A Few Model-Based Answers to Monetary Policy Questions in the United States and the Euro-Area

Ricardo Reis*

Princeton University, NBER, and CEPR

October 2007

Abstract

This paper estimates a sticky-information general-equilibrium model for the United States and the Euro-area, and uses it as a laboratory to answer monetary policy questions. The first set of questions is positive and concerns describing past monetary policy: what policy rule has best described policy? What has been the role of stabilization policy? How large have policy errors been? What is the role of policy announcements? What is the result of having interest rates move gradually? The second set of questions is normative and concerns the design of optimal policy: what is the optimal Taylor rule? What is the optimal elastic price-level standard? What rule maximizes social welfare? How does parameter uncertainty affect optimal policy? The answers to these questions suggest a few lessons for applied monetary policy.

JEL classification: E30, E10

*I am grateful to Tiago Berriel for excellent research assistance. Contact: rreis@princeton.edu.

1 Introduction

In a famous passage, John Maynard Keynes aspired to the day when economists would be as useful as dentists. Robert E. Lucas Jr. (1980) in turn argued that this usefulness should 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), the computer program that Lucas asked for has taken the form of dynamic stochastic general equilibrium (DSGE) models, which are quickly growing in richness and being used in central banks.[1]

The initial versions of these models faced one problem. They implied rapid adjustment of many macroeconomic variables to shocks, while in the data, these responses tend to be gradual and delayed. In their predictions for investment, consumption, real wages, or inflation, the standard classical model lacks "stickiness" in the words of Sims (1998) and Mankiw and Reis (2006). The most popular approach to deal with this disconnect between theory and data follows the influential work of Christiano, Eichenbaum and Evans (2005) by adding many rigidities that stand in the way of adjustment: habits in consumption, adjustment costs on investment, norms in wage bargaining, and indexed sticky prices.

One promising alternative is inattention. If people face costs of acquiring, absorbing and processing information, they will optimally choose to stay inattentive for stretches of time, only sporadically updating their information (Reis, 2006a, 2006b). Because different people update their information at different dates, news dissipates gradually through the economy, leading to delayed and sluggish adjustments in the aggregate. One virtue of inattention is that the limitations to absorbing information should affect all agents in all of their actions. Therefore, inattention naturally leads to a parsimonious explanation for the pervasiveness of stickiness in the macroeconomy through the stickiness of information.

The objective of this paper is to provide a sticky-information DSGE model in which to perform the experiments envisioned by Lucas. To do so, the model must fulfill two requisites. First, it must be sufficiently rich to make predictions on the key variables that policymakers care about. Second, there must be a way to assign values to the parameters of

[1] For a few examples, these type of models are now in use at the ECB (Smets and Wouters, 2003), the Board of Governors (Erceg, Guerrieri and Gust, 2006, Edge, Kiley, Laforte, 2007), and the IMF (Bayoumi, 2004).

the model so that it is able to match the important features of the data. Given a model and parameter values, one has the “laboratory” that Lucas asked for to study the consequences and merits of different policies.

The model in this paper is the sticky information general-equilibrium (SIGE) setup developed by Mankiw and Reis (2007). Section 2 presents the model, adding to the previous work a discussion of its limitations and the derivation of a criteria for evaluating social welfare under different policies. Section 3 picks parameters by fitting the model's predictions to data using Bayesian methods. It extends previous work by expanding the number of parameters being estimated, and applying it to both U.S. and E.U. data.

The heart of this paper is in sections 4 and 5. They ask different hypothetical questions on monetary policy and answer them through the lenses of the estimated SIGE model. For each policy experiment, I compare the dynamics of macroeconomic variables under the status quo and under the proposed change and I calculate the policy's welfare benefits.

Section 4 asks questions about the current policy status quo. It starts by describing the policy rule that best describes current and past behavior. Then, it examines the role that countercyclical stabilization policy has had. Next, it asks what has been the contribution and effects of discretionary policy errors. Finally, it investigates the effect of expectations regarding policy both by asking whether disinflations should be announced ahead of time, and whether it is best to move interest rates gradually or all at once.

Section 5 turns to optimal policy. Taking the parameter estimates as fixed, it starts by computing the Taylor rule coefficients that maximize welfare and comparing them with the actual coefficients estimated in the data. It then considers an alternative simple policy, an optimal “elastic price standard” that targets a path for the price level allowing for temporary deviations in response to output gaps. Next, it computes the more involved policy rule that maximizes welfare under commitment. I compare each of these policies in terms of the social welfare benefits they bring. Finally, I ask how parameter uncertainty affects these optimal policy rules, first by assessing their expected performance over the estimated distribution of parameter values, and second by computing the policy rules that maximize this expected performance.

I picked these policy experiments both because they clarify the properties of the SIGE model and because they connect to the typical questions asked by central bankers. Insofar as the model captures important features of the data, it leads to lessons on how to best conduct monetary policy. Section 6 concludes by stating these lessons.

2 The SIGE model

The SIGE model of Mankiw and Reis (2007) belongs to the wide class of general-equilibrium models with monopolistic competition and price-setting by firms that have become the workhorse for the study of monetary policy (surveyed in Woodford, 2003a). There are three sets of markets where agents meet every period: a market for savings where households trade bonds and interest rates change to balance borrowing and lending, markets for different varieties of goods where monopolistic firms sell varieties of goods to households, and markets for labor where monopolistic households sell varieties of labor to firms. Aside from households and firms, there is a third agent, the central bank, setting monetary policy. I model the current policy status quo in terms of a Taylor rule.

To this otherwise classical framework, SIGE adds one new assumption: that consumers, workers, and firms only update their information sporadically, with a share of each randomly drawn every period being allowed to update their information. Reis (2006a, 2006b) provides micro-foundations for optimal inattentiveness by considering costs of acquiring, absorbing and processing information, and establishes some restrictive conditions under which this may imply the particular updating pattern just described. For a homoskedastic linearized economy with constant costs of planning, the optimal rate of arrival of information is fixed, so I will treat it as a parameter, but it is important to keep in mind that it can be mapped into a measure of the monetary cost of updating information.

Sub-section 2.1 presents the formal model and sub-section 2.2 discusses areas for improvement in future research. Readers less interested in the micro-foundations can skip these sections and jump straight to section 2.3 where I discuss the reduced-form equations describing the dynamics of macro variables. Section 2.4 derives a welfare criteria to evaluate different policies.

2.1 The micro-foundations

I describe the economic environment first, that is the set of agents and markets, and then solve for optimal behavior to finally define an equilibrium.

The environment: There is a continuum of households, distributed in the unit interval and indexed by j, that derive utility at each date t from consuming an aggregate $C_{t,j}$ of goods, and disutility from working $L_{t,j}$ hours according to the function:

$$U(C_{t,j}, L_{t,j}) = \ln(C_{t,j}) - \frac{\varkappa L_{t,j}^{1+1/\psi}}{1+1/\psi}. \tag{1}$$

Households live forever and discount the future by the factor ξ, so individual welfare is $E_0 \sum_{t=0}^{\infty} \xi^t U(C_{t,j}, L_{t,j})$. The parameter ψ measures the Frisch elasticity of labor supply, and the intertemporal elasticity of substitution is one so that the model is consistent with a balanced growth path with bounded hours. The aggregate good is a Dixit-Stiglitz composite of the consumption of a continuum of varieties of goods in the unit interval, indexed by i, with a time-varying and random elasticity of substitution $\hat{\nu}_t$:

$$C_{t,j} = \left(\int_0^1 C_{t,j}(i)^{\frac{\hat{\nu}_t}{\hat{\nu}_t - 1}} di \right)^{\frac{\hat{\nu}_t - 1}{\hat{\nu}_t}}. \tag{2}$$

Each good trades at price $P_{t,i}$, and the Dixit-Stiglitz form implies that there is a static price index:

$$P_t = \left(\int_0^1 {P_{t,i}}^{1-\hat{\nu}_t} di \right)^{\frac{1}{1-\hat{\nu}_t}}, \tag{3}$$

such that, conditional on the optimal choices of the consumer, $\int_0^1 P_{t,i} C_{t,j}(i) = P_t C_{t,j}$. The household faces a budget constraint every period:

$$M_{t+1,j} = \Pi_{t+1} \left(M_{t,j} - C_{t,j} + (1-\tau_w) W_{t,j} L_{t,j} / P_t + T_{t,j} \right) \tag{4}$$

The household enters a 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 the transfer $T_{t,j}$. The remaining funds accumulate at the real interest rate Π_{t+1} by participating in a financial market where 1-period bonds are traded. 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 the beginning of time that ensures that every period they are all left with the same wealth.

For each good i, there is a technology that produces it using as inputs the labor from each household with a constant elasticity of substitution among them and overall decreasing returns to scale:

$$Y_{t,i} = A_t N_{t,i}^{\beta}, \tag{5}$$

$$N_{t,i} = \left(\int_0^1 N_{t,i}(j)^{\frac{\hat{\gamma}_t}{\hat{\gamma}_t - 1}} dj \right)^{\frac{\hat{\gamma}_t - 1}{\hat{\gamma}_t}}. \tag{6}$$

The parameter β determines the degree of returns to scale, while $\hat{\gamma}_t$ is the random elasticity of substitution across labor varieties. $N_{t,i}$ is the labor aggregate used in the production of good i, that combines in a Dixit-Stiglitz way the labor from each household. The corresponding price index across labor varieties such that $W_t N_{t,i} = \int_0^1 W_{t,j} N_{t,i}(j) dj$ is:

$$W_t = \left(\int_0^1 W_t(i)^{1-\hat{\gamma}_t} di \right)^{\frac{1}{1-\hat{\gamma}_t}}. \tag{7}$$

The market-clearing conditions for each good and labor variety are:

$$G_t \int_0^1 C_{t,j}(i) dj = Y_{t,i}, \tag{8}$$

$$\int_0^1 N_{t,i}(j) di = L_{t,j}. \tag{9}$$

The time-varying and random G_t reflects changes in government spending, which here lead to wasted resources. I refer to them broadly as aggregate demand shocks. The fraction of output of each variety consumed by the government is $1 - 1/G_t$. I define aggregate output and labor as Dixit-Stiglitz aggregators across varieties

$$Y_t = \left(\int_0^1 Y_{t,i}^{\frac{\hat{\nu}_t - 1}{\hat{\nu}_t}} di \right)^{\frac{\hat{\nu}_t}{\hat{\nu}_t - 1}}, \tag{10}$$

$$L_t = \left(\int_0^1 L_{t,j}^{\frac{\hat{\gamma}_t - 1}{\hat{\gamma}_t}} dj \right)^{\frac{\hat{\gamma}_t}{\hat{\gamma}_t - 1}}. \tag{11}$$

although note that using instead the definitions $Y_t = \int Y_{t,i} di$ and $L_t = \int L_{t,j} dj$ leads to the same results up to a first-order approximation.

To describe the data, I assume that policy followed a Taylor rule:

$$i_t = \phi_y \log \left(\frac{Y_t}{Y_t^n} \right) + \phi_p \log \left(\frac{P_t}{P_{t-1}} \right) - \varepsilon_t, \tag{12}$$

where $i_t \equiv \log\left[E_t\left(\Pi_{t+1}P_{t+1}/P_t\right)\right]$ and ε_t is a policy shock. The level of output, Y_t^n, is that which would be realized if all agents were attentive.

Information and optimal behavior: I assume that consumers are separated into two decision-makers that do not communicate with each other: a shopper and a planner. The shopper's job is to allocate total spending on consumption by household j to the continuum of varieties. He has full information, so he solves the following problem at each data:

$$\min_{\{C_{t,j}(i)\}_{i\in[0,1]}} \int_0^1 P_{t,i}C_{t,j}(i)di, \tag{13}$$

subject to the labor aggregator (6). The standard solution is:

$$C_{t,j}(i) = C_{t,j}\left(P_{t,i}/P_t\right)^{-\hat{\nu}_t}. \tag{14}$$

Summing over all consumers and using the market clearing condition gives the total demand for variety i:

$$Y_{t,i} = \left(P_{t,i}/P_t\right)^{-\hat{\nu}_t} G_t \int_0^1 C_{t,j}dj. \tag{15}$$

The planner's job is to decide on total spending each period, $C_{t,j}$. He is inattentive, and δ is the probability that he updates his information each date. Therefore, there are δ agents who have current information, $\delta(1-\delta)$ that have 1-period old information, $\delta(1-\delta)^2$ with 2-period old information, and so on. Since agents that last updated their information at the same time are identical in everything, we can group planners according to when they last updated. The subscript j then denotes how long ago did the planner last update and there are $\delta(1-\delta)^j$ many agents in this group.

When the planner updates his plan at some date t, he chooses a plan for current and future consumption all the way into infinity $\{C_{t+l,l}\}_{k=0}^{\infty}$ since with a vanishingly small probability he may never update again. Letting $V(M_t)$ denote the value function of an agent that has just update his information, his problem is:

$$V(M_t) = \max_{\{C_{t+l,l}\}} \left\{ \sum_{l=0}^{\infty} \xi^l (1-\delta)^l \ln(C_{t+l,l}) + \xi\delta \sum_{l=0}^{\infty} \xi^l (1-\delta)^l E_t\left[V\left(M_{t+1+l}\right)\right] \right\} \tag{16}$$

subject to the sequence of budget constraints in (4), a no-Ponzi scheme condition, and the demand for labor variety j that this worker solely supplies. The optimality conditions are:

$$\frac{\xi^l(1-\delta)^l}{C_{t+l,l}} = \xi\delta\sum_{k=l}^{\infty}\xi^k(1-\delta)^k E_t\left[V'\left(M_{t+1+k}\right)\bar{\Pi}_{t+i,t+1+k}\right] \tag{17}$$

$$V'\left(M_t\right) = \xi\delta\sum_{l=0}^{\infty}\xi^l(1-\delta)^l E_t\left[V'\left(M_{t+1+l}\right)\bar{\Pi}_{t,t+1+l}\right], \tag{18}$$

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 (17) is the same as the right-hand side of (18). Therefore, $1/C_{t,0} = V'\left(M_t\right)$, 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 for the $V'\left(M_{t+1+l}\right)$ terms in (18) and writing the equation recursively gives the standard Euler equations linking the marginal utility of consumption today and tomorrow for an agent that updates her information at both dates. Equation (18) for $t+l$ and (17) imply that inattentive consumers set their marginal utility equal to the expected marginal utility of attentive consumers. These two relations are:

$$C_{t,0}^{-1/\theta} = \xi E_t\left[\Pi_{t+1}C_{t+1,0}^{-1/\theta}\right], \tag{19}$$

$$C_{t+l,l}^{-1/\theta} = E_{t-l}\left[C_{t+l,0}^{-1/\theta}\right]. \tag{20}$$

Next, I turn to firms. I assume that they are divided in two departments, one that purchases the cost-minimizing mix of inputs, and another that picks the overall amount of production. The purchasing department has full information, so for firm i at date t it minimizes costs by solving:

$$\min_{\{N_{t,i}(j)\}_{j\in[0,1]}}\int_0^1 W_{t,j}N_{t,i}(j)dj \tag{21}$$

subject to the labor aggregator (6). The solution to this problem is:

$$N_{t,i}(j) = N_{t,i}\left(W_{t,j}/W_t\right), \tag{22}$$

which, aggregating over all firms and using the labor market clearing condition, gives the total demand for labor variety j:

$$L_{t,j} = (W_{t,j}/W_t)^{-\hat{\gamma}_t} \int N_{t,i} di. \tag{23}$$

The sales department maximizes profits and is inattentive. Every period, a randomly drawn fraction of firms λ updates their information, so firms can be grouped into groups i of size $\lambda(1-\lambda)^i$ according to how long it has been since they last updated. Each firm is a monopoly provider of its good and chooses a nominal price at which it stands ready to satisfy demand. There is a fixed sales tax τ_p. The firm who last updated i periods ago sets $P_{t,i}$ to maximize expected real after-tax profits:

$$\max_{P_{t,i}} E_{t-i} \left[\frac{(1-\tau_p)P_{t,i}Y_{t,i}}{P_t} - \frac{W_t N_{t,i}}{P_t} \right] \tag{24}$$

subject to the available technology (5) and taking into account the demand for its good in (15). The first-order condition, after some rearranging, is:

$$P_{t,i} = \frac{E_{t-i}\left[(1-\tau_p)\hat{\nu}_t W_t N_{t,i}/P_t\right]}{E_{t-i}\left[\beta(\hat{\nu}_t - 1)Y_{t,i}/P_t\right]}. \tag{25}$$

Finally, I turn to workers. They are inattentive, and each period a randomly drawn sample updates their information with probability ω. Their problem is similar to that of consumption planners and consists of:

$$\hat{V}(M_t) = \max_{\{W_{t+l,l}\}} \left\{ -\sum_{l=0}^{\infty} \xi^l (1-\omega)^l \varkappa E_t \left(\frac{L_{t+l,l}^{1+1/\psi} + 1}{1+1/\psi} \right) + \xi\omega \sum_{l=0}^{\infty} \xi^l (1-\omega)^l E_t \left[\hat{V}\left(M_{t+1+l}\right) \right] \right\}, \tag{26}$$

subject to the sequence of budget constraints in (4), a no-Ponzi scheme condition, and the demand for the variety of labor j (23) which this worker supplies monopolistically. The optimality conditions are:

$$\xi^l (1-\omega)^l \varkappa E_t \left(\hat{\gamma}_{t+l} L_{t+l,l}^{1+1/\psi} \right) (1-\tau_w)/W_{t+l,l} =$$

$$\xi\omega \sum_{k=l}^{\infty} \xi^k (1-\omega)^k E_t \left[V'\left(M_{t+1+k}\right) \bar{\Pi}_{t+l,t+1+k} \left(\hat{\gamma}_{t+l} - 1\right) L_{t+l,l}/P_{t+l} \right] \tag{27}$$

$$\hat{V}'\left(M_t\right) = \xi\omega \sum_{k=0}^{\infty} \xi^k (1-\omega)^k E_t \left[\hat{V}'\left(A_{t+1+k}\right) \bar{\Pi}_{t,t+1+k} \right]. \tag{28}$$

Now, as in the consumer problem, combining (27) for $l = 0$ with (28) leads to the conclusion:

$$\frac{\hat{V}_t'(M_t)W_{t,0}}{P_t} = \frac{(1-\tau_w)\hat{\gamma}_t \varkappa L_{t,0}^{1/\psi}}{\hat{\gamma}_t - 1}. \tag{29}$$

This expression shows that ψ is the Frisch elasticity of labor supply for attentive agents, and that the marginal disutility of working os 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. Similar manipulations as in the consumer problem lead to the two following condition:

$$\frac{\hat{\gamma}_t}{\hat{\gamma}_t - 1} \times \frac{L_{t,0}^{1/\psi} P_t}{W_{t,0}} = \xi E_t \left(\Pi_{t+1} \times \frac{\hat{\gamma}_{t+1}}{\hat{\gamma}_{t+1} - 1} \times \frac{L_{t+1,0}^{1/\psi} P_{t+1}}{W_{t+1,0}} \right), \tag{30}$$

$$W_{t+l,l} = \frac{E_t\left((1-\tau_w)\varkappa \hat{\gamma}_{t+l} L_{t+l,l}^{1/\psi}\right)}{E_t\left(\hat{\gamma}_{t+l} L_{t+l,l} L_{t+l,0}^{1/\psi - 1}/W_{t+l,0}\right)}. \tag{31}$$

Monopolistically competitive equilibrium. An equilibrium of this economy is a set of aggregate variables $\{Y_t, L_t\}$, output of each variety $\{Y_{t,i}\}$, labor of each variety $\{L_{t,j}\}$, prices of each good $\{Y_{t,i}\}$, wages $\{W_{t,i}\}$, and interest rates $\{i_t\}$, such that consumers behave optimally (15), (19), (20), firms behave optimally (23), (25), workers behave optimally (30), (31), markets clear, the aggregates for output, labor, prices and wages satisfy (10), (11), (3), (7) and monetary policy follows the Taylor rule (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 $\{\hat{\nu}_t\}$, and labor substitutability $\{\hat{\gamma}_t\}$.

2.2 Missing work on the micro-foundations

The SIGE model that I just presented provides a DSGE account of 5 macroeconomic series: output, inflation, real wages, hours, and nominal interest rates. In the tradition of Kydland and Prescott (1982) and Rotemberg and Woodford (1997), the model makes a few simplifying assumptions, some more common and other perhaps more unusual. Each of these presents an opportunity for future work that can improve the model. I now discuss a few that seem particularly promising.

First, the model lacks investment and capital accumulation. It is an issue of open debate whether this absence significantly affects the dynamics of the other variables in this class of models (Woodford, 2005, Sveen and Weinke, 2005), but a clear benefit from modelling investment is that it extends the model to explain one more macroeconomic variable. The reason why the SIGE model omits investment is that while there is previous work that studies in detail the micro-foundations and implications of inattentiveness on the part of consumers (Reis, 2006a), firms setting prices (Mankiw and Reis, 2002, Reis, 2006b) and workers (Mankiw and Reis 2003), the behavior of inattentive investors accumulating capital has not been studied yet. There has been some related work studying financial investment decisions with inattentiveness by Gabaix and Laibson (2002) and Abel, Eberly and Panageas (2007), but the step from this to study physical investment and capital accumulation remains to be taken.

The model also lacks international trade and exchange rates. The reason for this omission is the same as that for investment: the models of inattentive behavior in international markets are still missing. Also in this case, progress will likely come soon, as Bachetta and van Wincoop (2006) have already filled some of this gap. Once this is done, one can build an open economy SIGE to use in economies other than the U.S. 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 done.

Turning to the micro-foundations of inattentiveness, note that it assumes that when agents pays 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 be easily relaxed to allow people to observe only some things but not everything when they update (see e.g., Carroll and Slacalek, 2007). 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. There has been some recent work in this area, following the approach of Sims (2003), but these models are still not at the point where they can be put in general equilibrium and taken to the data, despite promising progress by Mackowiack and Wiederholt (2007).

One implication of removing the assumption that once agents learn, they 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 (2003b), followed by Hellwig (2002), Amato and Shin (2003), Morris and Shin (2006) and Adam (2007) have studied some of the implications of this behavior, and recent work by Lorenzoni (2006) moves towards turning these insights into a business cycle model that could be taken to the data. Hellwig and Veldkamp (2007) study another source of strategic interaction, on whether agents coordinate their attention times. These extra ingredients promise to enrich future models of inattentiveness.

There is another source of strategic interaction ignored by the SIGE model. The model assumed that consumers had inattentive planners and attentive shoppers, and firms had inattentive sales departments and attentive purchasing departments, so that at every market, monopolists would only face attentive agents. This is important because if a monopolist sold its product to some buyers that are inattentive, then it would 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 yet.

Overall, there are many features ignored by the SIGE model that can lead to new and interesting insights. They were omitted typically because they are not sufficiently well-understood to put them into the full DSGE setup that this paper wishes to deliver. Some of them will be easier to do and others more challenging, and I hope that future research will take them up.

2.3 The key reduced-form relations

The appendix shows how to derive a set of log-linear approximate relations for the aggregate variables in the model. The first relation is a Phillips curve or aggregate supply curve relating the price level (p_t) to marginal costs that depend on real wages ($w_t - p_t$), output (y_t) and aggregate productivity (a_t), and to shocks to desired markups due to changes in the elasticity of substitution for goods (ν_t):

$$p_t = \lambda \sum_{j=0}^{\infty} (1-\lambda)^j E_{t-j} \left[p_t + \frac{\beta(w_t - p_t) + (1-\beta)y_t - a_t}{\beta + \nu(1-\beta)} - \frac{\beta \nu_t}{(\nu - 1)[\beta + \nu(1-\beta)]} \right] \quad (32)$$

The parameter β measures the degree of returns to scale in production, while ν is the steady-state elasticity of substitution for goods.[2] Prices depend on past expectations of the marginal cost plus the desired markup because every period only a fraction λ of firms update their information and set their plans for current and future prices until the next adjustment.

The second relation is an aggregate demand curve or IS curve relating output to a measure of wealth ($y_\infty^n = \lim_{i\to\infty} E_t\left(y_{t+i}\right)$), the long real interest rate ($R_t = E_t \sum_{j=0}^{\infty}\left(i_{t+j} - \Delta p_{t+1+j}\right)$) and shocks to government spending (g_t):

$$y_t = \delta \sum_{j=0}^{\infty} (1-\delta)^j E_{t-j}\left(y_\infty^n - R_t\right) + g_t, \tag{33}$$

Higher expected future output raises wealth and increases spending, while higher expected interest rates encourage savings and lower spending. Every period only a randomly drawn share δ of consumers update their plan, so the larger is δ, the more consumption respond to shocks as they occur.

Next comes a wage or labor supply curve that relates current wages (w_t) to the expected value of five determinants:

$$w_t = \omega \sum_{j=0}^{\infty} (1-\omega)^j E_{t-j}\left[p_t + \frac{\gamma(w_t - p_t)}{\gamma+\psi} + \frac{l_t}{\gamma+\psi} + \frac{\psi\left(y_\infty^n - R_t\right)}{\gamma+\psi} - \frac{\psi\gamma_t}{(\gamma+\psi)(\gamma-1)}\right] \tag{34}$$

First, nominal wages rise one-to-one with prices since workers care about real wages. Second, the higher are real wages in the economy, the higher is demand for a worker's variety of labor so the higher the wage she will demand. Third, the more labor is hired (l_t) the better it must be compensated since the marginal disutility of working rises. Fourth, higher wealth discourages work through an income effect, and higher interest rates promote it by giving a larger return on saved earnings today. Fifth and finally, if the elasticity of substitution across labor varieties (γ_t) rises, workers' desired markup falls so they lower their wage demands. The two new parameters are ψ, the Frisch elasticity of labor supply, and γ the steady-state elasticity of substitution between different varieties of labor. A randomly drawn fraction of workers ω becomes informed every period and respond to shocks to these five determinants,

[2] All variables with a t subscript refer to log-linearized values around their non-stochastic steady state. Without any subscript are fixed parameters and steady state values.

while the remaining workers set wages to what they expected would be optimal when they last updated.

The fourth relation is a standard production function linking output to technology and labor with decreasing returns to scale:

$$y_t = a_t + \beta l_t, \tag{35}$$

The fifth and final relation is the Taylor rule:

$$i_t = \phi_y(y_t - y_t^n) + \phi_p \Delta p_t - \varepsilon_t, \tag{36}$$

where $y_t - y_t^n$ is the output gap, or the difference between actual output and its level if all agents were attentive, and ε_t are policy disturbances.

These 5 equations give the equilibrium values for output, wages, prices, labor, and nominal interest rates as a function of shocks to aggregate productivity growth, aggregate demand, goods markups, labor markups, and monetary policy. We assume that each of these shocks follows an autoregressive process of order 1 with coefficients $\rho_{\Delta a}$, ρ_g, ρ_ν, ρ_γ, and ρ_ε, and is subject to innovations $e_t^{\Delta a}$, e_t^g, e_t^ν, e_t^γ, and e_t^ε, that are independent and normally distributed with standard deviations $\sigma_{\Delta a}$, σ_g, σ_ν, σ_γ, and σ_ε. The model has a total of 20 parameters, that I collect in the vector $\boldsymbol{\theta}$, and is solved using the algorithm of Mankiw and Reis (2007).

2.4 Social welfare

To compare the merits of different policies, I use a measure of social welfare that sums over the utility of all the households. 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 this sum. Because I will compare different policy rules, I take an ex ante perspective, looking at the unconditional expectation of social welfare.

I denote social welfare as a function of the parameters by $\mathcal{W}(\boldsymbol{\theta})$. The appendix provides a formula for computing this function under the current policy rule and for calculating the benefits of alternative policies in units of steady-state consumption.

3 Estimating sticky information

Taking sticky information models to the data has been an active field of research. One approach has looked for direct evidence of inattentiveness using micro data. Carroll (2003) used surveys of inflation expectations to show that the public's forecasts lag the forecasts made by professionals.[3] Mankiw, Reis and Wolfers (2003) followed by showing that the disagreement in the inflation expectations in the survey data have properties consistent with sticky information.[4] Reis (2006a) and Carroll and Slacalek (2006) interpret some of the literature on the sensitivity and smoothness of micro consumption data in the light of sticky information and Klenow and Willis (2006) and Knotek (2006) find slow dissemination of information in the micro data on prices. For the most part, this literature has supported the sticky information assumption, and has obtained consistent estimates of the information-updating rates.

A second approach estimates Phillips curves assuming sticky information on the part of price setters only.[5] These limited information approaches typically use data on inflation, output, marginal costs and expectations to estimate simpler versions of (32) with typically good or mixed results. 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 instead 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. A few papers had attempted this before and typically found either mixed or poor fits between model and data.[6] Mankiw and Reis (2006) explained the contrast between the negative results in some of these papers and the mostly positive results found by the other two approaches. These papers assumed inattentiveness only in price-setting, while assuming that the other agents in the model were fully attentive. However, as discussed in the introduction, to fit the data stickiness should be pervasive, and for the internal coherence of the model, inatten-

[3]See also Dopke et al (2006a) and Nunes (2006).

[4]Also focussing on disagreement, see Branch (2007), Rich and Tracy (2006), and Gorodnichenko (2006).

[5]See Khan and Zhu (2006), Dopke et al (2006b), Korenok (2005), Pickering (2004), Coibion (2006), and Molinari (2007).

[6]See Trabandt (2003), Andres et al (2005), Kiley (2007), Laforte (2007), Korenok and Swanson (2006, 2007), and Paustiam and Pytlarczyz (2006).

tiveness should apply to all decisions. By assuming attentive consumer and workers, the general-equilibrium restrictions imposed in these papers were misspecified.

This paper follows this third approach, estimating sticky information in general equilibrium, but allowing for pervasive stickiness. Given a set of data $\mathbf{Y}$, this paper takes a Bayesian likelihood approach, starting with a prior density $p(\boldsymbol{\theta})$ and using the likelihood function $\mathcal{L}(\mathbf{Y} \mid \boldsymbol{\theta})$ to obtain the posterior density of the parameters $p(\boldsymbol{\theta} \mid \mathbf{Y})$. This is done numerically, using Markov Chain Monte Carlo simulations.

I set the value of five parameters: the Frisch elasticity of labor supply ψ to 4, the degree of returns to scale in labor β to 2/3, and because I observe hours and output in the data, I can back out a series for the technology shock and use least-squares to attribute a value to $\rho_{\Delta a}$ and $\sigma_{\Delta a}$.[7] All the other 15 parameters are estimated using the algorithms developed by Mankiw and Reis (2007)[8]. The prior distributions follow the convention in the DSGE literature (An and Schorfheide, 2006) for most parameters, while for the inattentiveness I use a flat prior in the unit interval. They are described in Table 1.

3.1 U.S. estimates of the SIGE model

I use U.S. quarterly data on real output per capita, total real compensation per hour, hours per capita, inflation, and the effective federal funds rate. The first four series refer to the non-farm business sector, and the price series used is the implicit price deflator for this sector. All the series are de-meaned and go from 1954:3 to 2006:1.

Table 2 displays summary statistics of the posterior distributions for each parameter. The estimates of the non-policy parameters are roughly similar to those in Mankiw and Reis (2007). The median elasticity of substitution across goods varieties implies a steady-state markup of 5%, while across labor varieties, markups are higher at 16%. All of the 5 exogenous shocks are quite persistent, and for most estimates, the 95% credible sets are considerably tighter than the priors so the data were informative. Consumers and workers update their information at almost the same frequency, between 5 and 6 quarters, while firms are considerably more attentive, once every 1.5 quarters.

Table 3 displays variance decompositions. Inflation is mostly driven by monetary shocks,

[7] For the United States, $\rho_{\Delta a} = .35$ and $\sigma_{\Delta a} = .01$, while for the Euro-area $\rho_{\Delta a} = .22$ and $\sigma_{\Delta a} = .005$.

[8] Different from Mankiw and Reis (2007), I estimate the parameters describing monetary policy: ϕ_y, ϕ_p, ρ_ε, and σ_ε, which they fixed using prior information.

while 81% of the variance of hours is due to monetary and aggregate demand shocks. For output growth, aggregate demand shocks account for almost half of its variance, with monetary, productivity and goods-markup shocks all accounting for roughly the same share of the remainder.

Figure 1 shows one-standard deviation impulse responses to the four non-policy shocks. Notably, all four shocks lead to hump-shaped dynamics for inflation. As in Gali (1999), a positive innovation to productivity lowers hours, and while it raises output, it does so by less than full-information output, leading to a negative output gap. The SIGE model can generate booms or recessions following a productivity shock, but the U.S. data seems to prefer parameter combinations where there is a recession. Increases in the elasticity of substitution across goods or labor varieties raise hours in the short run, but actually lead to a negative output gap because they immediately raise full-information output by more than actual output.

Table 4 compares a few relevant sample moments with the estimated model's predictions. The first two are the standard deviations of inflation and output growth, since these are typically used to assess policy trade-offs. In the data they are both .01, but at the prior parameters, their value was much higher. The posterior predictions from the model are significantly closer to the data. The third moment is the serial correlation of inflation, which has attracted considerable attention, both theoretically and empirically.[9] In the data, it is .83, which is within the posterior 95% credible set. The last three moments capture three puzzling observations in the macro data: the positive Phillips correlation that models with no role for monetary policy have trouble matching, the smoothness of real wages relative to labor productivity that models with frictionless labor markets struggle with, and the gradual response of real variables to shocks that classical models miss.[10] With full attention, the model in this paper would predict that these moments would be 0, 1 and 1.22. Inattention moves the model in the right direction for matching the facts on the Phillips correlation and the gradual response of output, although at the posterior estimates, it has little effect over the smoothness of real wages.

[9] See Fuhrer and Moore (1995) and Pivetta and Reis (2007).

[10] Mankiw and Reis (2006) discuss the central role of these facts.

3.2 Euro-area estimates of the SIGE model

For the Euro-area, I use the area-wide quarterly dataset that combines data from each country's national accounts to build consistent pseudo-aggregates for the whole region from 1970:1 to 2005:4. Inflation is the change in the log of the GDP deflator, output growth 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.

I restrict attention to the period from 1993:4 to 2005:4. On November 1st 1993, the Maastricht treaty was signed, creating the European Union and setting out the plan to the introduction of a single currency in 1999. The Euro-area data from this date forward is closer to the model's assumption of a common and stable monetary policy rule. The appendix describes the estimation results starting the sample instead at 1979:3 (when the European Monetary System was first created).

Table 5 shows the posterior distributions of the parameters for the Euro-area. Note that, in spite of the short sample, the posterior mean estimates are considerably different from their prior counterparts, and that the credible sets are fairly tight. The data were clearly informative. Turning to the estimates, the elasticities of substitution in both the goods and labor market are similar, implying desired markups of 9% and 7% respectively, in contrast with the United States where goods market are more "competitive" in the sense of significantly lower desired price markups. For the non-policy shocks, it stands out that aggregate demand shocks are very small.

The estimates of inattentiveness are more surprising and interesting. According to them, consumers are very inattentive, not updating their information for 3.5 years on average. Workers, in contrast, are quite well-informed, updating information every 4.5 months. In part, this may be due to the stronger role of unions in Europe than the United States. These may imply that, on the one hand, unions are constantly monitoring information and bargaining ensuring wages reflect current information, and on the other hand they reduce the incentive for each household to collect information for choosing consumption. As for firms, they are more inattentive than their U.S. counterparts, updating their price plans on average every 3 quarters.

Table 6 has the variance decompositions. While they are imprecisely estimated, one can detect some broad patterns that are different from the ones for the United States.

Inflation is mostly driven by productivity shocks, while output growth and hours are driven by aggregate demand shocks. Monetary shocks have a smaller role than in the United States, although they account for almost half of the variability of real wage growth, and for some parameters in the 95% set, they can account for between one third and one half of the variability of inflation, output growth, and hours. Shocks to the markups are essentially irrelevant in the Euro-area.

Figure 2 shows the impulse responses to the different shocks. The productivity and aggregate demand shocks have very persistent effects on all variables, with inflation peaking several years after the initial shock. Curiously, a positive aggregate-demand shock raises hours and the output gap, but induces such a strong raise in nominal interest rates that it ends up raising inflation. As expected from the variance decompositions, the markup shocks have a negligible impact.

Table 7 has the same predictive moments as table 4. As it was the case there, the posterior predictions for standard deviations are significantly closer to the data than their prior counterparts. The model predicts considerably more inflation persistence than in the data, however. The Phillips correlation is weaker in the E.U. data, and the posterior estimates reflect this, although it is estimated imprecisely. The main discrepancy between the model's predictions and the data is in the variance of real wages relative to labor productivity. As a result of having quite attentive workers, the posterior estimates predict very volatile real wages.

4 Monetary policy positive questions

Using the two estimated models, I now ask and answer a set of hypothetical questions about monetary policy.

4.1 What rule has best described policy?

The estimation assumed that policy followed a Taylor rule for nominal interest rates. An extensive literature, starting with Taylor (1993) has documented that this provides a good description of policy in the United States, and a not-too-bad 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. The mean and standard-error of the posterior parameter distributions are:

$$i_t = \underset{(.20)}{1.28}\Delta p_t + \underset{(.10)}{0.43}(y_t - y_t^n) - \varepsilon_t \text{ and } \varepsilon_t = \underset{(.02)}{0.90}\varepsilon_{t-1} + \underset{(.002)}{0.01}z_t \text{ for the United States,}$$

$$i_t = \underset{(.16)}{1.23}\Delta p_t + \underset{(.37)}{1.49}(y_t - y_t^n) - \varepsilon_t \text{ and } \varepsilon_t = \underset{(.05)}{0.79}\varepsilon_{t-1} + \underset{(.01)}{0.02}z_t \text{ for the Euro-area.}$$

According to these estimates, monetary policy in the United States responds to fluctuations in the output gap and is very persistent. In contrast, the estimated Taylor rule for the Euro-area involves a response to output that is more than three times as large as the one for the United States, as well as less persistent shocks.[11]

Figure 3 plots the impulse responses to a 1-standard-deviation monetary-policy shock in the United States. Lowering nominal interest rates leads to a persistent increase in inflation, peaking 1 year after the shock. An expansion ensues, with hours and the output gap peaking before inflation, 3 quarters after the shock. This description matches relatively well the conventional wisdom on the impact of monetary policy shocks. Figure 4 plots the impulse responses for Europe. These are remarkably similar to the U.S. ones. The only slight difference is that the impulse responses die a little faster in the Euro-area, which is probably tied to the less persistent monetary shocks.

4.2 What has been the role of stabilization policy?

I interpret stabilization policy as setting nominal interest rates that respond to movements in the output gap, that is $\phi_y > 0$. The 95% posterior credible sets for the coefficient on output in the Taylor rule are [.26, .64] and [.90, 2.36] respectively for the United States. and the Euro-area. clearly, policy has reacted to the output gap in both areas, and significantly more so in the Euro-area.

Figures 5, 6, and 7 plot the impulse responses to shocks to monetary policy, aggregate demand and productivity shocks if instead U.S. nominal interest rates had not responded to the output gap. Had that been the case, the response of hours, the output gap, and also inflation to shocks would have been significantly more pronounced. Next, I calculate the variance of the macroeconomic variables if ϕ_y had been zero. The variances of inflation,

[11] These results are in contrast with those from the DSGE model with sticky prices of Christiano, Motto and Rostagno (2007). They find that nominal interest rates in the Euro-area are more persistent than in the United States, and equally responsive to the output gap. One key difference is that they specify the policy rule differently, assuming partial adjustment in interest rates rather than persistent ε_t shocks, and including money as an extra argument in the policy rule for the Euro-area.

hours and the output gap would have been 138%, 116%, and 106% higher respectively. Only nominal interest would have been more stable, by 27%. As a result, stabilization policy has a clearly beneficial role in the United States: welfare would have been 12% lower in units of steady-state consumption without it.

In the Euro-area, depicted in figures 8, 9 and 10, we again observe that hours and the output gap would respond more to shocks if there were no stabilization concerns. However, both nominal interest rates and now also inflation are significantly less responsive to shocks in this counterfactual scenario. Looking at variances, without stabilization policy, hours and the output gap would have been 62% and 228% more volatile respectively, whereas inflation and interest rates would have been 55% and 49% less volatile respectively. The welfare loss of eliminating stabilization policy is 4 times smaller than in the United States at 3%.

Interestingly, while ϕ_y was higher in the Euro-area, these results show that stabilization policy has been more important in the United States at curtailing volatility and improving welfare.

4.3 How large have policy errors been?

Changes in interest rates not captured by changes in inflation or in output are attributed to ε_t. In the data, these ε_t shocks may come from many sources, from reaction to other macroeconomic and financial news, to the exercise of judgement and discretion, or just to policy mistakes (Blinder and Reis, 2006). In the model, these shocks are treated exclusively as policy errors. They bring no benefits and unambiguously lower welfare by introducing an extra source of variability.

According to the variance decomposition in Tables 3 and 6, these shocks have accounted for a large share of the variance of several macro variables in the United States, but a smaller amount in the Euro-area. If they had been eliminated, the variance of inflation would have fallen by a striking 74% in the United States, while hours and the output gap would be 46% and 26% as volatile as before respectively. In terms of welfare, eliminating these policy shocks would have raised U.S. well-being by an equivalent amount to 5% of steady-state consumption. In the Euro-area, the reduction in the variance is more modest, 10%, 8% and 35% for inflation, hours and the output gap, and the welfare benefit is smaller, 1.4%.

4.4 What is the role of policy announcements?

In the past decade, there has been an increasing emphasis on transparency in central banking. Part of the argument for transparency is that if the central bank acts in a predictable way it will reduce confusion and mistakes on the part of private decision-makers. If policy shocks must take place, then this point of view argues that they should be well-announced and 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 in the interim between announcement and action, a larger fraction of agents have time to become aware.

Figures 11 and 12 show the results from announcing a monetary policy shocks 1 or 2 years ahead, in the United States and the Euro-area respectively. In both regions, announcements lower the total impact of monetary policy shocks on hours and the output gap, while increasing their impact on inflation. The agents that update their information learn about the shocks before it happens and adjust their actions in response. Inflation and nominal interest rates rise immediately, even before the policy shock materializes.

4.5 What is the result of having interest rates move gradually?

As described by Bernanke (2004), the FOMC tends to change interest rates gradually. There have been some academic arguments in favor of such actions, typically involving financial stability, the gradual revelation of news, or the desire to move long-term interest rates. Woodford (2000) noted that in forward-looking models like SIGE, gradualism involves combining policy responses with announcements of future policy changes.

Figures 13 and 14 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 is a sequence of four consecutive shocks, each of size $\sigma_\varepsilon/4$ and each coming as a surprise to the agents. Finally, in the third scenario, the sequence of four shocks is accounted at date 0. Gradualist policy seems to have a slightly larger impact on macro variables than unexpected policy changes. This effect is significantly larger if it is announced.

5 Monetary policy normative questions

This section uses social welfare function from the model to compute optimal policy rules under different regimes.

5.1 What is the optimal Taylor rule?

At the mean posterior estimates, the optimal Taylor rule's are (in parenthesis are the estimated status quo coefficients):

$$i_t = \underset{(1.28)}{1.62}\Delta p_t + \underset{(0.43)}{1.81}(y_t - y_t^n) \text{ for the United States,}$$

$$i_t = \underset{(1.23)}{19.00}\Delta p_t + \underset{(1.49)}{0.54}(y_t - y_t^n) \text{ for the Euro-area.}$$

The optimal response of nominal interest rates to inflation is higher at the optimal rules than in the status quo. In the Euro-area, it is extreme; however, the welfare function is very insensitive to values of ϕ_π above 3, so the exact value of 19 need not be taken too seriously. Looking at the response of nominal interest rates to the output gap, it is 3 times lower at the optimum than in the status quo in the Euro-area, but it would have to increase by more than 4 times in the United States.

Implementing this optimal Taylor rule in place of the estimated rule would raise welfare in the United States by 5.5% of steady-state consumption. Recalling from section 4 that eliminating the policy errors led to a welfare gain of 5%, this implies that adjusting the parameters of the Taylor rule to their optimal levels contributes 0.5% to welfare. The welfare gains for Europe are more modest in total, 2%, but higher in terms of the benefit from adjusting the coefficients in the Taylor rule, 0.6%.

5.2 What is the optimal elastic price-level standard?

Ball, Mankiw and Reis (2005) showed that in an economy with inattentive firms, the optimal policy was an "elastic price standard" as described in Hall (1984). It keeps the price level close to a deterministic target K_t, allowing for deviations of the price level from it in response to output gaps:

$$p_t = K_t - \phi(y_t - y_t^n) \tag{37}$$

With inattentiveness on the part of consumers and workers as well, there is no reason to expect that this simple policy continues to be the best. I computed the optimal ϕ if such a rule was implemented in the United States and the Euro-area. In the United States, the optimal ϕ is 0.95, so there is quite a bit of elasticity of the price path relative to the target, while in the Euro-area, it is close to zero (0.01), so that policy resembles a strict price-level target. This type of policy would raise welfare in the United States by 2.6%, less than half of the benefits from an optimal Taylor rule (5.5%). In the European Union, a strict price-level target raises welfare by 2% units of steady-state consumption relative to the status quo. This is also less than the optimal Taylor rule, but now only slightly so (0.1%).

5.3 What rule maximizes social welfare?

Having discussed two popular policy rules, I now turn to the unconditionally optimal policy rule. To calculate it, I compute the optimal path for the price level, p_t, in response to each of the shocks. The appendix explains the calculations.

Table 8 collects the welfare gains from different policies. Moving to the unconditionally optimal rule would raise U.S. welfare by 6.3% in steady-state consumption units, a substantial amount. This rule raises welfare by 0.7% more than the next-best alternative, an optimal Taylor rule. For the Euro-area, the benefits are less than half, 2.7%, but still substantial, and about 0.7% larger than the benefits from either an optimal Taylor rule or an optimal price-level standard.

5.4 How does parameter uncertainty affect policy?

The calculations so far have computed optimal welfare at the mean parameter estimates. Splitting the parameter vector $\boldsymbol{\theta}$, into two vectors, one with the policy parameters $\boldsymbol{\theta}^p$ and another with the non-policy parameters $\boldsymbol{\theta}^{np}$, the social welfare function becomes $\mathcal{W}(\boldsymbol{\theta}^p, \boldsymbol{\theta}^{np})$. Given the posterior distributions for the non-policy parameters, $p(\boldsymbol{\theta}^{np})$, the optimal policy rules so far were $\hat{\boldsymbol{\theta}}^p = \arg\max_{\boldsymbol{\theta}^p} \mathcal{W}\left(\boldsymbol{\theta}^p, \int \boldsymbol{\theta}^{np} dp(\boldsymbol{\theta}^{np})\right)$.

These calculations therefore ignored the parameter uncertainty associated with the $\boldsymbol{\theta}^{np}$ by evaluating welfare at its their mean estimates. I can take this estimation uncertainty into account to, following Levin et al. (2006), compute instead $\tilde{\boldsymbol{\theta}}^p = \arg\max_{\boldsymbol{\theta}^p} \int \mathcal{W}(\boldsymbol{\theta}^p, \boldsymbol{\theta}^{np}) dp(\boldsymbol{\theta}^{np})$. These policy rules are robust, in the sense that they perform well when averaged over the many different models that correspond to each of the $\boldsymbol{\theta}^{np}$. By using the Bayesian posterior density over these models, the policy rules are optimal, in the sense of Bayesian model-averaging.

Table 9 shows the robustly optimal parameters of the three policy rules that I considered, as well as their welfare gain relative to two benchmarks: the estimated status quo, and the expected welfare of using the optimal policy rules in sections 5.1 to 5.3 that ignored parameter uncertainty. Figure 15 shows the posterior distributions for welfare under the optimal policies $\mathcal{W}(\hat{\boldsymbol{\theta}}^p, \boldsymbol{\theta}^{np})$, and the robustly optimal policies $\mathcal{W}(\tilde{\boldsymbol{\theta}}^p, \boldsymbol{\theta}^{np})$.

As Giannoni (2007) found, these calculations show that a concern for robustness leads monetary policy to be more aggressive. In the Taylor rule, nominal interest rates respond more strongly to movements in both inflation and output.[12] The expected welfare gain of implementing the robustly optimal rule, instead of the optimal rule at the mean estimate is negligible, not even reaching 1 basis point of steady-state consumption. The robust price standards are less elastic, but again the welfare benefits are minimal.

6 Lessons for applied monetary policy

The aim of this paper was to use one particular model of the macroeconomy to give policy advice. That required being explicit about all of the details of the model, and some may have left readers with an unpleasant after-taste. It also required assigning plausible values to parameters in order to match important features of the data, and there is room for disagreement on how well it did so. It is clear that the model' performance is still far from the level of success one should demand to confidently give precise policy recommendations. Sections 2 and 3 tried, as much as possible, to alert the reader to 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 Altig et al (2006), Smets and Wouters (2007) and Schmitt-Grohe and Uribe (2007).

Keeping in mind these caveats, I reached some conclusions regarding current policy:

- In the United States, monetary policy shocks have had a persistent and delayed impact, with the output gap and hours peaking only 3 quarters after the shock, and inflation one quarter later. Interest rates have responded strongly to output fluctuations, and this has had a strong beneficial stabilizing effect. Interpreting all deviations for the policy rule as costly mistakes provides an upper bound on their welfare costs at 5% of consumption. Announcing monetary policy shocks in advance increases their

[12] While ϕ_π falls for Europe, again it is very hard to calculate its exact value because the welfare function is close-to-flat for high values of this parameter. The difference in welfare between using the values of 18.25 and 19 is less than 0.01%.

effectiveness at changing inflation, while mitigating their impact on output and hours, and moving interest rates gradually enhances their overall impact.

- In the Euro-area, monetary policy shocks have a similar delayed and persistent effect on inflation, hours and the output gap. Interest rates are more sensitive to output than in the United States, but the benefits from this stabilization policy are smaller. Policy errors have been smaller than in the United states, so the welfare benefit of eliminating them is only 1.4% of steady-state consumption. Announcements and gradual movements in interest rates have the same effects as in the United States.

Turning from what policy has been, to what it could (and, according to the model should) have been, I concluded that:

- In the United States, the optimal Taylor rule has interest rates responding much more strongly to the output gap than is currently the case and, doing so together with eliminating policy errors, could raise welfare by as much as 5.5% of consumption. The optimal price-level standard involves allowing for large departures from the target price-level in response to fluctuations in output. The best performing policy rule under commitment would raise welfare by 6.3%. Taking into account parameter uncertainty, the robustly-optimal Taylor rule responds more aggressively to both output and inflation, but the welfare benefits relative to the optimal rule that ignores parameter uncertainty are tiny.

- In the Euro-area, the optimal Taylor rule responds much more strongly to inflation, but less strongly to output fluctuations. Adjusting the coefficients of the Taylor rule raises welfare by 0.6% of steady-state consumption, which together with the 1.4% benefit of eliminating policy errors, leads to an overall benefit of 2% of implementing the optimal Taylor rule. The optimal price-level standard is essentially strict price-level target, and it performs almost as well as the optimal Taylor rule. The best policy rule under commitment would raise welfare by 2.7% of consumption. The robustly-optimal policy rules perform only marginally better than the rules that ignore parameter uncertainty.

Appendix

A.1. The log-linear equilibrium for the full model. At the non-stochastic 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

$$\varkappa L^{1+1/\psi} = \frac{\beta G(\nu - 1)(\gamma - 1)}{(1-\tau_w)(1-\tau_p)\nu\gamma}. \tag{38}$$

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: ν_t and γ_t which are the log-deviations of $\hat{\nu}_t$ and $\hat{\gamma}_t$, 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 $\lim_{k\to\infty} E_t[\bar{\Pi}_{t,t+1+k}]$. Log-linearizing the optimality and market clearing conditions in the consumers' problem:

$$c_{t,0} = E_t\left(c_{t+1,0} - r_t\right) \tag{39}$$

$$c_{t,j} = E_{t-j}\left(c_{t,0}\right), \tag{40}$$

$$y_{t,j} = y_t - \nu\left(p_{t,j} - p_t\right), \tag{41}$$

Log-linearizing the conditions in the firms' problem:

$$y_{t,i} = a_t + \beta l_{t,i}, \tag{42}$$

$$p_{t,i} = E_{t-i}\left[p_t + \frac{\beta(w_t - p_t) + (1-\beta)y_t - a_t - \nu_t\beta/(\bar{\nu}-1)}{\beta + \bar{\nu}(1-\beta)}\right], \tag{43}$$

$$l_{t,j} = l_t - \gamma(w_{t,j} - w_t). \tag{44}$$

Log-linearizing the conditions in the workers' problem:

$$w_{t,0} - p_t - l_{t,0}/\psi + \gamma_t/(\bar{\gamma}-1) = E_t[-r_t + w_{t+1,0} - p_{t+1} - l_{t+1,0}/\psi + \gamma_{t+1}/(\bar{\gamma}-1)], \tag{45}$$

$$w_{t,j} = E_{t-j}\left(w_{t,0}\right). \tag{46}$$

Log-linearizing the price indices and Taylor rule:

$$
\begin{aligned}
p_t &= \lambda \sum_{i=0}^{\infty} (1-\lambda)^i p_{t,i}, && (47)\\
w_t &= \omega \sum_{j=0}^{\infty} (1-\omega)^j w_{t,j}, && (48)\\
i_t &= \phi_y (y_t - y_t^n) + \phi_p \Delta p_t - \varepsilon_t, && (49)\\
i_t &= r_t + E_t (\Delta p_{t+1}) && (50)
\end{aligned}
$$

And finally, log-linearizing the aggregate quantity indices:

$$
\begin{aligned}
y_t &= g_t + c_t, && (51)\\
c_t &= \delta \sum_{j=0}^{\infty} (1-\delta)^j c_{t,j}, && (52)
\end{aligned}
$$

This set of 14 equations over time provide the equilibrium solution for the set of 14 variables ($y_{t,j}$, y_t, $c_{t,0}$, $c_{t,j}$, c_t, $l_{t,0}$, $l_{t,j}$, l_t, $w_{t,j}$, w_t, p_t, $p_{t,i}$, i_t, r_t) as a function of the 5 exogenous processes (Δa_t, ε_t, g_t, γ_t, ν_t).

A.2. The reduced-form aggregate relations. The natural levels of the variables are defined as the equilibria values when all agents are attentive (so $\delta = \lambda = \omega = 1$). In this case , since all agents are identical, $y_{t,j} = y_t$, $c_{t,0} = c_{t,j} = c_t$, $l_{t,0} = l_{t,j} = l_t$, $w_{t,j} = w_t$, and $p_{t,i} = p_t$. Solving the set of linear equations, tedious algebra shows that:

$$
\left(\frac{1+\psi}{\psi\beta}\right) y_t^n = \left(\frac{1+\psi}{\psi\beta}\right) a_t + g_t + \frac{\gamma_t}{\gamma-1} + \frac{\nu_t}{\nu-1} \tag{53}
$$

I am then ready to derive the five equations. Starting with the Phillips curve, replace $y_{t,j}$ using (41) and $p_{t,i}$ using (43) into (47) and rearrange to obtain (32). Moving to the IS, iterate (39) forward and take the limit as time goes to infinity. Then, the fact that there is complete insurance plus the fact that eventually all become aware of shocks implies that $\lim_{i\to\infty} E_t (c_{t+i,0}) = \lim_{i\to\infty} E_t \left[y_{t+i}^n\right] \equiv y_t^\infty$. Using the definition of the long rate R_t and replacing for $c_{t,0}$ in (40) and (52) gives an expression for aggregate consumption. Replacing it in (51) and using the fact that $\lim_{i\to\infty} E_t [g_{t+i}] = 0$ gives the IS curve in (33). Next, I turn to the wage curve. Taking very similar steps as in the IS, iterate (45) forward and use the solution to replace for $w_{t,0}$ in (46) Combining the $w_{t,j}$ in the aggregator for w_t in (48) and replacing out $l_{t,j}$ using (44) gives the wage curve in (34). Fourth, aggregating (42) over i gives the aggregate production function in (35). Fifth and finally, the expressions for the nominal interest rate in (49) and (50) give the Taylor rule in (36).

A.3. Numerical solution of the model. The 5 equations in section 2.3 together with the initial condition $p_{-1} = 0$ define an equilibrium in the 5 variables $(y_t, p_t, w_t, l_t, i_t)$ as function of the five stochastic variables $(\varepsilon_t, \Delta a_t, g_t, \nu_t, \gamma_t)$. Mankiw and Reis (2007) showed how to use a method of undetermined coefficients to solve this model. Letting $s \in S = \{\Delta a,\ g,\ \nu,\ \gamma,\ \varepsilon\}$ denote the different shocks, they write $p_t = \sum_{s\in S} \sum_{n=0}^{\infty} \hat{p}_n(s) e^s_{t-n}$ and their Proposition 1 solves for the scalars $\hat{p}_n(s)$ that measure the impact of shock s at lag n on the price level. Their Corollary 1 in turn solves for the equilibrium coefficients $\hat{y}_n(s)$, $\hat{\imath}_n(s)$, $(\hat{w} - \hat{p})_n(s)$ and $\hat{l}_n(s)$ that govern the solution for output, nominal interest rates, real wages and hours.

A.4. Evaluating welfare. The utilitarian social welfare function is:

$$\int_0^1 E_0 \left[\sum_{t=0}^{\infty} \xi^t \left(\ln(C_{t,j}) - \frac{\varkappa L_{t,j}^{1+1/\psi}}{1 + 1/\psi} \right) \right] dj, \tag{54}$$

and I focus on its unconditional expectation. Multiplying the expression above by $(1 - \xi)$ and taking the limit as $\xi \to 1$, the policy objective is:

$$\int_0^1 \left(E\left(\ln(C_{t,j})\right) - \frac{\varkappa E\left(L_{t,j}^{1+1/\psi}\right)}{1 + 1/\psi} \right) dj. \tag{55}$$

Recall the definition of the log-linearized values: $c_{t,j} = \ln(C_{t,j}) - \ln(C)$ and $l_{t,j} = \ln(L_{t,j}) - \ln(L)$. Social welfare then is:

$$\ln(C) + \int_0^1 \left(E\left(c_{t,j}\right) - \frac{\varkappa L^{1+1/\psi} E\left(e^{(1+1/\psi) l_{t,j}}\right)}{1 + 1/\psi} \right) dj. \tag{56}$$

At this point I make one assumption common in this literature: 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 non-stochastic steady state is an efficient equilibrium without uncertainty. These assumptions allow us to focus monetary policy on the task of stabilizing economic activity (Woodford, 2003a). From (38), they imply that $\varkappa L^{1+1/\psi} = \beta$.

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:

$$\ln(C) - \frac{\beta}{1+1/\psi}\int_0^1 \exp\left((1+1/\psi)^2 Var(l_{t,j})\right) dj. \tag{57}$$

Because $l_{t,j}$ is a normal variable, $Var(l_{t,j})$ is a linear function of the variance of the exogenous shocks. As we will see, these are small in the data, so approximating $\exp\left(Var(l_{t,j})\right)$ by $1 + Var(l_{t,j})$ involves very little numerical error and social welfare becomes:

$$\ln(C) + \beta(1+1/\psi) - \beta(1+1/\psi)\int_0^1 Var(l_{t,j})dj \tag{58}$$

Using the distribution of workers according to when they last updated, this becomes:

$$\ln(C) + \beta(1+1/\psi) - \beta(1+1/\psi)\omega\sum_{j=0}^{\infty}(1-\omega)^j Var(l_{t,j}) \tag{59}$$

Next, note that combining (44) with (45) and (46) to replace out for $w_{t,0}$ gives the following expressions:

$$l_{t,j} = l_t - \gamma(w_{t,j} - w_t) \tag{60}$$

$$w_{t,j} = E_{t-j}\left(p_t + \frac{l_{t,j}}{\psi} - \frac{\gamma_t}{\bar{\gamma}-1} - R_t + y_n^{\infty}\right) \tag{61}$$

Using a method of undetermined coefficients, guess that $l_{t,j} = \sum_{s\in S}\left(\sum_{n=0}^{j-1}\tilde{l}_n(s) + \sum_{n=j}^{\infty}\check{l}_n(s)\right)e_{t-n}^s$ and solve to find that:

$$\tilde{l}_n(s) = \hat{l}_n(s) + \gamma\hat{w}_n(s) \text{ for all } s, \tag{62}$$

$$\frac{(\gamma+\psi)\check{l}_n(s)}{\gamma\psi} = \frac{\hat{l}_n(s)}{\gamma} + (\hat{w}-\hat{p})_n(s) + \frac{\hat{y}_n(s)}{\Delta_n} + \begin{cases} 0 & \text{for } s = \varepsilon, a, \nu \\ \rho_\gamma^n/(\gamma-1) & \text{for } s = \gamma \\ -\rho_g^n/\Delta_n & \text{for } s = g \end{cases} \tag{63}$$

where $\Delta_n = \delta\sum_{i=0}^{n}(1-\delta)^i$. From this, it follows that:

$$Var(l_{t,j}) = \sum_{s\in S}\left(\sum_{n=0}^{j-1}\tilde{l}_n(s)^2 + \sum_{n=j}^{\infty}\check{l}_n(s)^2\right)\sigma^2(s) \tag{64}$$

Finally, some grouping shows that

$$\omega \sum_{j=0}^{\infty}(1-\omega)^j \left(\sum_{n=0}^{j-1} \tilde{l}_n(s)^2 + \sum_{n=j}^{\infty} \check{l}_n(s)^2 \right) = \sum_{n=0}^{\infty} \left[(1-\Omega_n)\tilde{l}_n(s)^2 + \Omega_n \check{l}_n(s)^2 \right] \tag{65}$$

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 is:

$$\mathcal{W}(\boldsymbol{\theta}) = -\sum_{s \in S} \sum_{n=0}^{\infty} \left[(1-\Omega_n)\tilde{l}_n(s)^2 + \Omega_n \check{l}_n(s)^2 \right] \sigma_s^2, \tag{66}$$

wits $\tilde{l}_n(s)$ and $\check{l}_n(s)$ defined above. 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 (59) to calculate:

$$\exp(\beta(1+1/\psi)\left(\mathcal{W}(\boldsymbol{\theta}^{(1)}) - \mathcal{W}(\boldsymbol{\theta}^{(0)})\right))$$

A.5. Optimal policy. The set of non-policy parameters is always: $\boldsymbol{\theta}^{np} = (\beta, \psi, \nu, \gamma, \rho_{\Delta a}, \sigma_{\Delta a}, \rho_g, \sigma_g, \rho_\nu, \sigma_\nu, \rho_\gamma, \sigma_\gamma, \delta, \omega, \lambda)$. The set of policy parameters for the Taylor rule is $\boldsymbol{\theta}^p = (\phi_y, \phi_\pi, \rho_\varepsilon, \sigma_\varepsilon)$. Since the welfare function declines monotonically with σ_ε^2 and the policy controls it, it is obvious that $\sigma_\varepsilon^2 = 0$ at the optimum (and, as result, ρ_ε is irrelevant.

To find the optimal ϕ_y and ϕ_π I numerically maximized welfare subject to the restriction that $\phi_\pi > 1$.

For an elastic price standard, the only policy coefficient is ϕ. Using the corollary in Mankiw and Reis (2007), one can show that, regardless of the policy rule, in this model $\hat{y}_n(s) = \Psi_n \hat{p}_n(s) + \Phi_n(s)$ where Ψ_n and $\Phi_n(s)$ are messy functions of $\boldsymbol{\theta}^{np}$ while for the natural rate of output $\hat{y}_n^n(s) = \Xi_n(s)$. Combining this with the elastic price standard, it follows that $\hat{p}_n(s) = \phi(\Phi_n(s) - \Xi_n(s))/(1-\phi\Psi_n)$. Given this solution for $\hat{p}_n(s)$, welfare can be evaluated using the Corollary 1 in Mankiw and Reis (2007), and the results in A.4. This was then numerically maximized with respect to ϕ in the real line.

For the optimal unconditional policy, let $\bar{S}$ be the set of non-policy shocks $\{\Delta a, g, \nu,$

$\gamma\}$. The problem of finding the optimal policy is to:

$$\min_{\left\{\{\hat{p}_n(s)\}_{n=1}^{\infty}\right\}_{s\in\bar{S}}} \left\{ \sum_{s\in\bar{S}} \sum_{n=0}^{\infty} \left[(1-\Omega_n)\tilde{l}_n(s)^2 + \Omega_n \check{l}_n(s)^2 \right] \sigma_s^2 \right\} \tag{67}$$

Since the problem is additively separable across the different elements of $\bar{S}$, we can solve the four separate problems instead:

$$\min_{\{\hat{p}_n(s)\}_{n=1}^{\infty}} \sum_{n=0}^{\infty} \left[(1-\Omega_n)\tilde{l}_n(s)^2 + \Omega_n \check{l}_n(s)^2 \right] \sigma_s^2 \text{ for each } s \in \bar{S} \tag{68}$$

Using (62)-(63) together with the solutions in the Mankiw-Reis corollary, one can derive, after much algebra, that:

$$\tilde{l}_n(s) = a_n \hat{p}_n(s) + b_n(s) \tag{69}$$

$$\check{l}_n(s) = c_n \hat{p}_n(s) + d_n(s) \tag{70}$$

for some messy functions a_n, $b_n(s)$, c_n, and $d_n(s)$ that depend on the non-policy structural parameters of the economy. But then, the first-order conditions of (68) lead to the solution:

$$\hat{p}_n(s) = -\frac{(1-\Omega_n)a_n b_n(s) + \Omega_n c_n d_n(s)}{(1-\Omega_n)a_n^2 + \Omega_n c_n^2}. \tag{71}$$

for $n = 1, 2...$ and $s \in \bar{S}$. A simple algorithm uses the non-policy structural parameters of the economy to evaluate the expressions: a_n, $b_n(s)$, c_n, and $d_n(s)$ and calculates the optimal policy directly using (71). This takes only a few seconds. Given the optimal $\hat{p}_n(s)$, the Mankiw-Reis corollary gives the solution for the other variables, and the results in A.4. gives optimal welfare.

A.6. Robustly optimal policy. This exercise consists of finding:

$$\hat{\boldsymbol{\theta}}^p = \arg\max_{\boldsymbol{\theta}^p} \mathcal{W}\left(\boldsymbol{\theta}^p, \int \boldsymbol{\theta}^{np} dp(\boldsymbol{\theta}^{np})\right).$$

For the Taylor rule case, $\boldsymbol{\theta}^p = (\phi_y, \phi_\pi)$, and the integral was calculated by averaging over 10,000 draws from the posterior density. A first grid search for the optimum with 0.2 jumps started from the non-robust optimal policy rules. Two further grid searches, with 0.05 and 0.01 jumps, were then undertaken around the candidate optimum.

For the elastic price standard, $\boldsymbol{\theta}^p = \phi$ and the procedure was the same.

A.6. Europe EMS sample. Table A.1 shows the posterior estimates using Euro-data since the start of the European Monetary System (EMS) and the practice of pegging exchange rates. It is a stretch to apply the model to this extended dataset, since it includes changes in the members of the European Community, and involves frequent devaluations and exits from the EMS.

References

Abel, Andrew B., Janice C. Eberly and Stavros Panageas (2007) "Optimal Inattention to the Stock Market.." *American Economic Review*, forthcoming.

Adam, Klaus (2007) "Optimal Monetary Policy with Imperfect Common Knowledge." *Journal of Monetary Economics*, vol. 54 (2), pp. 276-301.

Altig, David, Lawrence Christiano, Martin Eichenbaum, and Jesper Linde (2005) "Firm-Specific Capital, Nominal Rigidities and the Business Cycle." NBER working paper 11034.

Amato, Jeffery D. and Hyun Song Shin (2003). "Public and Private Information in Monetary Policy Models." Bank of International Settlements working paper 138.

Andrés, Javier, Edward Nelson, and David López-Salido (2005). "Sticky-Price Models and the Natural Rate Hypothesis." *Journal of Monetary Economics*, 52 (5), 1025-1053.

Bacchetta, Philippe and Eric van Wincoop (2006) "Incomplete Information Processing: A Solution to the Forward Discount Puzzle," Federal Reserve Bank of San Francisco working paper 2006-35.

Ball, Laurence, N. Gregory Mankiw, and Ricardo Reis (2005). "Monetary Policy for Inattentive Economies." *Journal of Monetary Economics*, 52 (4), 703-725.

Bayoumi, Tamim (2004). "GEM: A New International Macroeconomic Model." IMF Occasional Papers 239.

Bernanke, Ben (2004). "Gradualims." Remarks at an economics luncheon co-sponsored by the Federal Reserve Bank of San Francisco (Seattle Branch) and the University of Washington, Seattle, Washington, May 20, 2004.

Branch, William A. (2007). "Sticky Information and Model Uncertainty in Survey Data on Inflation Expectations." *Journal of Economic Dynamics and Control*, vol. 31 (1), pp. 245-276.

Carroll, Christopher D. (2003). "Macroeconomic Expectations of Households and Professional Forecasters." *Quarterly Journal of Economics*, vol. 118, pp. 269-298.

Carroll, Christopher D. and Jiri Slacalek (2006) "Sticky Expectations and Consumption Dynamics." Manuscript, Johns Hopkins University.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans (2005) "Nominal Rigidities and the Dynamic Effects of an Shock to Monetary Policy." *Journal of Political Economy*, vol. 113 (1), pp. 1-45.

Christiano, Lawrence J., Roberto Motto, and Massimo Rostagno (2007) "Shocks, Sturctures, or Monetary Policies? The Euro-Area and the US after 2001." NBER working paper 13521.

Coibion, Olivier (2006). "Testing the Sticky Information Phillips Curve." Manuscript, University of Michigan.

Dopke, Jorg, Jonas Dovern, Ulrich Fritsche and Jirka Slacalek (2006a). "The Dynamics of European Inflation Expectations." DIW Berlin Discussion Paper 571.

Dopke, Jorg, Jonas Dovern, Ulrich Fritsche and Jirka Slacalek (2006b). "Sticky Information Phillips Curves: European Evidence." DIW Berlin Discussion Paper 615.

Edge, Rochelle M., Michael T. Kiley, and Jean-Philippe Laforte (2007). "Natural rate measures in an estimated DSGE model of the U.S. economy." Finance and Economics Discussion Series 2007-08.

Erceg, Christopher J., Luca Guerrieri, and Christopher Gust (2006). "SIGMA: A New Open Economy Model for Policy Analysis." *International Journal of Central Banking*, vol. 2 (1), pp. 1-50.

Gabaix, Xavier and David Laibson (2002). "The 6D Bias and the Equity Premium Puzzle." In Ben Bernanke and Kenneth Rogoff, eds., *NBER Macroeconomics Annual 2001*, pp. 257-312, Cambridge: MIT Press.

Gabaix, Xavier and David Laibson (2006). "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *Quarterly Journal of Economics*, vol. 121 (2), pp. 505-540.

Giannoni, Marc (2007). "Robust Optimal Monetary Policy in a Forward-Looking Model with Parameter and Shock Uncertainty." *Journal of Applied Economics*, vol. 22 (1), pp. 179-213.

Gorodnichenko, Yuriy (2006). "Monetary Policy and Forecast Dispersion: a Test of the

Sticky Information Model." Manuscript, University of Michigan.

Hall, Robert E. (1984). "Monetary Strategy with an Elastic Price Standard." in *Price Stability and Public Policy: A Symposium Sponsored by the Federal Reserve Bank of Kansas City*, Federal Reserve Bank of Kansas City, Jackson Hole.

Hellwig, Christian (2002). "Public Announcements, Adjustment Delays and the Business Cycle." Manuscript, U.C.L.A.

Hellwig, Christian and Laura Veldkamp (2007). "Knowing What Others Know: Coordination Motives in Information Acquisition." Manuscript, U.C.L.A.

Khan, Hashmat and Zhenhua Zhu (2006). "Estimates of the Sticky-Information Phillips Curve for the United States." *Journal of Money, Credit and Banking*, vol. 38 (1), pp. 195-207.

Kiley, Michael T. (2007). "A Quantitative Comparison of Sticky-Price and Sticky-Information Models of Price Setting." *Journal of Money, Credit and Banking*, vol. 39 (s1), pp. 101-125.

Klenow, Peter J. and Jonathan L. Willis (2007) "Sticky Information and Sticky Prices." *Journal of Monetary Economics*, forthcoming.

Knotek, Edward S. (2006) "A Tale of Two Rigidities: Sticky Prices in a Sticky-Information Environment." Manuscript, Federal Reserve Bank of Kansas City.

Korenok, Oleg (2005). "Empirical Comparison of Sticky Price and Sticky Information Models." VCU School of Business working paper 501.

Korenok, Oleg and Norman 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), 905-930.

Korenok, Oleg and Norman R. Swanson (2006). "How Sticky Is Sticky Enough? A Distributional and Impulse Response Analysis of New Keynesian DSGE Models." *Journal of Money, Credit and Banking*, forthcoming.

Krusell, Per and Anthony A. Smith Jr. (1998). "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy*, vol. 106 (5), pp. 867-896.

Kydland, Finn E. and Edward C. Prescott (1982) "Time to Build and Aggregate Fluctuations," *Econometrica*, vol. 50 (6), pp. 1345-1370.

Laforte, Jean-Philippe (2007). "Pricing Models: A Bayesian DSGE approach for the US Economy." *Journal of Money, Credit and Banking*, vol. 39 (s1), pp. 127-154.

Levin, Andrew T., Alexei Onatski, John C. Williams and Noah Williams (2006). "Monetary Policy Under Uncertainty in Micro-Founded Macroeconometric Models." In *NBER Macroeconomics Annual 2005*, edited by Mark Gertler and Kenneth Rogoff, MIT Press: Cambridge.

Lorenzoni, Guido (2006). "A Theory of Demand Shocks." Manuscript, M.I.T.

Lucas, Robert E., Jr. (1980) "Methods and Problems in Business Cycle Theory," *Journal of Money, Credit and Banking*, vol. 12 (4), pp.696-715.

Mackowiak, Bartosz and Mirko Wiederholt (2007) "Optimal Sticky Prices under Rational Inattention." Manuscript, Northwestern University.

Mankiw, N. Gregory and Ricardo Reis (2002). "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *Quarterly Journal of Economics*, 117 (4), 1295-1328.

Mankiw, N. Gregory and Ricardo Reis (2003) "Sticky Information: A Model of Monetary Non-neutrality and Structural Slumps." In Philippe Aghion, Roman Frydman, Joseph E. Stiglitz and Michael Woodford, eds., *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, Princeton: Princeton University Press.

Mankiw, N. Gregory and Ricardo Reis (2006) "Pervasive Stickiness." *American Economic Review*, vol. 96 (2), 164-169.

Mankiw, N. Gregory and Ricardo Reis (2007) "Sticky Information in General Equilibrium," *Journal of the European Economic Association,* vol. 2 (2-3), pp. 603-613.

Mankiw, N. Gregory, Ricardo Reis and Justin Wolfers (2004) "Disagreement about Inflation Expectations." In Mark Gertler and Kenneth Rogoff, eds., *NBER Macroeconomics Annual 2003*, pp. 209-248.

Molinari, Benedetto (2007). "The Role of Sticky Information in Inflation Dynamics: Estimates and Findings." Ente Eunadi Temi di Ricerca no. 50.

Morris, Stephen and Hyun Song Shin (2006). "Inertia of Forward-Looking Expectations." *American Economic Review*, vol. 96 (2), 152-157.

Nunes, Ricardo C. (2006) "On the Epidemiological Microfoundations of Sticky Information." Manuscript, Universitat Pompeu Fabra.

Paustian, Matthias and Ernest Pytlarczyk (2006). "Sticky Contracts or Sticky Information? Evidence from an Estimated Euro Area DSGE Model." Manuscript, Bowling Green

State University.

Pickering, Andrew (2004). "Sticky Information and the Phillips Curve: A Tale of Two Forecasts." Manuscript, University of Bristol.

Reis, Ricardo (2006a) "Inattentive Consumers," *Journal of Monetary Economics*, vol. 53 (8), pp. 1761-1800.

Reis, Ricardo (2006b) "Inattentive Producers," *Review of Economic Studies*, vol. 73 (3), pp. 793-821.

Rich, Robert J. and Joseph S. Tracy (2006). "The Relationship between Expected Inflation, Disagreement, and Uncertainty: Evidence from Matched Point and Density Forecasts." Federal Reserve Bank of New York Staff Report No. 253.

Rotemberg, Julio and Michael Wodford (1997). "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy." *NBER Macroeconomics Annual 1997*, pp. 297-346.

Schmitt-Grohe, Stephanie and Marin Uribe (2007). "Optimal Inflation Stabilization in a Medium-Scale Macroeconomic Model." In *Monetary Policy Under Inflation Targeting*, edited by Frederic S. Mishkin and Klaus Schmidt-Hebbel. Central Bank of Chile, Santiago, Chile.

Sims, Christopher A. (1998). "Stickiness." *Carnegie-Rochester Conference Series on Public Policy*, vol. 49 (1), pp. 317-356.

Sims, Christopher A. (2003). "Implications of Rational Inattention." *Journal of Monetary Economics*, vol. 50, pp. 665-690.

Smets, Frank and Rafael Wouters (2003). "An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area." *Journal of the European Economic Association*, 1 (5), pp. 1123–1175.

Smets, Frank and Rafael Wouters (2007). "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review*, 97 (3), pp. 586-606.

Sveen, Tommy and Lutz Weinke (2005). "New Perspectives on Capital, Sticky Prices, and the Taylor Principle," *Journal of Economic Theory*, vol. 123 (1), pp. 21-39.

Taylor, John B. (1993) "Discretion versus Policy Rules in Practice." *Carnegie-Rochester Conference Series in Public Policy*, vol. 39, pp. 195-214.

Woodford, Michael (2003a). *Interest and Prices*, Princeton: Princeton University Press.

Woodford, Michael (2003b). "Imperfect Common Knowledge and the Effects of Monetary

Policy." In Philippe Aghion, Roman Frydman, Joseph E. Stiglitz and Michael Woodford, eds., *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, Princeton: Princeton University Press.

Woodford, Michael (2005). "Firm-Specific Capital and the New Keynesian Phillips Curve." *International Journal of Central Banking*, vol. 1 (2), pp. 1-46.

Figure 1. Impulse responses to one-standard deviation non-policy shocks in the United States

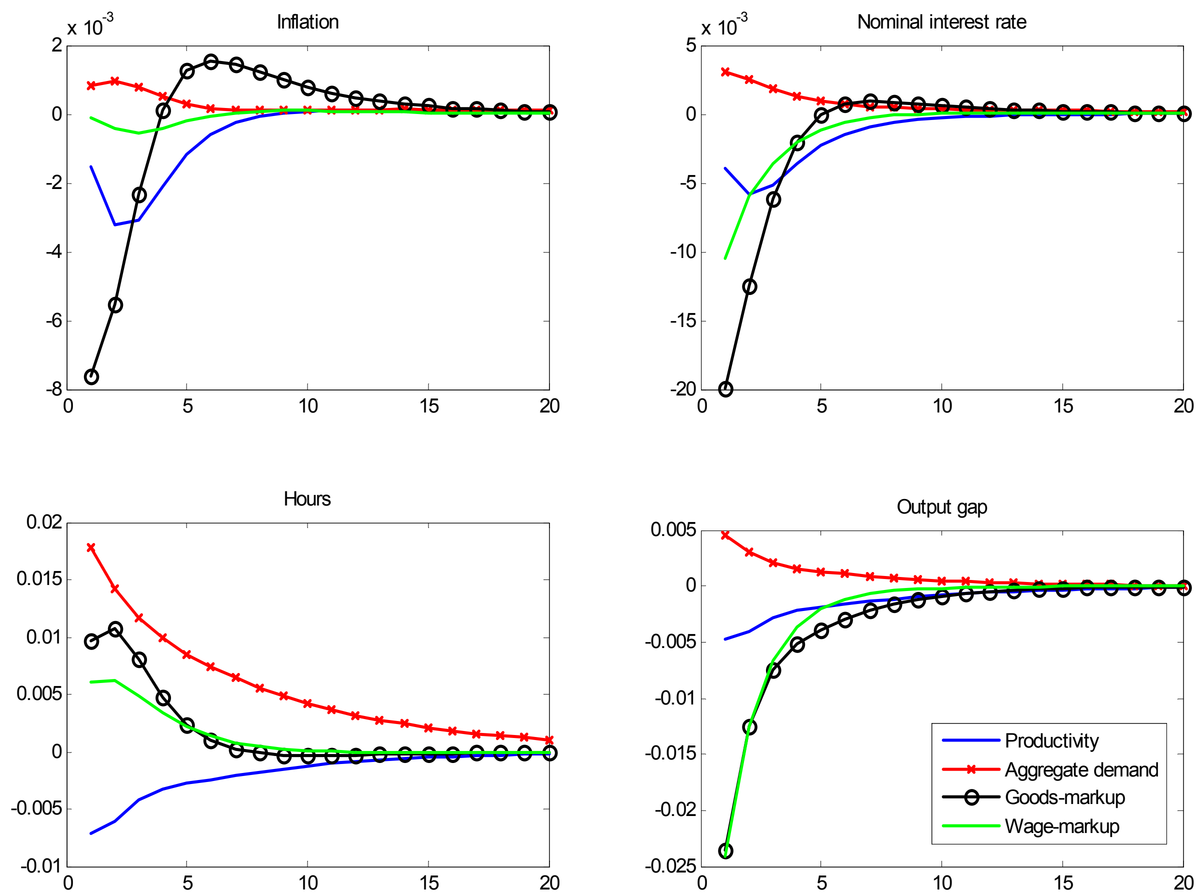

Figure 2. Impulse responses to one-standard deviation non-policy shocks in the Euro-area

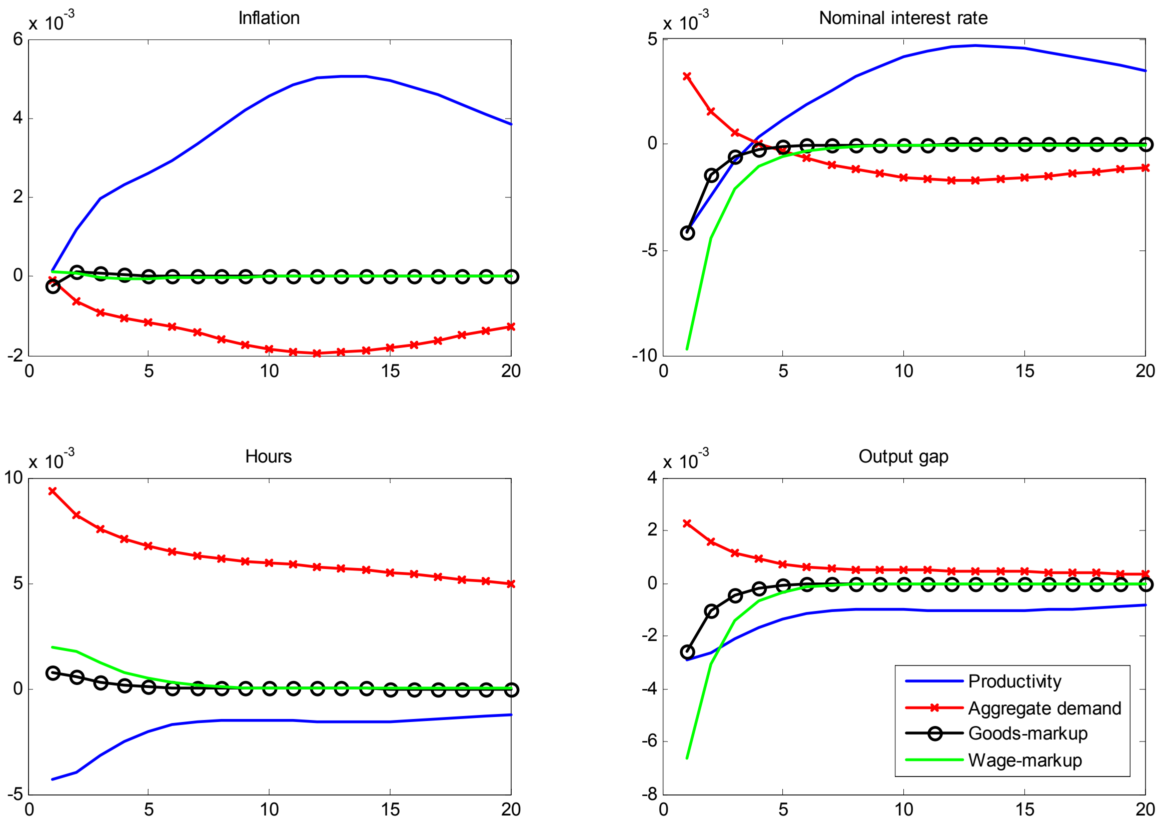

Figure 3. Impulse responses to a one-standard deviation monetary policy shocks in the United States

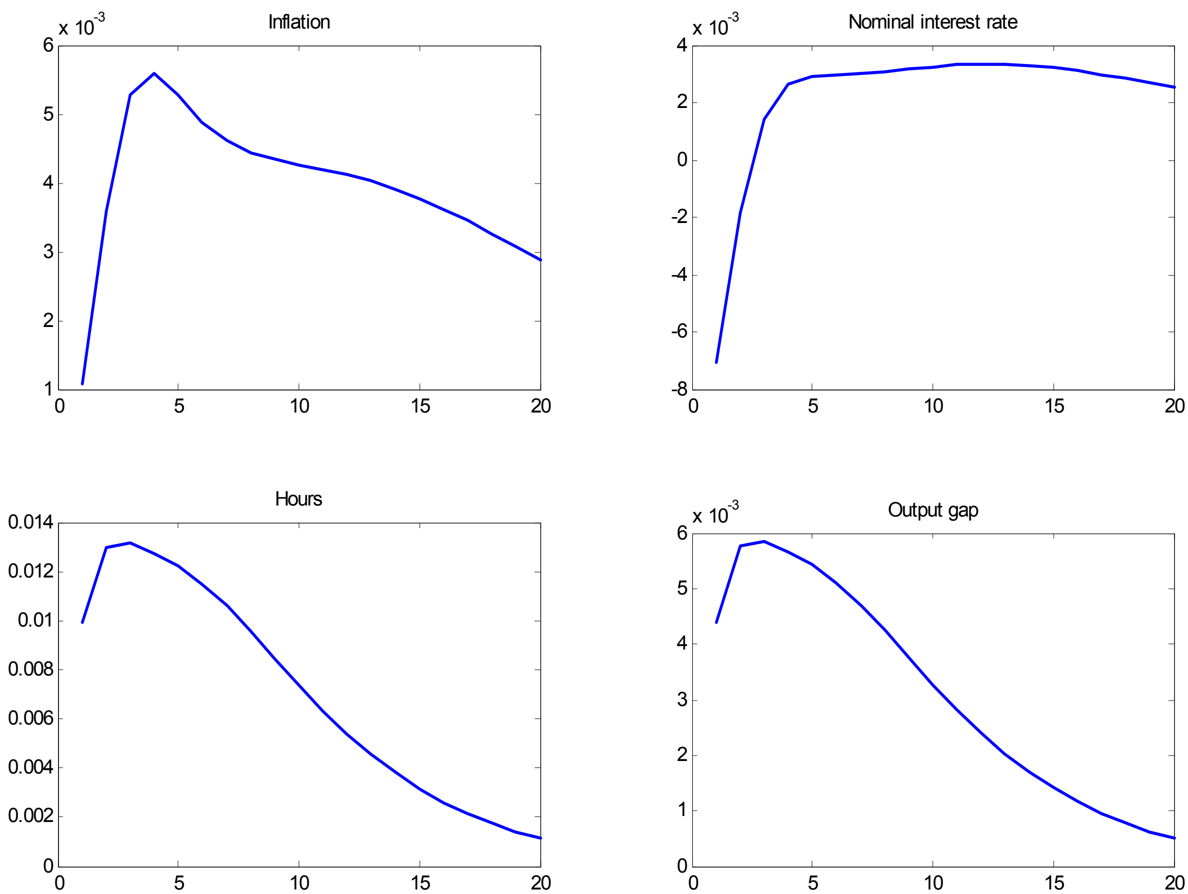

Figure 4. Impulse responses to a one-standard deviation monetary policy shocks in the Euro-area

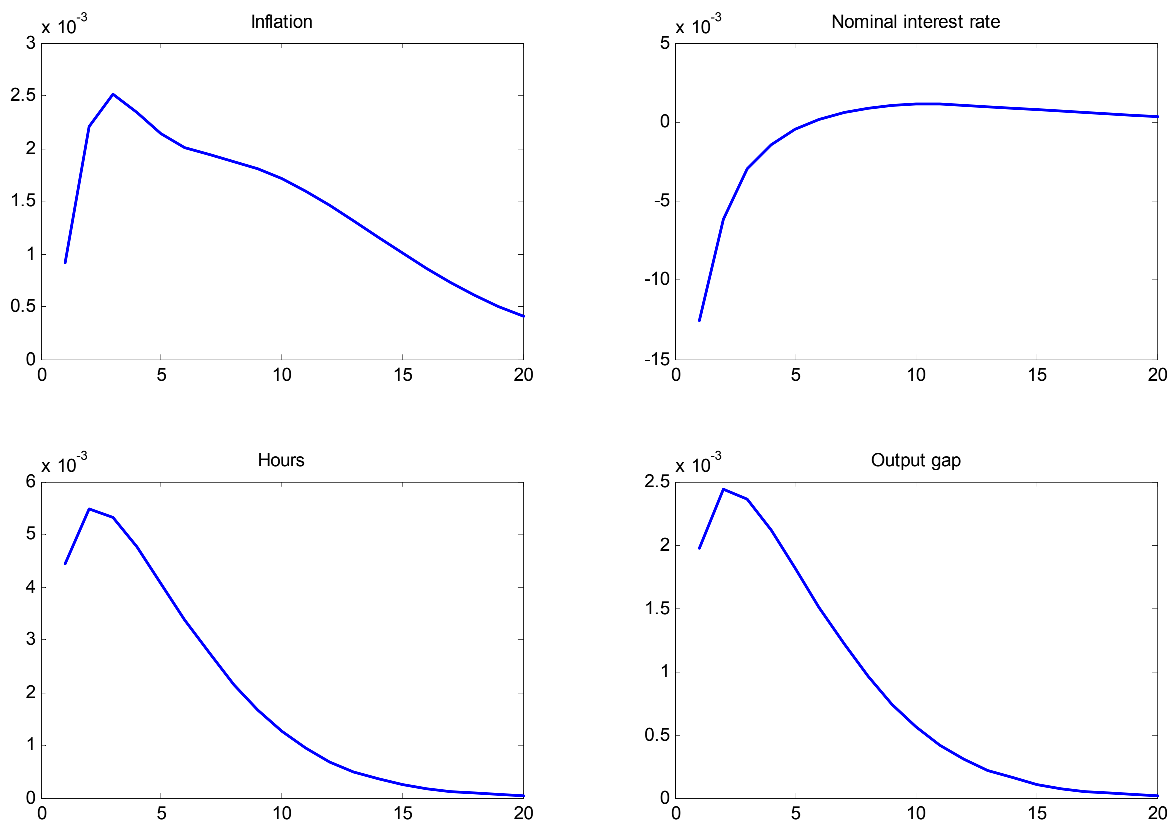

Figure 5. Impulse responses to a one-standard deviation monetary policy shock in the United States with and without output stabilization concerns.

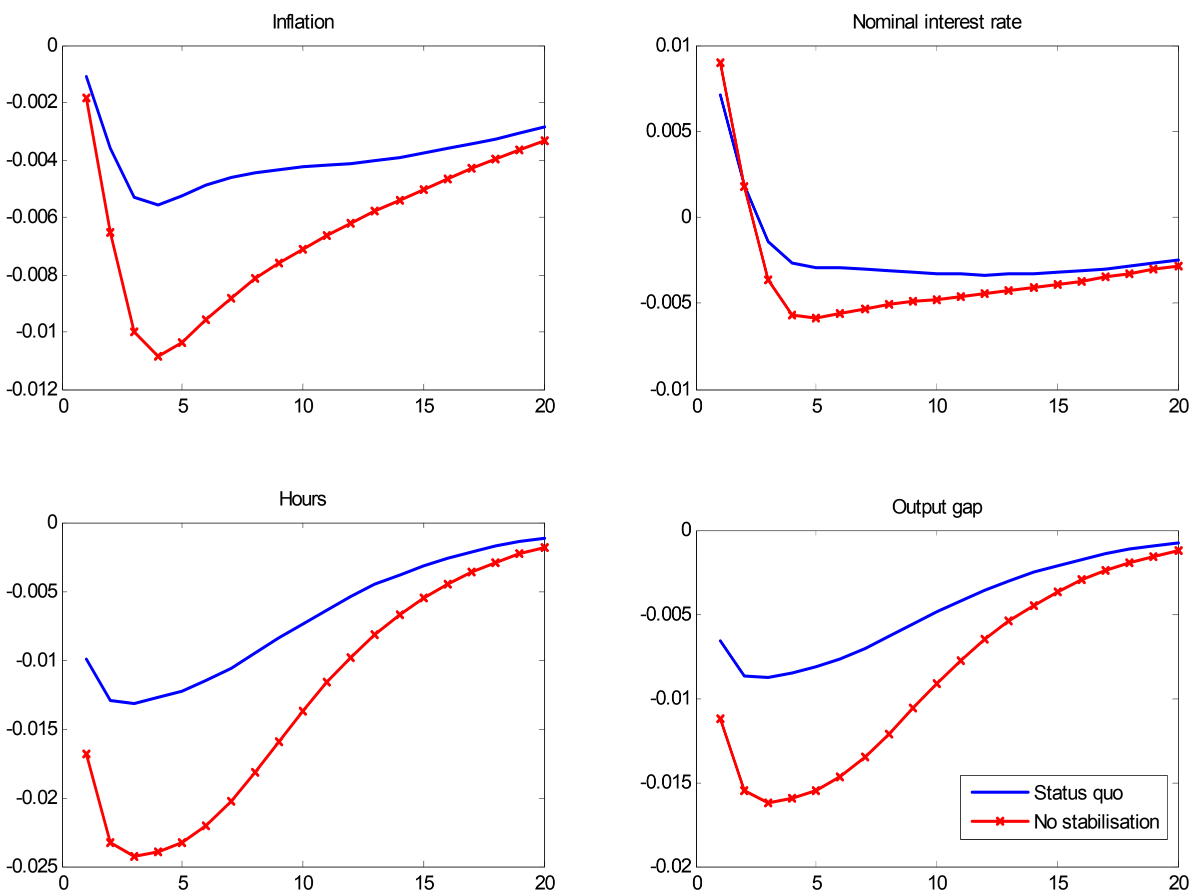

Figure 6. Impulse responses to a one-standard deviation aggregate productivity shock in the United States with and without output stabilization concerns.

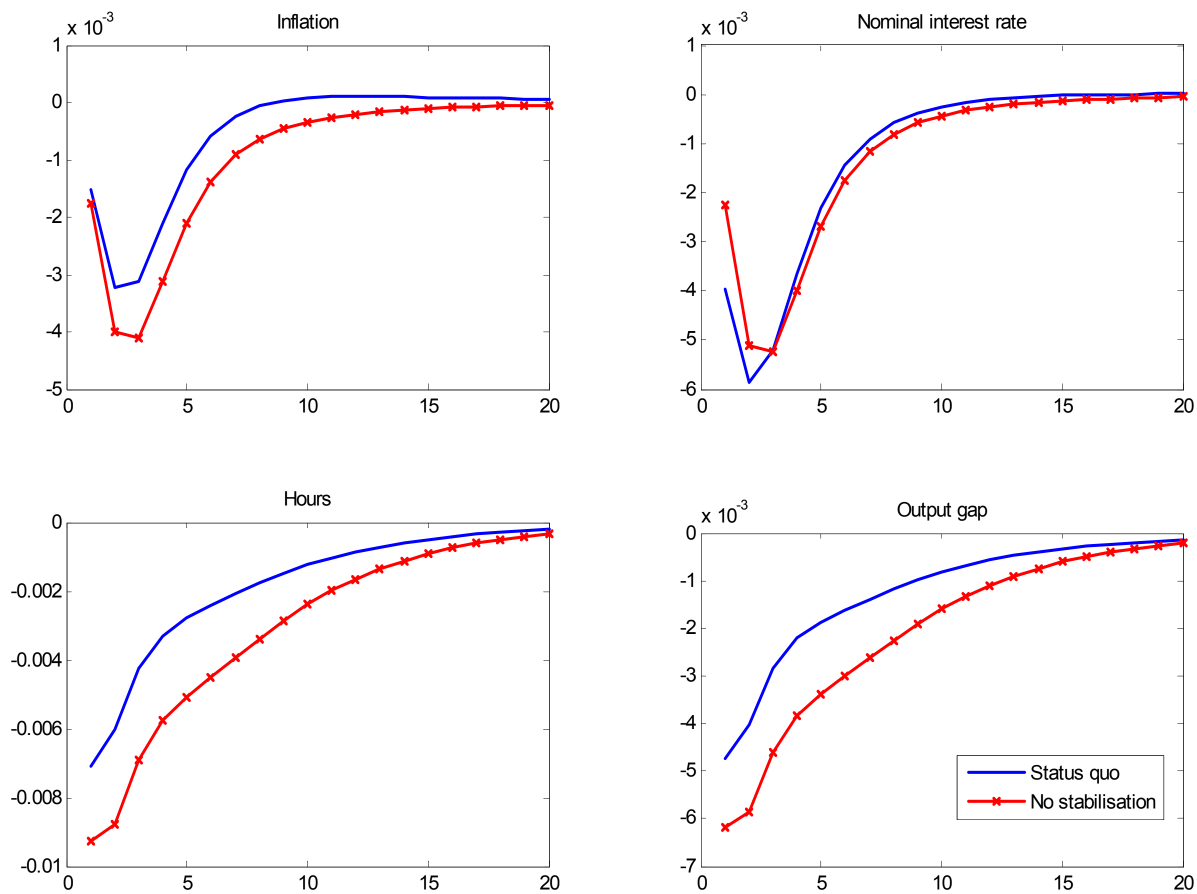

Figure 7. Impulse responses to a one-standard deviation aggregate demand shock in the United States with and without output stabilization concerns.

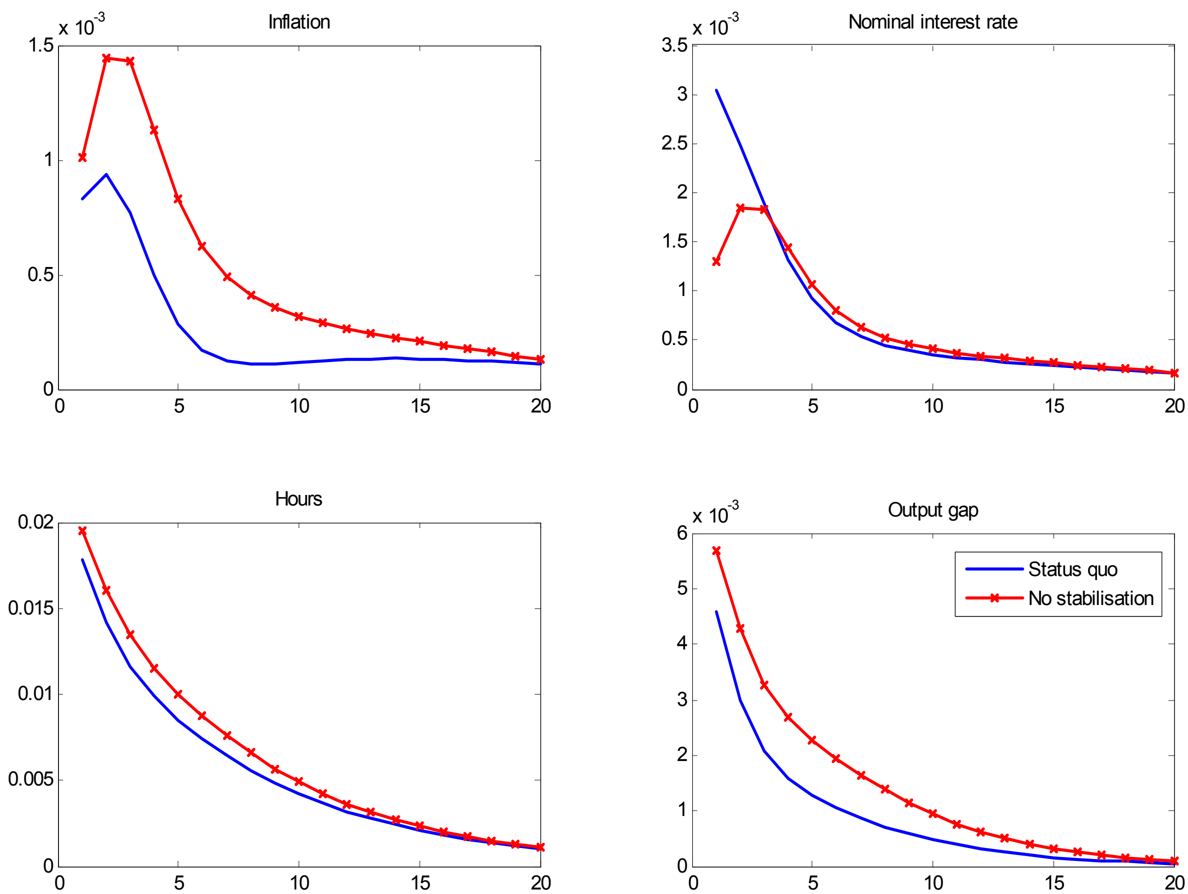

Figure 8. Impulse responses to a one-standard deviation monetary policy shock in the Euro-area with and without output stabilization concerns.

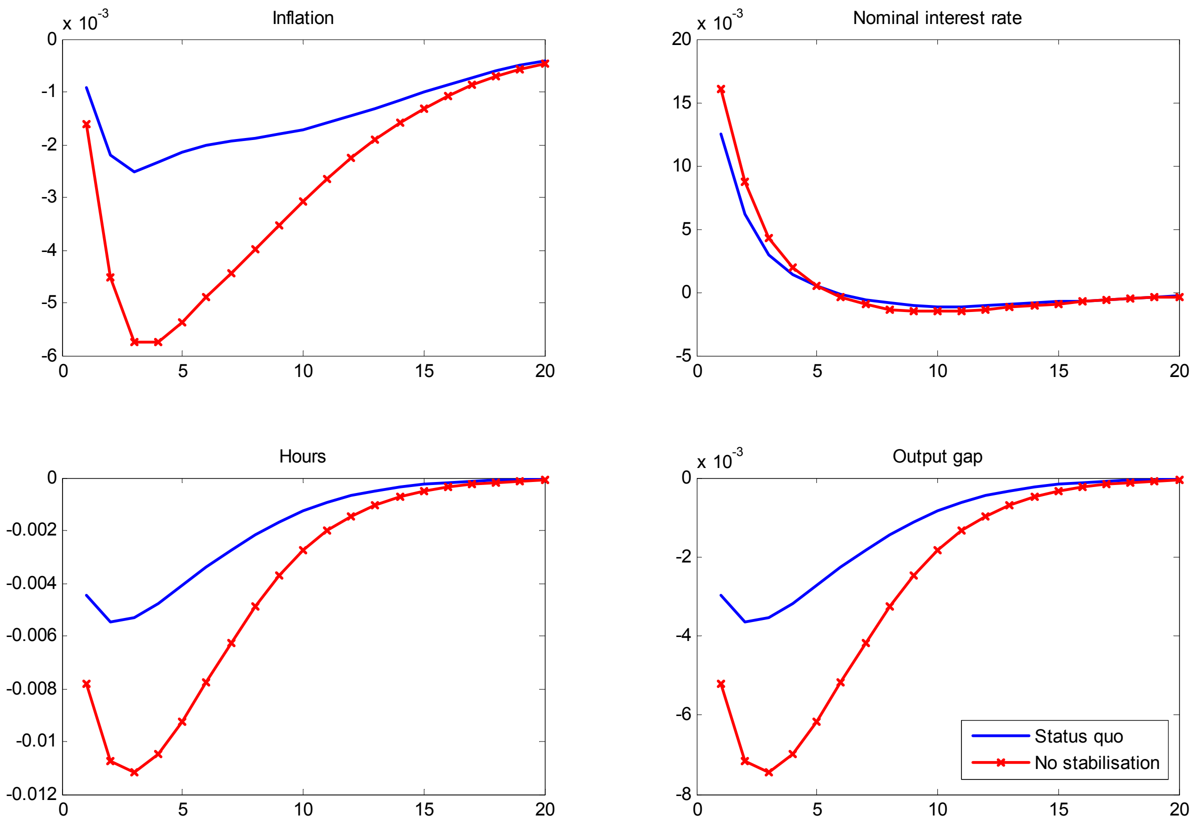

Figure 9. Impulse responses to a one-standard deviation aggregate productivity shock in the Euro-area with and without output stabilization concerns.

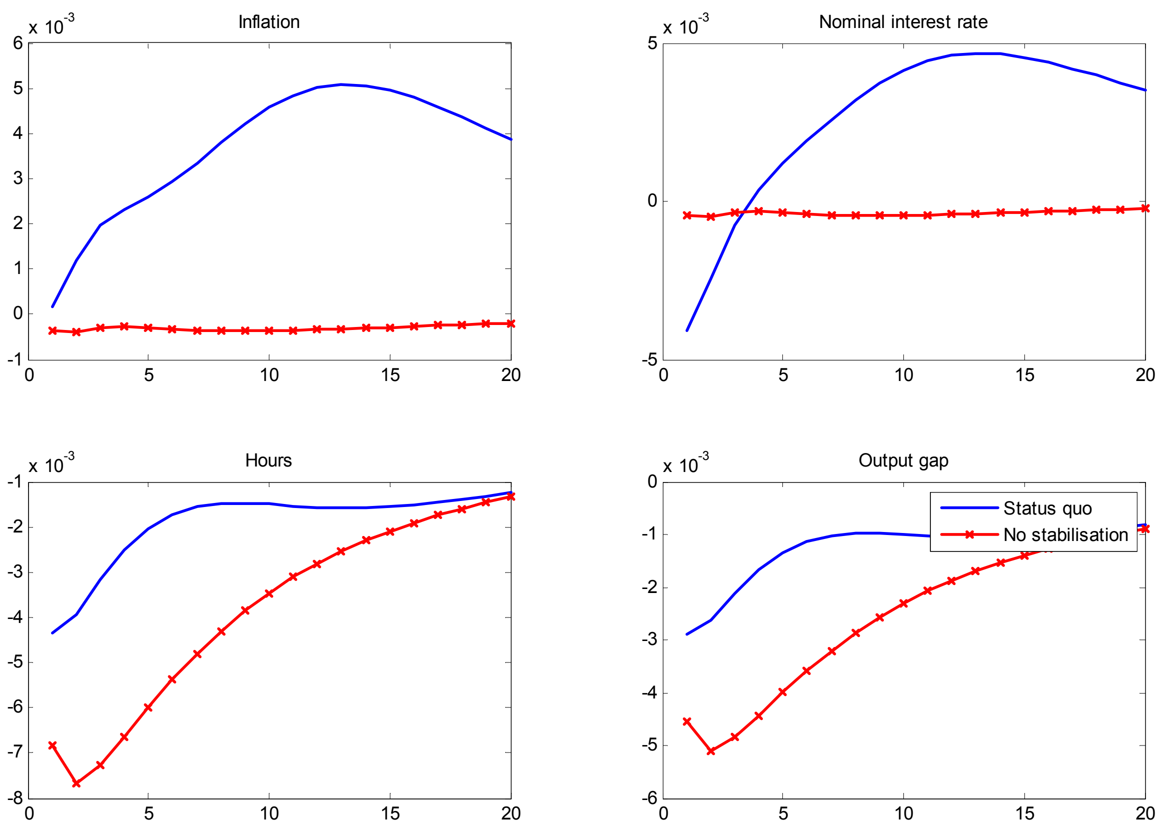

Figure 10. Impulse responses to a one-standard deviation aggregate demand shock in the Euro-area with and without output stabilization concerns.

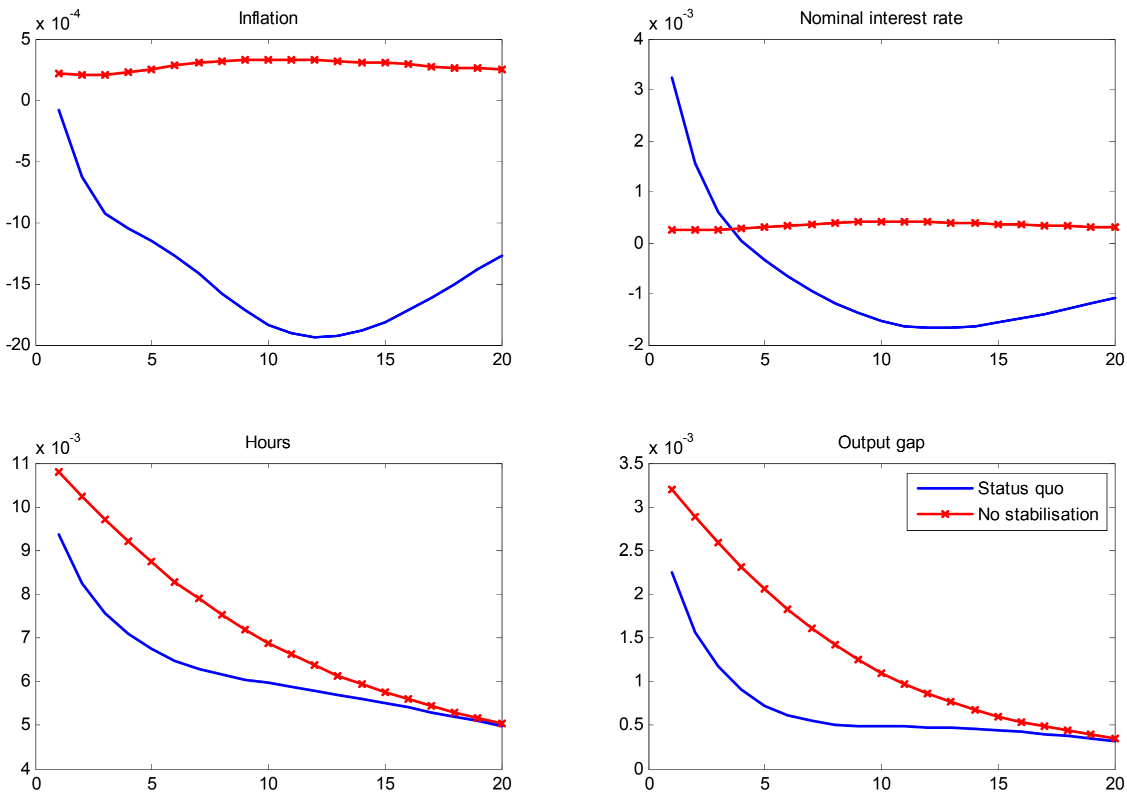

Figure 11. Impulse responses to a one-standard deviation monetary policy shock in the United States announced in advance or not

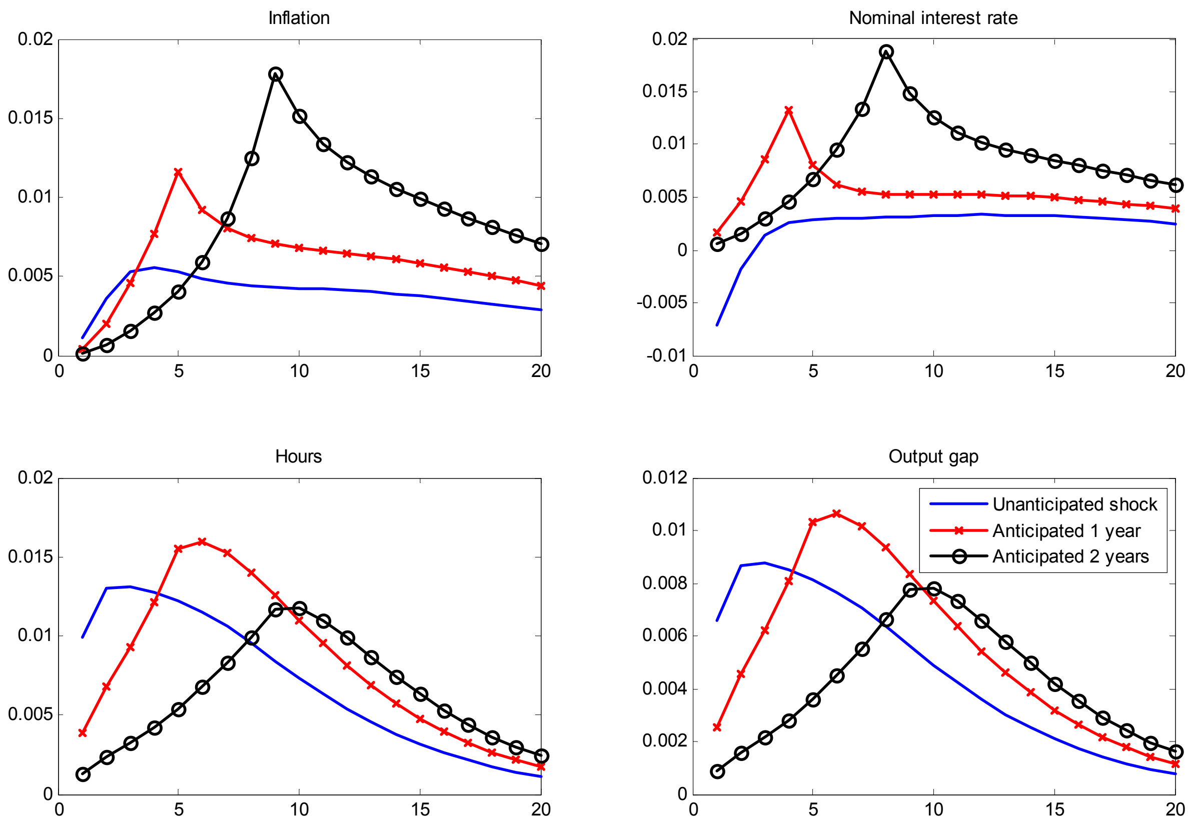

Figure 12. Impulse responses to a one-standard deviation monetary policy shock in the Euro-area announced in advance or not

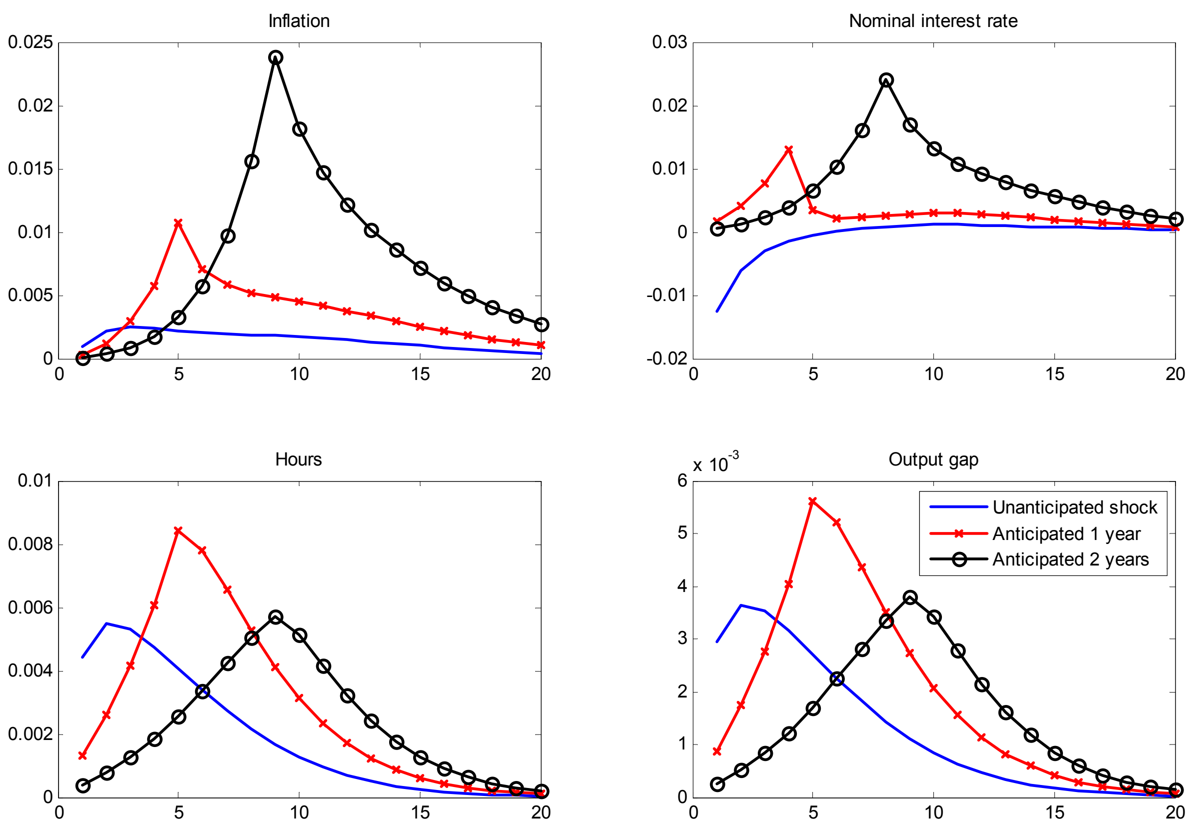

Figure 13. Impulse responses to a one-standard deviation monetary policy shock in the United States that gradually takes place or not

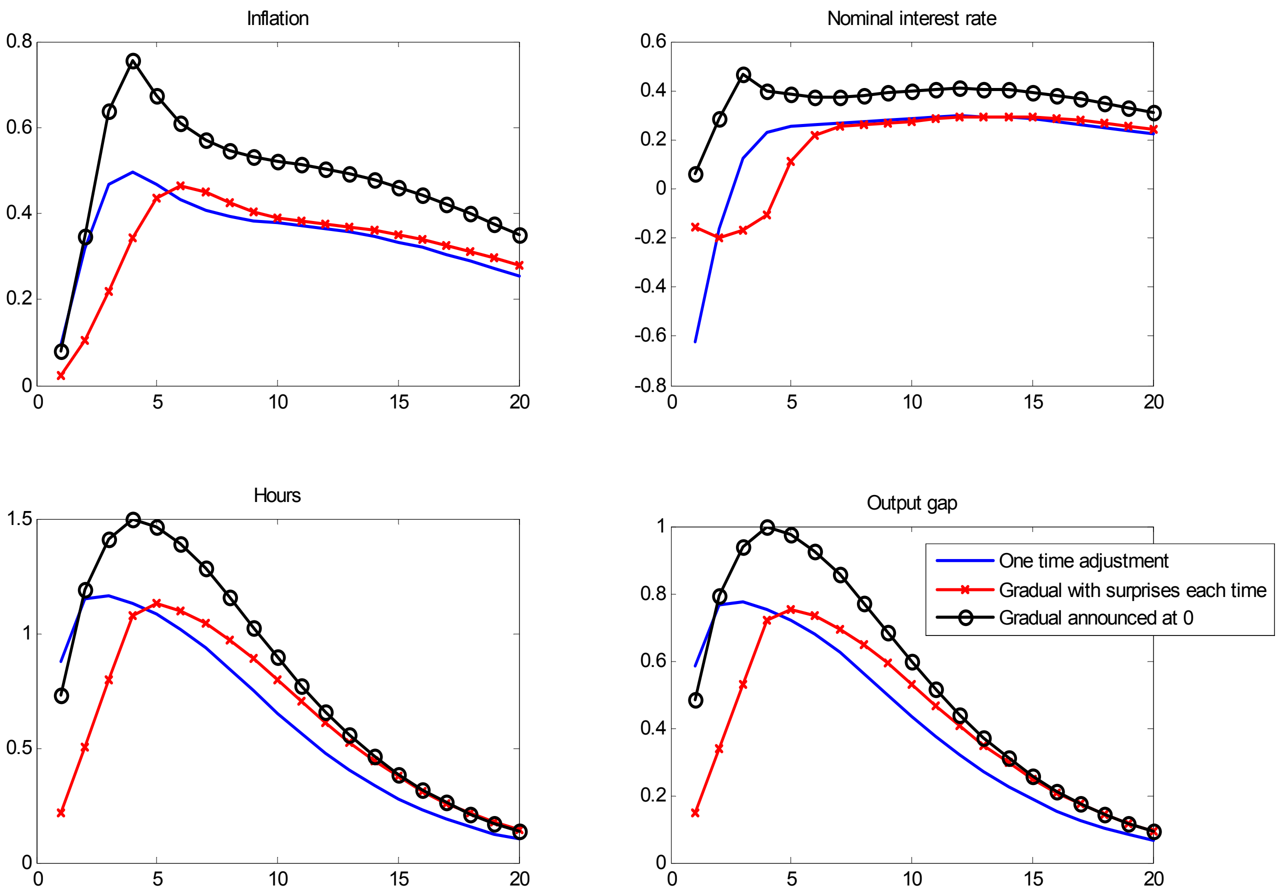

Figure 14. Impulse responses to a one-standard deviation
monetary policy shock in the Euro-area
that gradually takes place or not

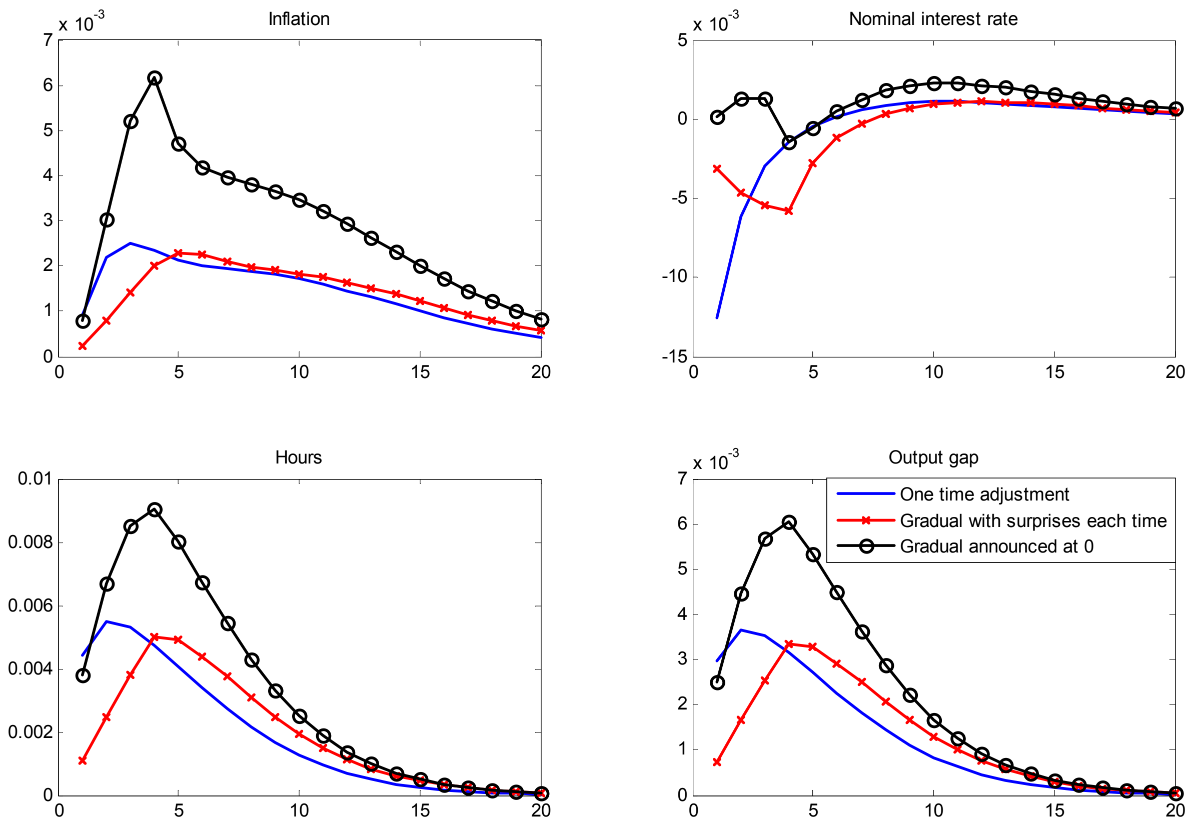

Figure 15. Density plots of welfare for different optimal policies in the United States

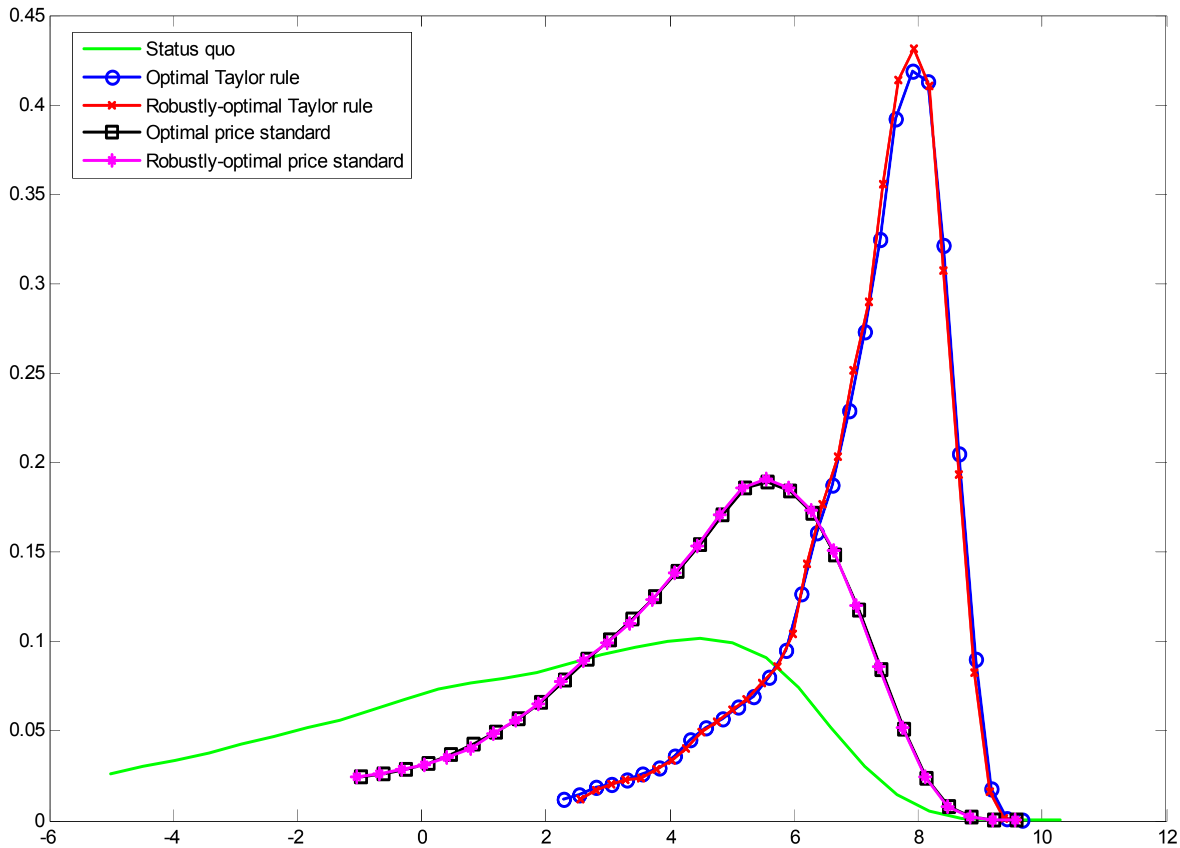

Figure 16. Density plots of welfare for different optimal policies in the Euro-area

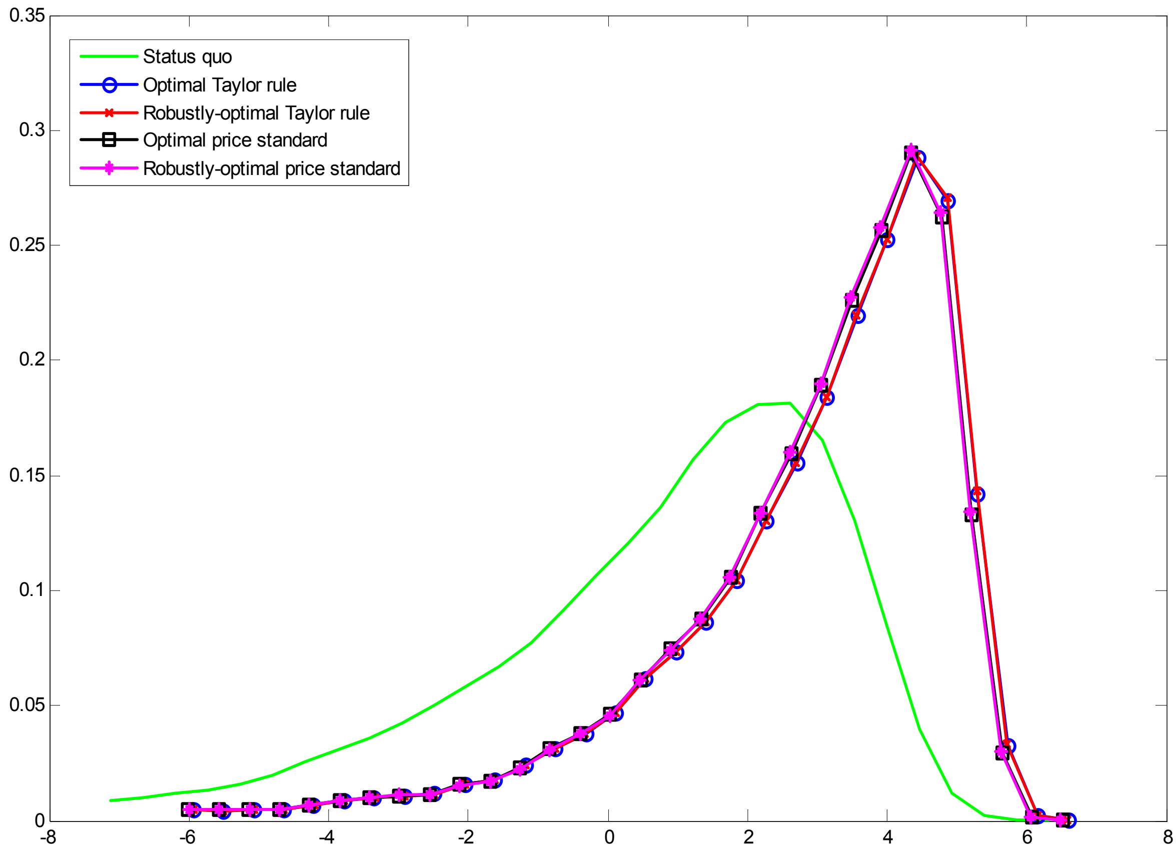

Table 1. Prior distributions

Parameters	Density	Mean	Standard Deviation	Percentiles (2.5; 50; 97.5)	Mode
Substitution elasticities					
ν	1+G	11	3.16	5.80 ; 10.67 ; 18.09	10.09
γ	1+G	11	3.16	5.80 ; 10.67 ; 18.09	10.09
Non-policy shocks					
ρ_g	B	.70	.22	.20 ; .75 ; .99	1.03
σ_g	$IG^{1/2}$	.22	.11	.11 ; .19 ; .51	.15
ρ_ν	B	.70	.22	.20 ; .75 ; .99	1.03
σ_ν	$IG^{1/2}$	.22	.11	.11 ; .19 ; .51	.15
ρ_γ	B	.70	.22	.20 ; .75 ; .99	1.03
σ_γ	$IG^{1/2}$	.22	.11	.11 ; .19 ; .51	.15
Monetary policy					
φ_y	G	.33	.25	.03 ; .27 ; .97	.14
φ_π	1+G	1.24	.25	1.00 ; 1.16 ; 1.92	1.19
ρ_ε	B	.70	.22	.20 ; .75 ; .99	1.03
σ_ε	$IG^{1/2}$	.22	.11	.11 ; .19 ; .51	.15
Inattentiveness					
δ	U	.50	.29	.03 ; .50 ; .98	[0,1]
ω	U	.50	.29	.03 ; .50 ; .98	[0,1]
λ	U	.50	.29	.03 ; .50 ; .98	[0,1]

Notes: The densities are the gamma (G), beta (B), inverse gamma (IG) and uniform (U).

Table 2. Posterior distributions for the United States

Parameters	Mean	Standard Deviation	Percentiles (2.5; 50; 97.5)	Mode
Substitution elasticities				
ν	21.83	3.64	15.21 ; 21.69 ; 29.52	20.94
γ	7.25	1.40	4.72 ; 7.18 ; 10.25	5.76
Non-policy shocks				
ρ_g	.88	.05	.76 ; .89 ; .98	.88
σ_g	.01	.002	.01 ; .01 ; .02	.01
ρ_ν	.65	.03	.60 ; .65 ; .70	.66
σ_ν	1.17	.41	.58 ; 1.10 ; 2.20	.98
ρ_γ	.59	.05	.46 ; .60 ; .67	.63
σ_γ	.33	.12	.15 ; .31 ; .61	.25
Monetary policy				
φ_y	.43	.10	.26 ; .42 ; .64	.40
φ_π	1.28	.20	1.01 ; 1.26 ; 1.72	1.14
ρ_ε	.90	.02	.85 ; 90 ; .94	.89
σ_ε	.01	.002	.01 ; .01 ; .02	.01
Inattentiveness				
δ	.17	.04	.12 ; .17 ; .26	.18
ω	.22	.03	.17 ; .22 ; .28	.22
λ	.68	.03	.63 ; .68 ; .74	.67

Notes: The moments are computed using 100,000 draws from the posterior, which come from mixing 5 independent MCMC simulated Metropolis chains of 40,000 draws discarding the first 20,000. Convergence was assessed using the Brooks-Gellman scale reduction factors and by plotting between-chain and within-chain variances.

Table 3. Variance decompositions for the United States

	Shock				
Variable	Monetary	Aggregate productivity	Aggregate demand	Goods markup	Labor markup
Inflation	.76	.05	.01	.17	.00
	[.52 ; .89]	[.03 ; .08]	[.00 ; .08]	[.07 ; .40]	[.00 ; .01]
Output growth	.16	.17	.47	.14	.05
	[.10 ; .25]	[.11 ; .23]	[.33 ; .61]	[.07 ; .25]	[.02 ; .12]
Hours	.45	.05	.36	.08	.03
	[.19 ; .69]	[.01 ; .12]	[.15 ; .69]	[.03 ; .17]	[.01 ; .10]
Interest rate	.26	.09	.02	.48	.13
	[.12 ; .44]	[.06 ; .13]	[.01 ; .08]	[.28 ; .70]	[.05 ; .25]
Wage growth	.08	.26	.01	.61	.03
	[.04 ; .14]	[.14 ; .38]	[.00 ; .03]	[.44 ; .78]	[.01 ; .07]

Notes: Median and 95% percentiles cell-by-cell using 50,000 draws of the posterior density, so rows will not add up to one exactly.

Table 4. Predictive moments for key U.S. facts

Moment	Data	Prior Mean	Posterior Mean	Posterior Percentiles (2.5; 50; 97.5)
$StDev(\pi_t)$	.01	.03	.02	.02 ; .02 ; .03
$StDev(\Delta y_t)$	.01	.08	.02	.01 ; .02 ; .02
$Corr(\pi_t , \pi_{t-1})$	.83	.94	.89	.78 ; .90 ; .96
$Corr(\pi_{t+2} - \pi_{t-2} , y_t - y_t^{trend})$	.46	.49	.31	.09 ; .31 ; .50
$StDev(\Delta(w-p)_t)/StDev(\Delta(y-l)_t)$	.70	.60	1.05	.82 ; 1.03 ; 1.39
$StDev(y_t-y_{t-1})/0.5StDev(y_t-y_{t-4})$	.79	1.19	.91	.80 ; .91 ; 1.04

Notes: Posterior moments computed using 50,000 draws from the posterior density and simulating pseudo-samples of the same length as original sample. The output trend is calculated using an HP filter with smoothing parameter 1600.

Table 5. Posterior distributions for the Euro-area

Parameters	Mean	Standard Deviation	Percentiles (2.5; 50; 97.5)	Mode
Substitution elasticities				
ν	12.32	2.98	7.27 ; 12.03 ; 18.67	9.20
γ	14.90	1.96	10.95 ; 14.92 ; 18.96	15.82
Non-policy shocks				
ρ_g	.98	.01	.95 ; .99 ; 1.00	.99
σ_g	.01	.002	.005 ; .01 ; .01	.01
ρ_ν	.46	.12	.21 ; .46 ; .67	.38
σ_ν	.07	.02	.03 ; .06 ; .13	.03
ρ_γ	.53	.10	.28 ; .56 ; .66	.55
σ_γ	.21	.09	.06 ; .20 ; .41	.17
Monetary policy				
φ_y	1.49	.37	.90 ; 1.43 ; 2.36	1.61
φ_π	1.23	.16	1.01 ; 1.21 ; 1.60	1.17
ρ_ε	.79	.05	.69 ; .80 ; .89	.88
σ_ε	.02	.01	.01 ; .02 ; .03	.01
Inattentiveness				
δ	.07	.02	.03 ; .07 ; .11	.08
ω	.66	.04	.59 ; .66 ; .72	.68
λ	.34	.05	.22 ; .34 ; .43	.26

Notes: Same as table 2, but now using 900,000 draws from 5 independent chains of 200,000 draws each, and discarding the first 20,000 draws.

Table 6. Variance decompositions for the Euro-area

	Shock				
Variable	Monetary	Aggregate productivity	Aggregate demand	Goods markup	Labor markup
Inflation	.08	.78	.09	.00	.00
	[.01 ; .53]	[.42 ; .95]	[.01 ; .29]	[.00 ; .00]	[.00 ; .00]
Output growth	.16	.14	.63	.00	.03
	[.04 ; .44]	[.05 ; .26]	[.35 ; .88]	[.00 ; .03]	[.00 ; .14]
Hours	.06	.05	.87	.00	.00
	[.01 ; .36]	[.01 ; .19]	[.52 ; .98]	[.00 ; .01]	[.00 ; 04]
Interest rate	.24	.47	.06	.02	.12
	[.05 ; .60]	[.19 ; .85]	[.02 ; .19]	[.00 ; .10]	[.01 ; .44]
Wage growth	.47	.40	.07	.02	.00
	[.24 ; .77]	[.19 ; .58]	[.02 ; .21]	[.00 ; .08]	[.00 ; .04]

Notes: Same as table 3.

Table 7. Predictive moments for key E.U. facts

Moment	Data	At Prior Mean	Posterior Mean	Posterior Percentiles (2.5; 50; 97.5)
StDev(π_t)	.002	.03	.02	.005; .01 ; .05
StDev(Δy_t)	.004	.08	.01	.005 ; .01 ; .01
Corr(π_t , π_{t-1})	.39	.92	.98	.91 ; .98 ; .1.00
Corr(π_{t+2} - π_{t-2} , y_t - y_t^{trend})	.36	.46	.15	-.39 ; .16 ; .63
StDev($\Delta(w\text{-}p)_t$)/StDev($\Delta(y\text{-}l)_t$)	1.22	.59	2.51	1.46 ; 2.39 ; 4.28
StDev(y_t-y_{t-1})/0.5StDev(y_t-y_{t-4})	.64	1.23	.99	.75 ; .97 ; 1.34

Notes: Same as table 4.

Table 8. Optimal policy and welfare

Policy rule	Parameters	Welfare gain from status quo (in % of consumption)
United States		
Eliminate policy errors	1.28 ; 0.43	5.00
Optimal Taylor rule	1.62, 1.81	5.54
Elastic price standard	0.95	2.60
Unconditional optimum		6.29
Euro-area		
Eliminate policy errors	1.22 ; 1.48	1.44
Optimal Taylor rule	19.00 ; 0.54	2.04
Elastic price standard	0.01	1.95
Unconditional optimum		2.70

Notes: Numbers in the last column are in percentage units.

Table 9. Robustly optimal policy and expected welfare

Policy rule	Optimal Parameters	Robust Parameters	Expected welfare gain 1	Expected welfare gain 2
United States				
Optimal Taylor rule	1.62 ; 1.81	2.31 ; 1.98	0.02	6.89
Elastic price standard	0.95	0.67	0.01	3.69
Euro-area				
Optimal Taylor rule	19.00 ; 0.54	18.25 ; 0.68	0.00	2.60
Elastic price standard	0.01	0.00	0.00	2.39

Notes: Expected welfare gain 1 has the gain in expected welfare from using robustly optimal parameters rather than optimal parameters. Welfare gain 2 has expected welfare gain relative to status quo. The numbers in the last two columns are in percentage units.

Table A.1. Posterior distributions for the Euro-area usiong data from 1979:3

Parameters	Mean	Standard Deviation	Percentiles (2.5; 50; 97.5)	Mode
Substitution elasticities				
ν	9.34	2.63	5.11 ; 9.04 ; 15.41	6.97
γ	9.50	2.12	5.85 ; 9.32 ; 14.15	8.88
Non-policy shocks				
ρ_g	.95	.04	.86 ; .96 ; 1.00	.94
σ_g	.01	.001	.004 ; .007; .01	.01
ρ_ν	.29	.06	.17 ; .30 ; .41	.32
σ_ν	.57	.37	.14 ; .47 ; 1.56	.17
ρ_γ	.92	.03	.84 ; .92 ; .97	.88
σ_γ	.68	.38	.26 ; .58 ; 1.76	.47
Monetary policy				
φ_y	.20	.07	.08 ; .20 ; .33	.28
φ_π	1.19	.10	1.05 ; 1.18 ; 1.42	1.22
ρ_ε	.23	.08	.07 ; .23 ; .39	.17
σ_ε	.01	.008	.01 ; .01 ; .02	.01
Inattentiveness				
δ	.01	.01	.003 ; .01 ; .04	.02
ω	.95	.04	.86 ; .95 ; 1.00	.98
λ	.28	.12	.16 ; .23 ; .67	.42

Notes: Same as table 5.