation along the boom phase of the cycle. In a two-sector small open economy with tradable and nontradable goods, a real currency appreciation requires an increase in real wages. However, the fact that nominal wages are sticky in our model prevents an upward adjustment in real wages. This nominal stickiness is necessary to produce a sizable boom and to generate a procyclical response of employment to the shock.

3.2.2 Productivity trend shocks

Aguiar and Gopinath (2007) argue that in the case of emerging market economies, stochastic productivity trends, rather than productivity level shocks, are a major source of business cycle fluctuations. Moreover, these types of shocks are able to explain the observed comovement in major aggregate variables in these economies. In particular, shocks to the trend are better equipped to produce strongly countercyclical current accounts, as observed in emerging economies. More importantly, these shocks can generate these comovements without relying on household preferences that remove wealth effects in the labor supply.[13]

In what follows, we incorporate the approach of Aguiar and Gopinath (2007) to our analysis by assuming that news shocks refer to future changes in productivity trends. We assume that the natural logarithm of the stochastic trend of labor productivity, T_t, evolves according to the following expression:

13. See Correia, Neves, and Rebelo (1995) for an analysis of the aggregate dynamics in a small open economy without wealth effects in the labor supply.

$$s_{T,t} = s_{T,t-1} + (1-\rho_T)\ln(1+g_y) + \rho_T \Delta s_{T,t-1} + \zeta_{T,t-p} + \varepsilon_{T,t}, \tag{13}$$

where $s_{T,t} = \ln(T_t)$ and $\varepsilon_{T,t} \sim N(0, \sigma^2_T)$ are i.i.d. innovations. A shock $\zeta_{T,t-p}$ corresponds to a news of an increase in the labor productivity trend p periods ahead. As in the previous case, we assume that this shock is uncorrelated with $\varepsilon_{T,t}$. If agents receive a signal, $\zeta_{T,t} > 0$, at time t, they expect that productivity p periods ahead will grow faster:

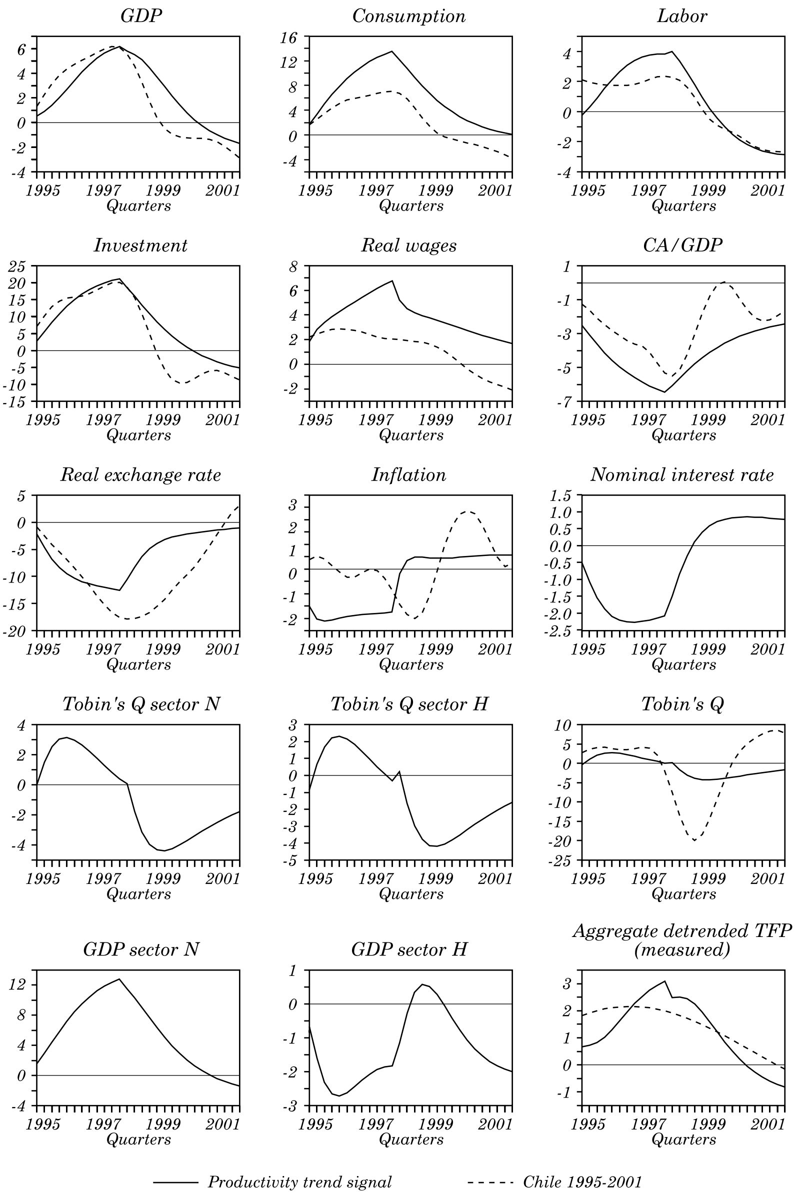
$$E_t\left[\Delta s_{T,t+p}\right] = \rho_T^p\left[\Delta s_{T,t} + (1-\rho_T)\ln(1+g_y)\right] + \zeta_{T,t}.$$

As in the case of news about productivity levels, we consider a shock $\varepsilon_{T,t+p} < 0$ in period $t+p$ to capture the idea that the news about expected productivity growth turns out to be overoptimistic ex post.

Figure 7 presents the trajectories of the endogenous variables to an expected shock to the trend in the future that does not materialize when $p = 12$ and $\rho_T = 0.999$. These trajectories were obtained using the baseline calibration of the model. The qualitative results of this shock are similar to those obtained with a positive signal to the productivity level in the future. We observe a boom-bust episode in output, labor, investment, and consumption. The quantitative pattern followed by the last three variables more closely resembles the data than in the previous case. Positive news regarding future productivity trend also generates a real appreciation of the exchange rate, as in the stylized facts reported earlier. The current account deficit reaches almost 7 percent, which is also very similar to what happened in Chile in the late 1990s, before the Asian crisis. In our model, the real appreciation explains why the output boom is mainly concentrated in the nontradable goods sector. This is completely different from the case of a productivity level signal, where the boom is explained by the expansion of the tradable goods sector. In the bust phase, as the expected increase in productivity growth does not materialize, the real exchange rate depreciates and the current account deficit reverses. There is a recession in output, and aggregate demand falls.

Despite the fact that productivity does not change, the measured TFP in the model rises above trend during the boom phase and falls during the bust phase. This pattern resembles the observed evolution of TFP constructed with actual Chilean data, which highlights the strong procyclicality of this variable. The model also predicts an

Figure 7. Productivity Trend Signal



Source: Authors' calculations.

increase in Tobin's *Q* during the boom and a subsequent fall during the recession. However, the size of the cycle of this variable is smaller than the observed pattern in Chilean stock prices in the 1990s. The model is also not able to closely replicate the behavior of inflation in Chile.

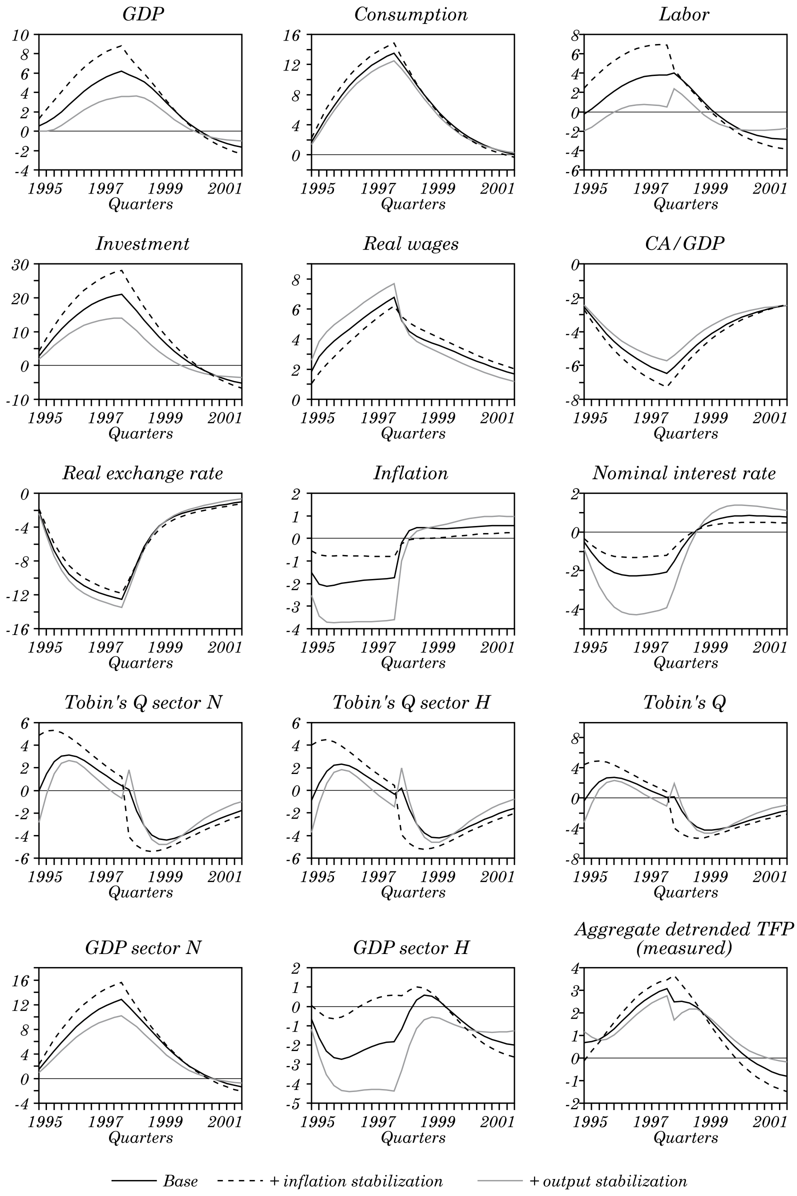
In our model, the boom-bust episode does not arise as a consequence of a loosening in monetary policy in response to a fall in inflation, as in Christiano and others (2007). Moreover, the dynamics of several variables in response to an overoptimistic signal regarding future productivity trends are observational equivalent to those obtained from a reversal in foreign financial conditions. Thus, overconfidence in productivity prospects is able to satisfactorily generate the boom-bust episode observed in emerging economies without any actual change in the economic fundamentals.

3.3 Monetary Policy Trade-offs

To explore the different monetary policy trade-offs in a boom-bust episode such as the one described here, we analyze the implications of alternative policy rules. First, we consider two alternative rules: one that reacts strongly to inflation and another that responds strongly to output. Second, we consider a rule in which monetary policy responds not only to output and inflation, but also to real exchange rate fluctuations. In all simulations below, we consider the responses after news about a future change in the productivity trend.

Figure 8 presents the baseline scenario, together with the results under a rule that is more aggressive to inflation and under a rule that is more aggressive to output fluctuations. If monetary policy focuses on following a more strict inflation target ($\psi_\pi = 3$), the boom in output, consumption, and investment would be larger because monetary policy takes a more expansive stance. The current account deficit would therefore also be larger, and the real appreciation would be slightly smaller. On the other hand, if monetary policy aggressively tries to stabilize output ($\psi_y = 0.8$), then it would induce a larger negative deviation of inflation from target and a larger currency appreciation. Given this currency appreciation, output stabilization is based proportionally more on tradables output than on nontradables output. The higher interest rate implied by this policy reduces the boom in Tobin's *Q* in both sectors and the current account deficit.

Figure 8. Stabilization of Inflation versus Output



Source: Authors' calculations.

In the case of a central bank that responds to exchange rate fluctuations, we modify the policy rule as follows:

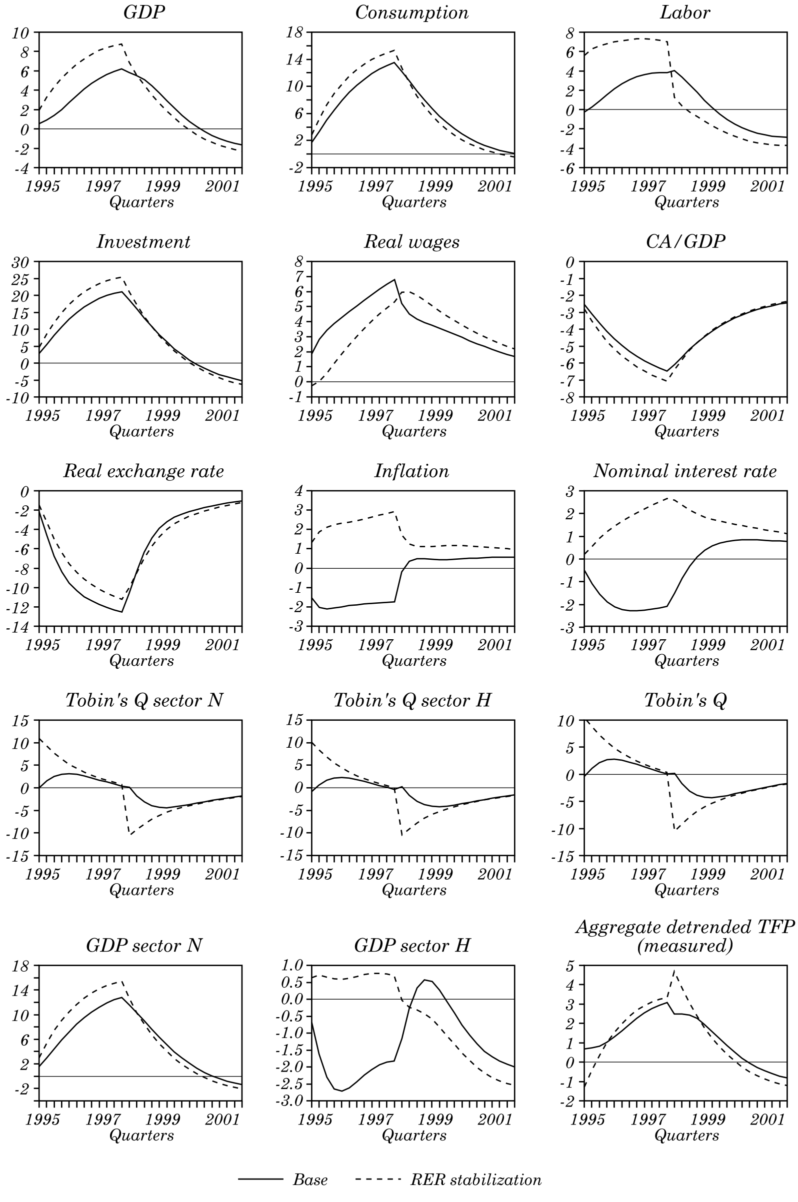
$$\frac{1+i_t}{1+i} = \left(\frac{1+i_{t-1}}{1+i}\right)^{\psi_i} \left(\frac{Y_t}{\overline{Y}_t}\right)^{(1-\psi_i)\psi_y} \left(\frac{1+\pi_t}{1+\overline{\pi}}\right)^{(1-\psi_i)\psi_\pi} \left(\frac{\mathrm{RER}_t}{\overline{\mathrm{RER}}}\right)^{(1-\psi_i)\psi_{rer}},$$

where RER_t is the real exchange rate and $\overline{\mathrm{RER}}$ is its steady-state value. We calibrate ψ_{rer} to 0.2. The rest of the parameters of the rule are same as in the baseline calibration. This policy rule is motivated by the Chilean experience in the 1990s, when the Central Bank of Chile simultaneously specified an inflation target and a target zone for the exchange rate to avoid excessive fluctuation in the latter. Figure 9 presents the results. Under this policy, monetary policy tends to be more expansive in response to the expected productivity gain. As a result, the increases in output, consumption, investment, and labor are larger than in the baseline case. The alternative rule reduces the volatility of the exchange rate, which prevents the perverse effects of the boom on the domestic tradable goods sector in the short run, but the current account deficit responds more sharply to the shock than in the baseline case, as a result of the investment and consumption booms. Inflation initially rises, but it falls after the bust because the reduction in marginal cost dominates the inflationary effects of the subsequent currency depreciation. Finally, by stabilizing the real exchange rate, the monetary policy exacerbates the boom-bust cycle in Tobin's Q and makes the predictions of the model quantitatively closer to the evolution of stock prices in Chile in the 1990s.

4. Conclusions

Using a small open economy DSGE model, we have shown that expected future gains in productivity that are not materialized ex post can generate a boom-bust cycle in output similar to what occurred in several emerging market economies in the 1990s. However, when people expect future productivity gains to be transitory level changes, the model predictions for the current account and the real exchange rate are not consistent with the observed pattern in those episodes. Moreover, the quantitative predictions for investment and consumption fall short with respect to what we observe in the data. This is the case even if we assume a strong monetary policy response to inflation and a low degree of wage indexation to past

Figure 9. Stabilization of the Real Exchange Rate



Source: Authors' calculations.

inflation. The reason is that in an open economy setup, the amplifying mechanism of monetary policy is unable to induce large consumption and investment booms.

When the expected future improvement in productivity corresponds to a trend shock, for which the productivity growth rate is expected to increase above its steady-state rate during some periods, the model predictions satisfactorily match the stylized facts observed in the data. Also, the boom generated by a news shock about future productivity trend affects the nontradable goods sector more deeply than the tradables sector. In fact, the real currency appreciation induced by the shock leads to a fall in output in the tradable goods sector. These results almost exactly replicate the results obtained under an exogenous reversal in the foreign financial conditions faced by the country.

Monetary policy faces important trade-offs in a boom-bust episode driven by overoptimistic perceptions about productivity improvements. On the one hand, if the central bank tries to stabilize output, it will exacerbate the fall in inflation and contraction in output in the tradable goods sector. On the other, if the central bank targets inflation more strictly, then the boom in activity, the current account deterioration, and the exchange rate appreciation will be larger and the subsequent recession more severe.

In the period under study, the Central Bank of Chile simultaneously pursued a target zone for the exchange rate and an inflation target. If we modify the policy rule in our model to capture this behavior by including an endogenous response of the interest rate to exchange rate fluctuations, then it does a better job of fitting the data. This type of policy only prevents the perverse effects of the boom on the domestic tradable goods sector in the short run, but it amplifies the boom-bust cycle in the other aggregate variables.

References

Aguiar, M. and G. Gopinath. 2007. "Emerging Market Business Cycles: The Cycle Is the Trend." *Journal of Political Economy* 115(1): 69–102.

Banerjee, A. 1992. "A Simple Model of Herd Behavior." *Quarterly Journal of Economics* 107(3): 797–817.

Barandiarán, E. 1983. "Nuestra crisis financiera." *Estudios Públicos* 12: 89–107.

Beaudry, P. and F. Portier. 2004. "An Exploration into Pigou's Theory of Cycles." *Journal of Monetary Economics* 51(6): 1183–216.

————. 2007. "When Can Changes in Expectations Cause Business Cycle Fluctuations in Neo-classical Settings?" *Journal of Economic Theory* 127(1): 458–77.

Calvo, G.A. 1983. "Staggered Prices in Utility-Maximizing Framework." *Journal of Monetary Economics* 12(3): 383–98.

Caplin, A. and J.V. Leahy. 1993. "Sectoral Shocks, Learning, and Aggregate Fluctuations." *Review of Economic Studies* 60(4): 777–94.

Chamley, C. and D. Gale. 1994. "Information Revelation and Strategic Delay in a Model of Investment." *Econometrica* 62(5): 1065–85.

Christiano, L.J., M. Eichenbaum, and C. Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113(1): 1–45.

Christiano, L.J., C. Ilut, R. Motto, and M. Rostagno. 2007. "Monetary Policy and Stock Market Boom-Bust Cycle." Northwestern University.

Correia, I., J. Neves, and S. Rebelo. 1995. "Business Cycles in a Small Open Economy." *European Economic Review* 39(6): 1089–113.

Erceg, C.J., D.W. Henderson, and A.T. Levin. 2000. "Optimal Monetary Policy with Staggered Wage and Price Contracts." *Journal of Monetary Economics*, 46(2): 281–313.

Farmer, R. 1993. *The Macroeconomics of Self-fulfilling Prophecies.* MIT Press.

Jadresic, E. and R. Zahler. 2000. "Chile's Rapid Growth in the 1990s: Good Policies, Good Luck, or Political Change?" Working paper 00/153. Washington: International Monetary Fund.

Jaimovich, N. and S. Rebelo. 2006. "Can News about the Future Drive the Business Cycle?" Working paper 12537. Cambridge, Mass.: National Bureau of Economic Research.

———. 2007. "News and Business Cycles in Open Economies." Working paper 13444. Cambridge, Mass.: National Bureau of Economic Research.

Marfán, M. 2005. "Fiscal Policy Efficacy and Private Deficits: A Macroeconomic Approach." In *Beyond Reforms Structural Dynamic and Macroeconomic Vulnerability*, edited by J.A. Ocampo. Stanford University and World Bank.

Mertens, K. 2007. "The Role of Expectations in Sudden Stops." Working paper ECO2007/16. European University Institute, Economics Department.

Neumeyer, P. and F. Perri. 2005. "Business Cycles in Emerging Economies: The Role of Interest Rates." *Journal of Monetary Economics* 52(2): 345–80.

Pigou, A. 1926. *Industrial Fluctuations*. London: MacMillan.

Schmitt-Grohé, S. and M. Uribe. 2003. "Closing Small Open Economy Models." *Journal of International Economics* 61(1): 163–85.

Schmidt-Hebbel, K. 1988. "Consumo e inversión en Chile: 1974-82: Una interpretación real del *Boom*" In *Del auge a la crisis de 1982*, edited by F. Morandé and K. Schmidt-Hebbel. ILADES/Georgetown.

Uhlig, H. 1997. "A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily." University of Tilburg.

Uribe, M. and V. Yue. 2006. "Country Spreads and Emerging Countries: Who Drives Whom?" *Journal of International Economics* 69(1): 6–36.

Valdés, R. 2007. "Policy Responses to Sudden Stops in Capital Flows: The Case of Chile in 1998." Working paper 430. Santiago: Central Bank of Chile.

Zeira, J. 1994. "Informational Cycles." *Review of Economic Studies* 61(1): 31–44.

Monetary Policy under Uncertainty and Learning

"The global financial crisis is a powerful reminder of how uncertainty affects the effectiveness and design of demand policies.
This book offers a unique collection of contributions that throw light on how monetary authorities can best set, operate and communicate policy when their information about the economy and shocks is imperfect and learning is underway. Undoubtedly a must-read for policymakers and scholars alike as we slide further into unchartered economic territory."

Nicoletta Batini, International Monetary Fund

"Two of the most pervasive features of the practice of monetary policy are uncertainty and learning. Both policymakers and the public have imperfect knowledge about the governing structural relations of the economy, as well as the value of key macroeconomic variables at the time when decisions have to be made. Consequently, both uncertainty and learning affect the way monetary policy is conducted. This book provides important theoretical insights by leading scholars and practitioners about the robustness of traditional monetary policy schemes and the construction of new ones that explicitly take into account uncertainty and learning. This collection also offers several empirical assessments of how some dimensions of uncertainty affect the performance of monetary policy. The book is a must for both scholars and practitioners."

Vittorio Corbo, Centro de Estudios Públicos
and former Governor of the Central Bank of Chile

"This impressive and timely volume is about how economic actors and policy makers should and do characterize and cope with uncertainties about events and models. These state of the art articles provide good views of the role of the most promising approaches – Bayesian methods for hidden Markov models, adaptive models incorporating least squares learning, models of costly information processing, and even decision making with multiple priors. Recent events have made the issues discussed here only more important."

Thomas Sargent, New York University

BANCO CENTRAL DE CHILE