Examining Fisher Information in Multi-Sector DGE Models

Jaromir Benes*
Reserve Bank of New Zealand

David Vavra
International Monetary Fund and National Bank of Serbia

This version: August 2006

Abstract

Dynamic general equilibrium models and their deep parameters may, for various reasons, feature lack of identification or other deficiencies related to the objective function (such as the likelihood or a distance function) used in their estimation or empirical validation. These deficiencies can be often an inherent property of a particular model structure, independent of observed data, and modellers should be aware of them (and their implications) when designing the experiment. Routinely performing, checking, and reporting several basic diagnostics of the objective function, such as its surface and Hessian, *prior* to any estimation procedure or empirical validation has been suggested by Canova and Sala (2006). We illustrate this modelling strategy by examining local identification of a simple DGE model in a full information likelihood based framework focusing mainly on the Fisher information matrix, and provide some computational details for a class of state-space representations with multiple unit roots. These can be seen as a convenient and practical modelling device, especially for analysing multi-sector models.

1 INTRODUCTION AND CONCLUSIONS

Suprisingly little attention has been paid thus far in the applied dynamic general equilibrium modelling literature to the issue of the identifiability of deep structural parameters as an inherent feature of a given model structure[1] *independent* of particular sample data observed. Only recently Canova and Sala (2006) suggested in their–potentially–pathbreaking paper that a proper investigation of various aspects of the objective function that underlies the subsequent estimation step (typically distance or likelihood functions) should *precede* any analysis based on real data. This is because such an exercise can, a priori, detect some of identification problems (such as underidentification or weak identification) embedded in a particular model structure itself, help to distinguish them from problems arising because of some particular properties of observed

*Corresponding author.
[1] Including the desired identification scheme, such as the choice of observables or suitable reparameterizations.

data, and lead ultimately to the most appropriate choice of the experiment design. This claim is certainly valid in both classical and Bayesian econometrics across virtually all limited and full information methods.

Canova and Sala illustrate their points within a framework of a limited information method based on minimising the distance between an econometrically estimated impulse response function and its model-implied counterpart. For a simple forward-looking model with a relatively small number of deep parameters they inspect various aspects of the curvature of their objective function. They detect its potential deficiencies and, therefore, lack of identification prior to estimating the model, and infer basic implications.

We perform a similar analysis in a full information likelihood based framework examining the Rothenberg (1971) local identifiability of a simple multi-sector DGE model. To this end, we focus mainly on the properties of the Fisher, or expected, information matrix, which provides an upper limit for information that can be obtained from observed data for inferences about a model's parameters. In this context, we suggest the following steps:

(i) Formal tests of the rank of the information matrix. These tests are, in general, helpful for detecting rank deficiencies but obviously cannot identify properly their sources.

(ii) Examination of various subblocks of the information matrix. This procedure may, in practice, well correspond to fixing certain parameters, or to simplifying the model structure[2] and provide basic information on the potential sources of identification problems;

(iii) Eyeball checks of the behaviour of the expected likelihood function in a neighbourhood of a given point in the parameter space by plotting its surface along two or three (contour plot) dimensions at a time.

(iv) Implications of different choices of observables (i.e. the composition of the vector of observables), constrained e.g. by the observability condition.

(v) Examination of various sample lengths. Such exercises are often used to a priori determine an appropriate sample size for achieving a desired level of accuracy in the experiment design in various fields, as e.g. in Dharan (1985);

(vi) Implications of different choices of structural innovations and/or measurement innovations introduced into the model, again constrained again e.g. by the stochastic non-singularity condition.

It is clear from Sections 2 and 3 that these diagnostics are computationally easy and cheap, at least relative to subsequent estimation procedures, and have the power to provide extra insight into the properties and, more importantly, general limitations of a model. It seems therefore natural that such diagnostics should be performed, checked and report ed routinely, and could become best practice in applied modelling.

As part of our exposition we describe the computation of Fisher information matrices for a class of state space representations that feature an unlimited number of unit roots. We show that this class is especially convenient for multi-sector DGE models characterised by a number of stochastic trends driving the economy in the long run, and hence by different rates of growth for different variables, or sectors, along a balanced growth path. We claim that these models, when properly treated, need not be stationarised and their first-order accurate policy functions may retain multiple unit roots. This can, in certain cases, considerably simplify the model analysis, estimation or empirical validation, and the mapping between model variables and their observed counterparts.

The paper is organised as follows. We review the design and estimation of the Fisher informa-

[2] Such as setting elasticiticies of substitution to one corresponds to replacing general CES functions with their Cobb-Douglas counterparts.

tion matrix in Section 2, and address some computational related to a class of state-space representations suitable for multi-sector DGE models in Section 3. The theory and basic diagnostic chechks are applied to a simple DGE model in Section 4. Two appendices provide more technical details: In Appendix A we outline the model used in Section 4; in Appendix B we construct a quasi-triangular state-space representation for DGE models with mutliple unit roots.

2 THE FISHER INFORMATION MATRIX

In this section we review the design of, and propose a simple estimation procedure for, the Fisher information matrix for Gaussian state-space models in the time domain, and discuss several issues related to the computational efficiency, and to the presence of unit roots.

The Fisher information matrix (FIM) provides by the virtue of the Cramér-Rao inequality an upper limit for the amount of information that is potentially available from observed data for inferences about parameters. It is usually defined as the sampling variance-covariance of the score vector of the log-likelihood function,[3] i.e.

$$\mathscr{I}(\theta,T) = \mathrm{E}_{Y|\theta}\left[s(Y|\,\theta)\cdot s(Y|\,\theta)'\right] = \mathrm{E}_{Y|\theta}\left[\frac{\partial \log \mathscr{L}(Y|\,\theta)}{\partial\theta}\cdot\frac{\partial \log \mathscr{L}(Y|\,\theta)}{\partial\theta'}\right],$$

which is evaluated at the true parameter vector, θ, by integrating over all possible realizations of the sample Y of a given length T. Furthermore, as we restrict ourselves to Gaussian models for which the usual regularity conditions hold,[4] the information matrix equality applies, and the FIM can be therefore computed as the sampling expectation of minus the Hessian,

$$\mathscr{I}(\theta,T) = -\mathrm{E}_{Y|\theta}\,\mathscr{H}(Y|\theta) = -\mathrm{E}_{Y|\theta}\,\frac{\partial^2 \log \mathscr{L}(Y|\,\theta)}{\partial\theta\,\partial\theta'}.$$

Consequently, our computation and estimation of the FIM is based on simple Monte Carlo integration the design of which follows here from the law of large numbers and the law of iterated expectations,

$$\widehat{\mathscr{I}}(\theta,T) = \frac{1}{N}\sum_{i=1}^{N} -\mathrm{E}\,\mathscr{H}(Y^i|\theta),$$

where Y^i, $i=1,\ldots,N$ is the i-th random sample of a length T drawn from the joint distribution of observables given by the model for a particular parameterization θ, and $-\mathrm{E}\,\mathscr{H}(Y^i|\theta)$ is the expectation of minus the correspoding sample Hessian,[5] which is also called the sample information matrix. The computation of the expected Hessian, and hence of the FIM estimate, is described in more detail in Section 3; we now turn to the formal testing of the estimated FIM and its rank. Formal tests can be employed as an efficient way to detect potential deficiencies of the FIM and its subblocks, as suggested in the introduction.

There have been developed a number of formal tests of the rank of stochastic matrices, such as Anderson (1951), Cragg and Donald (1996), Cragg and Donald (1997), or Ratsimalahelo (2002). Many of these tests have been generalised by Kleibergen and Paap (2006) in a statistic based on the singular value decomposition. However, their test statistic rk_q is derived under the assumption that the examined matrix is unrestricted. More specifically, the matrix estimator is assumed have the following limiting behaviour:

$$\sqrt{N}\left[\widehat{\mathscr{I}}(\theta,T)-\mathscr{I}(\theta,T)\right] \xrightarrow{d} \mathscr{N}(0,\Sigma),$$

which is not true in the present context, as the symmetry of the information matrix (or Hessians in general) imposes restrictions on the asymptotic covariance matrix, Σ. As a consequence, the test is, strictly speaking, inapt for information matrices or Hessians; however, the effect of symmetry on the limiting distribution of the test

[3]The sampling expectation of the score vector is obviously zero.

[4]See econometrics textbooks for the regularity conditions, e.g. Green (2003), Definition 17.3.

[5]Evaluating the expected Hessian is more efficient than the Hessian itself in the class of Gaussian models considered here.

[6]This shortcoming is present in other tests, too.

statistic has not been analyzed yet in the literature, and we stick to the use of the test despite this reservation. [6]

The null of the test is that the rank of an $n \times n$ matrix is $q < n$, or that the q smallest singular values are indistinguishable from zero. The rk_q statistic converges under the null to a χ^2 distribution with $(n-q)^2$ degrees of freedom,

$$\mathrm{rk}_q \xrightarrow{d} \chi^2_{(n-q)^2},$$

and is constructed as follows. We first perform the singluar values decomposition of the estimated FIM, $\widehat{\mathscr{I}}(\theta, T) = U \cdot \Lambda \cdot V'$, so that $U \cdot U' = \mathrm{I}_n$, $V \cdot V' = \mathrm{I}_n$, and Λ contains singular values on the main diagonal and zeros elsewhere, and partition these conformably with the number of tested singular values, $U = [U_1, U_2]$, $V' = [V_1', V_2']$, and $S = \mathrm{diag}[\Lambda_1, \Lambda_2]$, with U_2 and V_2 being both $n \times q$, and Λ_2 being $q \times q$. Then

$$\mathrm{rk}_q = N \cdot \lambda_q' \Omega^{-1} \lambda_q,$$

where

$$\lambda_q = \mathrm{vec}\,\Lambda_2,$$
$$\Omega = (V_2' \otimes U_2')\,\Sigma\,(V_2 \otimes U_2).$$

Finally, the asymptotic covariance matrix of the estimator, Σ, can be replaced with its consistent estimate which is

$$\widehat{\Sigma} = \frac{1}{N} \sum_{i=1}^{N} \left[-\mathrm{E}\mathscr{H}(Y^i|\theta) - \widehat{\mathscr{I}}(\theta, T) \right]^2.$$

3 COMPUTATIONAL ISSUES

We now describe the computation of the FIM[7] for the following class of time-invariant Gaussian state-space models:

$$y_t = Z(\theta)\,\alpha_t + \delta(\theta) u_t + H(\theta)\,\varepsilon_t, \tag{1a}$$
$$\alpha_t = T(\theta)\,\alpha_{t-1} + k(\theta) + R(\theta)\,\varepsilon_t, \tag{1b}$$

where[8]

(i) y_t, α_t, and u_t are vectors of observables, state variables, and regressors, respectively;

(ii) the state vector, α_t, is a rotation of the underlying vector of predetermined model variables, $\alpha_t = U x_t^p$;

(iii) the vector of regressors, u_t, is deterministic;

(iv) the structural innovations are Gaussian, $\varepsilon_t \sim N(0, \Omega)$, with Ω diagonal;

(v) the effect of innovations on the measurement equation, $H\varepsilon_t$, is uncorrelated the effect on the transition equation, $R\varepsilon_t$, i.e. $HR' = 0$.

(vi) the transition matrix, T, is real and quasi-triangular (in that it has 1×1 or 2×2 blocks along the main diagonal, and zeros below) and its eigenvalues lie inside and/or on the unit circle; in particular, we assume that T takes the following form:

$$T = \begin{bmatrix} I & T_1 \\ 0 & T_2 \end{bmatrix},$$

where I is an $n \times n$ identity matrix with n being the number of stochastic trends driving the model, and T_2 is obviously quasi-triangular, too, with eigenvalues within the unit circle only.

As a consequence of assumption (vi) we split the state vector into an upper, non-stationary, part and a lower, stationary, one,

$$\alpha_t = \begin{bmatrix} \alpha_{1t} \\ \alpha_{2t} \end{bmatrix}.$$

The particular representation (1a)-(1b) with assumptions (i)–(vi) is by no means restrictive in terms of some underlying DGE model: there are infinitely many state-space representations of a particular first-order accurate policy function (distinguished by the rotations of the vector of predetermined variables), and we choose the one

[7] The algorithms described here have been implemented in the IRIS Toolbox, an economic modelling package for Matlab developed by the authors of this paper, and available from www.iris-toolbox.com. The codes for the application found in Section 4 are available upon request from the corresponding author.

[8] For notational convenience we drop the obvious dependence on θ of system matrices, Z, δ, H, T, k, and R, henceforth.

with the matrix T featuring the desired properties. Furthermore, we allow for any number of unit root processes to drive the dynamics of the model, without the necessity of non-stationary variables being stationarised by suitable transformations (e.g. scaling by underlying technology processes). Appendix B describes the derivation of the representation (1a)-(1b) in models with multiple stochastic trends (multi-sector models).

There are at least two major advantages of the above representation over any other one (expecially over the one where the state vector is identical with the vector of predetermined variables). First, we decouple all unit root processes so that they are represented by the first n elements of α_t, and share an empty cointegrating space; the latter feature is extremely convenient when dealing with the initial condition for the upper (non-stationary) part of α_t in the Kalman filter. Second, the quasi-triangularity of T_2 permits us to compute the unconditional covariance matrix of the lower (stationary) part of α_t efficiently by successively solving the individual rows of the underlying Lyapunov equation.

Next, the explicit inclusion of a deterministic vector u_t in the state-space representation is motivated by the option of concentrating out of the transition equation some of the structural parameters (typically those determining the long run non-stochastic growth characteristics, such as the growth rates, inflation rates, or interest rates) and including them into the matrix δ. Parameters in δ can be then concentrated out of the ordinary likelihood function, and their maximum likelihood estimate turns out to be a linear function of observables given all other parameters in θ. As this have the power to improve computational efficiency in the estimation we assume that the researcher explores such linearities; an example of such an estimation strategy is given in Section 4.

To compute the log-likelihood function of (1a)–(1b) and its derivatives on a particular sample of observables, $Y = \{y_1, \ldots, y_T\}$, we cast the prediction error decomposition method making use of a sequence of one-step-ahead prediction errors, ν_t, and their MSE matrices, F_t (which are products of a Kalman filter run),

$$\begin{aligned} \nu_t &= y_t - y_{t|t-1}, \\ F_t &= \mathrm{E}(y_t - y_{t|t-1})^2, \end{aligned}$$

where $y_{t|t-1}$ denotes the conditional forecast (expectation) of the vector of observables based upon time $t-1$ information. Since the prediction errors are by construction uncorrelated the log-likelihood function for a particular sample can shown to be

$$\log \mathscr{L}(Y|\theta) \propto -\tfrac{1}{2}\textstyle\sum_{t=1}^{T}\left(\log|F_t| + \nu_t' F_t^{-1} \nu_t\right).$$

More importantly, the sample score vector and the sample information matrix can be both computed using *first derivatives* only.[9] Namely, their i-th and (i,j)-th entries, respectively, are

$$\begin{aligned} s_i(Y|\theta) &= -\tfrac{1}{2}\textstyle\sum_{t=1}^{T}\big[\nu_{ti}' F_t^{-1}\nu_t + \\ &\qquad \mathrm{tr}\left(F_t^{-1}F_{ti} - F_t^{-1}F_{ti}F_t^{-1}\nu_t\nu_t'\right)\big], \\ -\mathrm{E}\,\mathscr{H}_{ij}(Y|\theta) &= -\tfrac{1}{2}\textstyle\sum_{t=1}^{T}\big[2\,\nu_{ti}' F_t^{-1}\nu_{tj} + \\ &\qquad \mathrm{tr}(F_t^{-1}F_{ti}F_t^{-1}F_{tj})\big], \end{aligned}$$

where $F_{ti} = \partial F_t / \partial \theta_i$ and $\nu_{ti} = \partial \nu_t / \partial \theta_i$. Note that F_t and F_t^{-1} are independent of a particular realization of observables, and hence can be stored and reused for all N draws.

To perform the prediction error decomposition we initialise the Kalman filter with the unconditional mean and covariance matrix for the lower, stationary, part of the state vector, $E\alpha_{2t}$ and $\mathrm{E}(\alpha_{2t} - \mathrm{E}\,\alpha_{2t})(\alpha_{2t} - \mathrm{E}\,\alpha_{2t})'$, respectively. These unconditonal moments are again independent of observables, and can be stored and reused. Next, the initial condition for the upper, non-stationary, part of the state vector is in our experiment design treated as unknown fixed

[9] See e.g. Harvey (1989) for the derivation of both.

[10] This assumption is to reduce the risk of stochastic singularity (in that the forecast MSE might result into zero for some of the non-stationary elements of the state vector at the initial time $t = 1$) for some model setups.

[11] Also reproduced in Harvey (1989).

quantities that are to be estimated. More specifically, we define the initial condition to be the value of the state vector at the last pre-sample time, $t = 0$,[10] and use an algorithm due to Rosenberg (1973).[11] Fixed initial conditions, α_{20}, are treated as extra parameters, and estimated by maximising the likelihood function. However, for any given θ, these initial conditions can be again concentrated out of the ordinary likelihood function and made a linear function of observables. Conditions under which α_{20} is identified are discussed in Appendix B. They, in general, relate to the observability of the system.

4 APPLICATION TO A SIMPLE MULTI-SECTOR MODEL

The model examined in this section is meant to describe the basic medium-term determination of consumer inflation in a small open economy that features (i) a permanent trend in the relative price of tradeables and non-tradeables, (ii) permanent shocks to the terms of trade, and (iii) an export production sector relatively disconnected from domestic consumption. Examples of such economies may include commodity exporters, or industralised countries specialised in agricultural exports.

The economy is driven by five (uncointegrated) stochastic trends: a production process in the export sector, a technology processes in the non-tradeable sector, the terms of trade, a domestic monetary (or nominal) trend, and a foreign monetary (or nominal) trend.

[To be continued.]

5 CONCLUSIONS

[To be continued.]

REFERENCES

Anderson, T. W. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions. *Annals of Mathematical Statistics* **22**(3), 327–351.

Calvo, Guillermo A. (1983). Staggered Prices in a Utility-Maximizing Framework. *Journal of Monetary Economics* **12**(3), 383–398.

Canova, Fabio and Luca Sala (2006). Back to Square One: Identification Issues in DSGE Models. *European Central Bank Working Paper*, 583.

Christiano, Lawrence J., Martin Eichenbaum and Charles L. Evans (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* **118**(1), 1–45.

Cragg, John G. and Stephen G. Donald (1996). On the Asymptotic Properties of LDU-Based Tests of the Rank of a Matrix. *Journal of the American Statistical Association* **91**(435), 1301–1309.

Cragg, John G. and Stephen G. Donald (1997). Inferring the Rank of a Matrix.. *Journal of Econometrics* **76**(1), 223–250.

De Gregorio, Jose and Holger C. Wolf (1994). Terms of Trade, Productivity, and the Real Exchange Rate. *NBER Working Paper*, 4807.

Dharan, Bala G. (1985). A Priori Sample Size Evaluation and Information Matrix Computation for Time Series Models. *Journal of Statistical Computation and Simulation* **21**(2), 171–177.

Edwards, Sebastian and Carlos A. Végh (1997). Banks and Macroeconomics Disturbances under Predetermined Exchange Rate. *Journal of Monetary Economics* **40**(2), 239–278.

Green, William H. (2003). *Econometric Analysis. 5th Edition*. Prentice-Hall.

Harvey, Andrew C. (1989). *Forecasting, structural time series models, and the Kalman filter*. Cambridge University Press.

Kleibergen, Frank and Richard Paap (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics* **133**(1), 97–126.

Ratsimalahelo, Zaka (2002). Rank test based on matrix perturbation theory. Unpublished mimeo. University of Franche-Comté, Besançon.

Ravn, Morten, Stephanie Schmitt-Grohé and Martín Uribe (2006). Deep Habits. *Review of Economic Studies* **73**(1), 195–218.

Rosenberg, Barr (1973). Random Coefficient Models: The Analysis of a Cross Section of Time Series by Stochastically Covergent Parameter Regression. *Annals of Economic and Social Measurement* **2**, 399–428.

Rotemberg, Julio J. (1982*a*). Monopolistic Price Adjustment and Aggregate Output. *Review of Economic Studies* **49**(4), 517–531.

Rotemberg, Julio J. (1982*b*). Sticky Prices in the United States. *Journal of Political Economy* **90**(6), 1187–1211.

Rotemberg, Julio J. and Michael Woodford (1999). The Cyclical Behavior of Prices and Costs. NBER Working Paper Series, 6909.

Rothenberg, Thomas J. (1971). Identification in Parametric Models. *Econometrica* **39**(3), 577–591.

Schmitt-Grohé, Stephanie and Martín Uribe (2003). Closing small open economy models. *Journal of International Economics* **61**(1), 163–185.

A THE MODEL

A small open economy consists of three sectors: domestic producers of non-tradeables, importers of tradeables, and producers of export goods. There is no domestic consumption of export goods nor there is domestic production of tradeables [a similar structure was examined e.g. by De Gregorio and Wolf (1994) in the context of the analysis of the terms of trade, productivity differentials, and real appreciation]. The exports and their foreign-currency price are both assumed to be exogenous, affecting thus the domestic consumption cycle only through the income effect in the intertemporal budget constraint.

There is a permanent trend in the relative price of tradeables and non-tradeables due to a differential between productivity growth in the non-tradeable sector and the real growth of exports. Furthermore, we allow for permanent shocks to drive the terms of trade. To achieve a steady state with stable nominal expenditure shares but permanent wedges in relative prices and respective real quantities we need to impose unit elasticity of substitution between the two goods in domestic consumption in the long run. However, to produce realistic responses of the model at business-cycle frequencies we need to reduce the short-run elasticity to considerably lower values. This is achieved by deep habit in the consumers' utility function. Finally, we close the small open economy model with a costly banking assumption to achieve a unique distribution of wealth and consumption along a balanced growth path.

In our notational conventions, uppercase letters are usually reserved for quantities supplied whereas lowercase ones for quantities demanded.

A.1 Technology

The domestic tradeable and non-tradeable markets are monopolistically competitive, with a continuum of differentiated producers existing in each of them. The *i*-th non-tradeable producer has the following labor-intensive technology to produce the *i*-th variety of non-tradeable goods:

$$C^n_{jt} = (n_{jt} - n_0)^\gamma \cdot \exp a^n_t,$$

where $\gamma \in (0, 1]$, $n_0 > 0$ is overhead labor such that $n_0/n_j = 1-\gamma$ in the steady state,[12] and a^n_t is

[12] See e.g. Rotemberg and Woodford (1999) for the motivation of such a calibration assumption.

a sector-wide unit-root technology process, $a_t^n = a_{t-1}^n + \alpha_n + \varepsilon_t^{an}$, with ε_t^{an} being a non-tradeable technology (or production frontier) shifter. The labor input, n_{jt}, is a CES index defined over the variety of labor skills supplied by the continuum of the representative household's members, with the elasticity of substitution ε. Furthermore, the wage bill needs to be paid in advance, and is therefore financed loans, $\ell_{jt} = W_t n_{jt}$, available from a financial intermediary at the lending rate i_t^ℓ.

The price and output choice of the producer is constrained by a downward-sloping demand curve, which derives from the household's optimization, $C_{jt}^n = (p_{jt}^n / P_t^n)^{-\varepsilon} c_t^n$, and by the assumption of quadratic cost associated with deviations of the producer's own inflation from the last observed aggregate rate of price changes,

$$\xi_n \cdot \left(\Delta \log p_{jt}^n - \Delta \log p_{t-1}^n - \varepsilon_t^{pn}\right)^2,$$

where ε_t^{pn} is a non-tradeable cost-push shifter. This is a modified version of the assumption first introduced by Rotemberg (1982*a*) and Rotemberg (1982*b*). The modification allows for a non-zero steady-state rate of inflation while preserving monetary super-neutrality in the long run, and is a counterpart to the Calvo (1983) and Christiano *et al.* (2005) pricing used also frequently in sticky-price models.

The *i*-th tradeable producer purchases import goods, m_{jt}, from abroad at the cost of p_t^{m*}/s_t, and transform them costlessly into the *i*-th type of domestic tradeables.

$$C_{jt}^\tau = m_{jt} \cdot \exp \varepsilon_t^{a\tau},$$

where $\varepsilon_t^{a\tau}$ is a tradeable production frontier shifter. Again, the producer is constrained by the demand curve, $C_{jt}^\tau = (p_{jt}^\tau / P_t^\tau)^{-\varepsilon} c_t^\tau$, and by the quadratic cost of price adjustments with a tradeable cost-push shifter, $\varepsilon_t^{p\tau}$,

$$\xi_\tau \cdot \left(\Delta \log p_{jt}^\tau - \Delta \log p_{t-1}^\tau - \varepsilon_t^{p\tau}\right)^2,$$

Import inflation follows a stationary process,

$$\Delta \log p_t^{m*} = \rho_{pm} \Delta \log p_{t-1}^{m*} + (1 - \rho_{pm})\pi^* + \varepsilon_t^{pm}.$$

Finally, there is a single exporter whose production is a unit-root proces affected by production frontier shifters, with no input factors involved,

$$\log x_t = \log x_{t-1} + \alpha_x + \varepsilon_t^{ax}.$$

The foreing-currency price of these exports is also exogenous: for analytical convenience we instead express, without loss of generality, the process for the country's terms of trade, $T_t = p_t^{x*}/p_t^{m*}$,

$$\log T_t = \log T_{t-1} + \varepsilon_t^{tot},$$

with a permanent terms-of-trade shifter, ε_t^{tot}.

A.2 Preferences

There is a single household that consists of a continuum of memebers indexed on the interval $[0, 1]$; each member is endowed with differentiated labor skills. The household as a whole decides on the consumption plan taking total labor income and labor effort as given, whereas the individual members choose the hours worked and the wage rate taking consumption as given. The household's expected lifetime utility,

$$\mathrm{E}_t \sum_{k=0}^{\infty} \beta^k \left[\log \Gamma_{t+k} - \kappa \cdot \frac{(\int_0^1 N_{jt+k} \, \mathrm{d}j)^{1+1/\eta}}{1 + 1/\eta} \right],$$

where Γ_t is a real consumption index defined later, $\beta \in (0, 1)$ is a discount factor, $\kappa > 1$ is a scale factor, and $\eta \geq 0$ is the wage elasticity of the labor supply schedule, is maximised subject to a budget constraint,

$$d_t + \int_0^1 (p_{jt}^\tau c_{jt}^\tau + p_{jt}^n c_{jt}^n) \, \mathrm{d}j = (1 + i_{t-1}^d) d_{t-1} + \int_0^1 W_t N_{jt} \, \mathrm{d}j + \Psi_t,$$

where d_t are deposits held with the financial intermediary and bearing the deposit rate, i_t^d, and Ψ_t are net cash flows collected from domestic firms owned by the household.

The real consumption index, Γ_t, is defined over the consumed quantities of individual consumption goods, and accounts for deep external habit formed at the level of the total consumption of tradeables and non-tradeables; this a modified version of the deep habit hypothesis proposed originally by Ravn *et al.* (2006). In particular,

$$\Gamma_t = \tfrac{1}{1-\chi}\,(c_t^\tau - \chi\, h_t^\tau)^\omega (c_t^n - \chi\, h_t^n)^{1-\omega},$$

where $\chi \in [0,1)$ is the importance of habit in utility, c_t^τ and c_t^n are CES consumption indeces related to tradeables and non-tradeables, respectively, defined over the continue of these goods with an elasticity of substitution ε, and h_t^τ and h_t^n are the respective habit reference levels the dynamics of which is linked to past consumption[13] but fully externalised from the household's optimization,

$$h_t^\tau = c_{t-1}^\tau \exp(\alpha_x + \varepsilon_t^{ch}),$$
$$h_t^n = c_{t-1}^n \exp(\alpha_n + \varepsilon_t^{ch}).$$

The two reference levels are exposed to a common habit shifter, ε_t^{ch}.

Finally, the wage setting at the level of individual members of the household is subjected to a quadratic cost of wage adjustments (expressed in utility equivalents) analogous to that in the price setting:

$$\xi_w \cdot \left(\Delta \log w_{jt} - \Delta \log w_{t-1} - \varepsilon_t^{ls}\right)^2,$$

where ε_t^{ls} is an aggregate labor supply shifter.

A.3 Financial Intermediation

Out financial intermediation is similar to the costly banking sector in Edwards and Végh (1997). Financial flows between the household (d_t) and non-tradeable producers (ℓ_t) are intermediated by an agency that is risk-neutral and behaves competitively in both deposit and lending markets. The agency maximises profits given by

$$D_t(i_t - i_t^d) + L_t(i_t^\ell - i_t) - f(D_t, L_t),$$

where D_t and L_t are deposits from the household and loans granted to producers, respectively, and $f(D_t, L_t)$ is a strictly increasing, strictly convex and linearly homogeneous cost function associated with intermediation. Costly banking activities have the capacity to close a small open economy model in yet another way than those surveyed by Schmitt-Grohé and Uribe (2003).

The intermediary have can refinance from the domestic interbank market (i.e. from the central bank) or form the international financial market; the net position of the economy is then given by the discrepancy between deposits and loans, $D_t - L_t$. Furthermore, the international financial market consists of a continuum of other risk-neutral national intermediaries: hence, the uncovered interest parity is assumed to hold in the form,

$$1 + i_t = (1 + i_t^*) \cdot \mathrm{E}_t(S_t/S_{t+1}) \cdot \exp u_t,$$

where u_t is an autonomous disparity evolving as $u_t = \rho_u u_{t-1} + \varepsilon_t^{fx}$, driven by foreign exchange market shifters, ε_t^{fx}.

A.4 The Government

The government sets the domestic interbank interest rate, i_t, according to an inflation forecast targeting rule,

$$i_t = \bar{\imath}_t + \mu \cdot \Delta_t + \varepsilon_t^{mp},$$

where $\bar{\imath}_t$ is a policy neutral level of the nominal interest rate, π is a point target, $\mu > 1$ describes the degree of responsiveness of monetary policy, ε_t^{mp} are deviations from its systematic behaviour, and Δ_t is a discounted sum of deviations from the target of consumer inflation forecasts at all horizons,

$$\Delta_t = \tfrac{1}{1-\psi}\,\mathrm{E}_t \textstyle\sum_{k=0}^{\infty} \psi^k \,(\log \Pi_{t+k}/\Pi_{t+k-1} - \pi),$$

with a policy discount factor $\psi \in (0,1)$ determining the average forecast horizon monitored by the bank, i.e. $\psi/(1-\psi)^2$. The policy neutral rate is determined by preference parameters and the inflation target, see subsection A.5.

[13] And adjusted for the respective consumtpion growth rates so that $h_t^\tau = c_t^\tau$ and $h_t^n = c_t^n$ in the steady state.

A.5 Balanced Growth Path

Along a balanced growth path, the nominal expenditures move in constant proportions; namely (we drop time subscripts):

$$\Pi \cdot \Gamma = p^\tau c^\tau / \omega = p^n c^n / (1-\omega),$$
$$\tfrac{\varepsilon-1}{\varepsilon} \cdot c^\tau P^\tau = c^\tau p^{m*}/s = x p^{x*}/s,$$
$$Wn = L = D = \tfrac{\varepsilon}{\varepsilon-1} \cdot \Pi \cdot \Gamma = \tfrac{\varepsilon-1}{\varepsilon} \cdot p^n c^n.$$

Furthermore, denoting by $\widehat{x_t} = \log(x_t/x_{t-1})$ the log of the gross rate of growth, or by $\widehat{\imath_t} = \log(1+i_t)$ the log of the gross rate of interest, a balanced growth path is characterised as follows:

$$\widehat{\Gamma} = \widehat{c} = \omega \widehat{c}^\tau + (1-\omega)\widehat{c}^n,$$
$$\widehat{c}^\tau = \widehat{h}^\tau = \widehat{x} = \alpha_x,$$
$$\widehat{c}^n = \widehat{h}^n = \alpha_n,$$
$$\widehat{\Pi} = \omega \widehat{p}^\tau + (1-\omega)\widehat{p}^n = \pi,$$
$$\widehat{p}^n - \widehat{p}^\tau = \alpha_x - \alpha_n,$$
$$\widehat{s} = \widehat{p}^{m*} - \widehat{p}^\tau,$$
$$\widehat{\imath} = \widehat{\imath}^{\,d} = \widehat{\imath}^{\,\ell} = \omega\alpha_x + (1-\omega)\alpha_n - \widehat{\beta} + \pi,$$
$$\widehat{\imath} = \widehat{\imath}^* - \widehat{s}.$$

where $\widehat{\beta} = \log(1+\beta)$. The last equation is conditioned by the assumption that the model is parameterised so that the net position of the economy, $D_t - L_t$, is zero.

B QUASI-TRIANGULAR STATE-SPACE FORM OF POLICY FUNCTIONS WITH MULTIPLE UNIT ROOTS

[To be continued.]