

WORKING PAPER SERIES 6

Martin Fukač and Adrian Pagan:
Issues in Adopting DSGE Models for Use in the Policy Process

6
2006

WORKING PAPER SERIES

Issues in Adopting DSGE Models for Use in the Policy Process

Martin Fukač
Adrian Pagan

6/2006

CNB WORKING PAPER SERIES

The Working Paper Series of the Czech National Bank (CNB) is intended to disseminate the results of the CNB's research projects as well as the other research activities of both the staff of the CNB and collaborating outside contributor, including invited speakers. The Series aims to present original research contributions relevant to central banks. It is refereed internationally. The referee process is managed by the CNB Research Department. The working papers are circulated to stimulate discussion. The views expressed are those of the authors and do not necessarily reflect the official views of the CNB.

Printed and distributed by the Czech National Bank. Available at <http://www.cnb.cz>.

Reviewed by: Jan Brůha (Czech National Bank)
Petr Zemčík (CERGE-EI)
Domenico Giannone (Universite Libre de Bruxelles)

Project Coordinator: Juraj Antal

© Czech National Bank, November 2006
Martin Fukač, Adrian Pagan

Issues in Adopting DSGE Models for Use in the Policy Process

Martin Fukač* and Adrian Pagan†

Abstract

Our discussion is structured by three concerns – model design, matching the data and operational requirements. The paper begins with a general discussion of the structure of dynamic stochastic general equilibrium (DSGE) models where we investigate issues like (i) the type of restrictions being imposed by DSGE models upon system dynamics, (ii) the implication these models would have for "location parameters", viz. growth rates, and (iii) whether these models can track the long-run movements in variables as well as matching dynamic adjustment. The paper further looks at the types of models that have been constructed in central banks for macro policy analysis. We distinguish four generations of these and detail how the emerging current generation, which are often referred to as DSGE models, differs from the previous generations. The last part of the paper is devoted to a variety of topics involving estimation and evaluation of DSGE models.

JEL Codes: C11, C13, C51, C52.

Keywords: DSGE model, Bayesian estimation, model evaluation.

* Martin Fukač, Czech National Bank, Economic Research Department, and CERGE-EI (e-mail: martin.fukac@cnb.cz).

† Adrian R. Pagan, Queensland University of Technology (e-mail: a.pagan@qut.edu.au).

We would like to thank Gunnar Bardsen, Mardi Dungey, Sophocles Mavroeides and Jim Nason for their comments on an earlier version of this paper. Research of the second author was supported by ARC Grant No. DP0449659.

Nontechnical Summary

Macroeconometric modelling has seen a steady progression away from single equation data-dominated approaches towards relatively complete, albeit small, systems. This trend has been driven by the desire to have a strong economic perspective which links much more closely with models developed by macro theorists. It does not seem likely that this movement will be reversed. Indeed, a number of central banks have already adopted variants of 4G models as the core model in their forecasting and policy assessment tasks.

It is notable that the more information that a central bank needs to provide in terms of forecasts and explanations of policy decisions, the more there is a tendency to utilize 3G and 4G models. Thus a central bank like the Reserve Bank of Australia, which has to publish little in the way of forecasts, still works with a small 2G model – see Stone et al. (2005). Since it seems unlikely that the future will see a decline in the amount of information and explanation that is demanded, *dynamic stochastic general equilibrium* (DSGE) type models are almost certainly "here to stay".

Given this fact it is imperative that procedures be developed for evaluating these models as well as the processes for constructing them, just as happened with the first and second generation of models. Our feeling is that too little has been done in this regard. In particular, it is still very hard to determine how well the models track the data. Estimation of the parameters has also often been poorly described, although it has to be said that the recent enthusiasm for Bayesian methods may well be going too far in the direction of formal estimation. Often this method seems to conceal as much as it reveals. Its implementation is not entirely straightforward either. Even apart from all the difficulties facing choice of priors, the need to construct posterior densities by simulation methods often requires a considerable amount of experimentation with tuning parameters, such as the scaling factor in proposal densities in a Metropolis-Hastings algorithm. Sometimes the answers seem to be quite sensitive to these choices. This means that replication of Bayesian results is not straightforward and it is unclear whether what has been presented in published studies is in fact a good representation of the "true" posteriors. One needs to have more scepticism about the Bayesian results that have been obtained.

Even if we accept that the models are performing adequately in their explanation of outcomes we still have to recognize that there are many sources of information and some of these have no obvious correspondence to variables in a DSGE model, e.g. confidence measures. Yet these are rarely ignored when making policy decisions. How we adapt these models to make them "policy friendly" is a fascinating question, but it is perhaps the biggest question of them all. Good work has begun on this topic but more attention will need to be paid to it in the years to come.

1. Introduction

That the literature on DSGE models and the resources devoted to experimenting with them by central banks has been rapidly growing can scarcely be disputed. Not only do we see many papers being produced with DSGEs by central bank researchers but we also see advertisements for employment that specify this as an area of expertise. However, since ultimately most research in central banks is designed to assist in making policy choices, it is natural to ask what issues arise if DSGE models are to be given a greater role in this process, today and in the future.

The paper begins with a general discussion of the structure of DSGE models. It uses a stylized representation but the points carry over to more complex models. In that section we investigate a number of issues. One is the type of restrictions imposed by DSGE models upon system dynamics. Another is the implication these models would have for "location parameters", viz. growth rates. The latter investigation is important since it has often been said that it is the non-constancy of such parameters that is the main source of forecast failure of macro-econometric models Clements and Hendry (1999) – and having a model to interpret them is important when it comes to discovering why they may have been non-constant. Finally we look at the question of whether these models can track the long-run movements in variables as well as matching dynamic adjustment, in particular, whether co-integration exists between variables and whether it is of the right quantitative magnitude. These questions have, until recently, largely been neglected with DSGE models, owing to a propensity to remove permanent components from data with filters such as Hodrick-Prescott before engaging in the modelling.

Section 3 of the paper looks at the types of models that have been constructed in central banks for macro policy analysis. We distinguish four generations of these and detail how the emerging current generation, which are often referred to as DSGE models, differs from the previous generations. In particular, since DSGE models have the fundamental feature of being driven by shocks, we enquire into whether this feature is to be found in the latest (fourth) generation of models, and in what sense.

Section 4 looks at a variety of topics involving estimation and evaluation of DSGE models. We provide some general analysis of this and illustrate the arguments by reference to an open economy model constructed by Lubik and Schorfheide (2005a) for the UK. Finally, the models that we have been considering earlier are what are often referred to as "core" or "base" models, and to make these usable for policy often requires some adjustments, either to match the data or to handle forecasting adequately. Accordingly, section 5 considers a miscellany of issues relating to the problems of making DSGE models operational.

Generally, our discussion is structured by three concerns – model design, matching the data and operational requirements. Throughout we will distinguish between the role that DSGE models have played as vehicles for experimenting with new ideas and the role that they might eventually play as policy oriented models.

2. Some Preliminaries

Consider the following stylized version of an economic system of the form

$$B_0 y_t = B_1 y_{t-1} + D x_t + C E_t y_{t+1} + G u_t \quad (2.1)$$

where y_t is a vector of $n \times 1$ variables, x_t is a set of observable and u_t a set of unobservable shocks.¹ There are p observable and less than or equal to n unobservable shocks. If there were more than n of the latter we would be looking at factor models and we sidestep that issue until later. By observable we will mean that the shocks can be recovered from a statistical model. By unobservable we will mainly mean that the shocks are defined by the economic model. Later, however, we will divide unobservable shocks into those that are defined by the economic model and those which are added simply to produce a better tracking of the data; the latter we will call tracking shocks.

We will assume that $u(t)$ is $I(0)$ while x_t can be either $I(1)$ or $I(0)$. In most DSGE models the permanent unobservable shock is technology and that appears in a production function such as

$$y_t = \alpha k_t + (1 - \alpha) l_t + u_{1t}, \quad (2.2)$$

so that, if u_{1t} is permanent, we could write this relation as

$$\Delta y_t = \alpha \Delta k_t + (1 - \alpha) \Delta l_t + e_{1t}, \quad (2.3)$$

where $\Delta u_{1t} = e_{1t}$. It is the fact that u_{1t} is a structural shock that enables us to eliminate it by differencing of the variables in the structural relation. One would then just re-write the system in (2.1) so as to reflect the fact that some relations involve differences. This will generally mean that higher order lags appear in the system. If the structural shock appeared in an equation with a forward looking expectation one would need to account for the fact that terms such as $E_{t-1}(y_t)$ appear in the system, but this doesn't pose any great difficulties with the solution method we outline below.

Let m_p of the x_t and n_p of the y_t be $I(1)$ processes i.e. have a permanent component. Assume all the u_t are $I(0)$. Then we can write

$$(B_0^p - B_1^p - C^p) y_t^p - D^p x_t^p \sim I(0)$$

where the "p" superscript corresponds to the $I(1)$ variables. These are then the long-run relations between the $I(1)$ variables. Since the number of co-integrating vectors will be $n_p + m_p - m_p = n_p$, provided the m_p permanent variables in x_t are not co-integrated, the $n_p \times (n_p + m_p)$ matrix $\begin{bmatrix} B_0^p - B_1^p - C^p & -D^p \end{bmatrix}$ must have rank n_p , and can be written as $\alpha \beta'$, where β is a set of co-integrating vectors. Thus the co-integrating vectors will potentially depend upon the values of B_0^p, B_1^p, C^p, D^p . Looking a bit more closely at this we note that, if the equations had the form (and assuming that all y_t, x_t are $I(1)$),

$$B_0 y_t = B_1 y_{t-1} + B_2 \Delta y_{t-1} + D x_t + C E_t y_{t+1} + G u_t$$

then the long-run relations are unchanged. So the distinction that needs to be made is whether or not all the parameters connected with dynamic terms (lagged values) enter into the long-run

¹ Generally G is composed on known values due to the presence of B_0 .

relations. Parameters such as B_0, B_1, D and C enter the long-run but B_2 do not. Accordingly, while the dynamics may be better matched by the addition of terms like Δy_{t-1} , replication of the dynamics does not mean that the resulting model will yield data-compatible co-integrating vectors and so it might track the long-run movements incorrectly.

To find a representation that eliminates the expectations we follow Binder and Pesaran (1995) and write $\xi_t = y_t - Py_{t-1}$, which is then substituted to obtain

$$\begin{aligned} B_0(\xi_t + Py_{t-1}) &= B_1y_{t-1} + Dx_t + CE_t(\xi_{t+1} + Py_t) + Gu_t \\ &= B_1y_{t-1} + Dx_t + CE_t(\xi_{t+1} + P(\xi_t + Py_{t-1})) + Gu_t \\ &= B_1y_{t-1} + Dx_t + CE_t(\xi_{t+1}) + CP\xi_t + CP^2y_{t-1} + Gu_t, \end{aligned}$$

so that we need $B_0P - B_1 - CP^2 = 0$ to eliminate the y_{t-1} term and to produce

$$B_0\xi_t = CE_t(\xi_{t+1}) + Dx_t + CP\xi_t + Gu_t.$$

This then implies

$$\begin{aligned} \xi_t &= (B_0 - CP)^{-1}CE_t(\xi_{t+1}) + (B_0 - CP)^{-1}Dx_t + (B_0 - CP)^{-1}Gu_t \\ &= \Pi_1E_t\xi_{t+1} + \Pi_2x_t + \Pi_3u_t \end{aligned}$$

and the solution to the latter would be

$$\xi_t = \sum \Pi_1^j E_t(\Pi_2x_{t+j} + \Pi_3u_{t+j})$$

Thus the solution is of the form

$$y_t = Py_{t-1} + \sum \Pi_1^j (\Pi_2E_t x_{t+j} + \Pi_3E_t u_{t+j})$$

and we need to specify the nature of x_t and u_t . In the case where the x_t and u_t are AR(1) processes we would get a Vector Autoregression with Exogenous Variables (VARX) system

$$y_t = Py_{t-1} + D_0x_t + G_0u_t. \tag{2.4}$$

Now the economic theory here is a statement about P . D_0 involves both P and a statistical process for x_t , with the latter capable of being inferred from the data independently of the model. This is not so for u_t as, although one might estimate a process for it, this can only be done by estimating the complete model. It seems important then that one might ask the question of how much of the explanation for y_t comes from the exogenous (and unknown) assumption that u_t is an AR(1) process and how much comes from the economic theory i.e. P . This is not an easy question to answer unambiguously since G_0 is a function of both P and its autoregressive structure, but we can certainly perform some experiments to get an idea of how useful the model is relative to the exogenous assumptions in matching the data.

A study of the nature of P reveals an interesting feature. Since it is defined as $P = (B_0 - CP)^{-1}B_1$, it can have rank no greater than the rank of B_1 . This is not a co-integration restriction as we have assumed that x_t is $I(0)$. Instead it is a common dynamic factor restriction as in Vahid

and Engle (1993). Such restrictions can be very useful in forecasting if they are correct but also very costly if they are not. Thus looking at the rank of P seems a worthwhile task in all cases. As we will see later there seem to be many DSGE models in which P is rank deficient.

We now look at the issue of location parameter changes. We assume that $\Delta x_t = \mu_x + v_t$ so that (2.4) can be written as

$$\begin{aligned}\Delta y_t &= (P - I)y_{t-1} + D_0x_{t-1} + D_0\Delta x_t + G_0u_t \\ &= \Psi z_{t-1} + D_0\Delta x_t + G_0u_t \\ &= \alpha_1\beta'z_{t-1} + D_0\Delta x_t + G_0u_t\end{aligned}\tag{2.5}$$

where $z_t = \begin{bmatrix} y_t \\ x_t \end{bmatrix}$, $\Psi = [P - I \quad D_0]$ and α_1 is the loading for Δy_t on the ECM terms $\psi_t = \beta'z_t$. Hence, taking expectations,

$$\mu_y = \alpha_1 E(\psi_{t-1}) + D_0\mu_x.$$

In turn we could write the system of equations consisting of (2.5) and $\Delta x_t = \mu_x + v_t$ in the form

$$\Delta z_t = F_1\psi_{t-1} + F_2\mu_x + F_3\zeta_t$$

where $F_1 = \begin{pmatrix} \alpha_1 \\ 0 \end{pmatrix}$, $F_2 = \begin{pmatrix} D_0 \\ I \end{pmatrix}$, $F_3 = \begin{pmatrix} G_0 & D_0 \\ & I \end{pmatrix}$, and $\zeta_t = \begin{pmatrix} u_t \\ v_t \end{pmatrix}$. Thus

$$\begin{aligned}\beta'\Delta z_t &= \beta'F_1\psi_{t-1} + \beta'F_2\mu_x + \beta'F_3\zeta_t \\ \psi_t &= (I + \beta'F_1)\psi_{t-1} + \beta'F_2\mu_x + \beta'F_3\zeta_t \\ \Rightarrow E(\psi_t) &= (I + \beta'F_1)^{-1}\beta'F_2\mu_x\end{aligned}$$

and so

$$\mu_y = [\alpha_1(I + \beta'F_1)^{-1}\beta'F_2 + D_0]\mu_x$$

showing that the predicted mean for the growth rate in y_t depends upon the mean of the growth rate in x_t , but also the complete set of parameters in the model. Shifts in μ_y , which will often be the source of forecast failures, can occur because of changes either in μ_x or in some of the model parameters.

Now, in many instances, including the example we look at later, the growth rates for y_t and x_t are "de-measured" before estimation. This is not done with the model but using the actual growth rates of y_t and x_t . Hence the model may imply a growth rate for y_t that is very different from what is in the data and one should clearly check whether this is a problem with the model or not. It is also the case, as Lubik (2005) observes, that the relation above implies restrictions that can be used for estimating the model parameters. Note that, if there is no co-integration, then $\alpha_1 = 0$, and so the prediction is that $\mu_y = D_0\mu_x$, which still depends on the model parameters.

The discussion above has proceeded as if y_t could be regarded as observed. But it is generally a log deviation from some value that represents a point of attraction for that variable. When the

system variables are stationary or have a deterministic growth path it is easy to relate y_t to observables. Where the situation becomes more complex is if the observed data is an $I(1)$ process so that y_t is measured as the observed variable relative to its $I(1)$ component as predicted by the model. Thus, if Y_t^0 is the level of observed output, and A_t is the level of technology, it is conventional to form $Y_t = \frac{Y_t^0}{A_t}$ i.e. $y_t = y_t^0 - a_t$, where the lower case letters represent logs of the capital letters. In this case another equation is needed relating the unobserved quantity y_t to the data, and that is generally provided by

$$\Delta y_t^0 = \Delta y_t + \Delta a_t. \tag{2.6}$$

Afterwards (2.6) combines with (2.4) when estimation is to be performed. Since these two sets of equations constitute a state space form, Kalman filtering methods are normally employed to extract estimates of y_t .

3. Base Policy Model Design

3.1 The First Three Generations of Models

Policy models can be regarded as evolving through four generations. The first generation was largely driven by the IS/LM framework and involved writing down equations which described the determinants of variables in the national accounting identity for GDP, e.g. investment and consumption. Dynamics were introduced through distributed lag relations. Many issues relating to model construction and evaluation were posed and solved by workers with these models.

Second generation models, such as the Canadian model RDX2 – Helliwell et al. (1971) – and the MPS model² at the Fed, introduced much stronger supply side features and also moved towards deriving some of the relationships as the consequence of static optimization problems solved by agents – in particular the consumption decision and factor choices often came from this perspective. Dynamics were again introduced by modifying the static relationships with the use of distributed lag ideas. But a new development was that the latter were often implemented through an error correction form, the use of which had been popularized in the late 1970s by Davidson et al. (1978), in which the static solution represented a target to which the decision variable adjusted. Like the previous generation of models there was considerable diversity within this class and it expanded over time. Often this diversity was the result of a continual incorporation of new features, e.g. rational expectations were introduced into financial market decisions. However, it was often the case that these new features were not easy to satisfactorily reconcile with the existing large-scale models and this often led to a good deal of dissatisfaction with the adapted versions – see the discussion in Coletti et al. (1996)

Third generation (3G) models responded to the latter difficulties, in particular the fact that these models rarely converged to a steady state solution when simulated.³ Consequently, these models became much smaller and emphasis was placed on the need to initiate their construction by designing a steady state that they were to converge to (more often a steady state growth path, but we will ignore that qualification), and to fully account for stock-flow interactions. In particular stocks had to change in such a way as to eventually exhibit constant ratios to flow variables.

² Gramlich (2004) observes that this was also called the Federal Reserve-MIT-Penn model and this is probably a better name given the Fed's role in its construction.

³ Gramlich comments on his work on the MPS model that " ..the aspect of the model that still recalls frustration was that whenever we ran dynamic full-model simulations, the simulations would blow up".

It was much easier to ensure that these characteristics held by setting up a model in which there were well-defined optimization choices for households and firms, along with rules for monetary and fiscal authorities, than trying to force them upon models in which these decisions were largely ad-hoc. An early version of this class of models that was used for forecasting and policy work was Murphy (1988), which was distinguished by the fact that it possessed a well-defined steady state.⁴

Perhaps the most influential of the 3G models, however, was the QPM model of the Bank of Canada. Basically this model was constructed in two stages. In the first stage optimization problems were posed and solved by agents, as detailed in Black et al. (1994). These Euler equations were then combined with identities and solved for a steady state equilibrium between variables. This *base model* incorporated some internal or *intrinsic dynamics* owing to the presence of identities connecting stocks and flows. A second stage involved adapting the base model to the data by the addition of external or *extrinsic dynamics*, as discussed in Colleti et al. (1996). In this second stage a variety of methods were used to introduce dynamics. A popular procedure was to set up a synthetic optimization problem in which agents attempted to close the gap between the steady state solution for a variable and the observed value, but were constrained from doing this instantaneously by there being some costs of adjustment. As Nickell (1985) showed, if the costs were quadratic, then this formulation led to error correction models, and that equivalence meant that the speed of adjustment could be estimated directly from applying these forms to the data.

Thinking of the origin of an ECM which has been estimated from data as coming from some optimization structure was probably useful in the context of explaining policy outcomes, but wasn't really needed if the objective was simply to fit the data closely.⁵ In most instances the equilibrium solutions were ones that described relations that should hold between different series rather than a computed set of observations on the steady state variables. These models became dominant in the 1990s, being used at the Reserve Bank of New Zealand (FPS), the Federal Reserve (FRBUS) and, more recently, the Bank of Japan (JEM). A distinctive feature of the models is that it was necessary to provide a solution to the open-economy problem caused by the fact that agents in a small open economy could borrow at a rate of interest that was fixed, and so could finance consumption streams indefinitely – see Schmidt-Grohe and Uribe (2003) for a discussion of methods for dealing with this issue. In practice, two strategies tended to be used to rule out such behaviour. In the first the representative consumer of closed economy models was replaced by agents with finite lives. This formulation could be shown to be equivalent to a model with a representative consumer but one whose discount rate depended on the probability of death – this is often known as the Blanchard-Yaari solution. A second approach was to have the risk premium attached to foreign debt rising with the level of foreign borrowing, so that eventually agents would not wish to borrow from foreign sources to finance consumption. The ratio of foreign debt to GDP became a crucial element in the latter models and decision rules had to be constructed to ensure that this prescribed ratio was achieved in steady state.

These models also allowed for the possibility that decisions about prices, consumption, investment etc. might be based upon expected future outcomes, rather than just those made about

⁴ The model was more fully described in Powel and Murphy (1995).

⁵ Although there is an isomorphism between ECMs and the strategy of maximizing some linear quadratic function defined over the gap between steady state and actual values to which is added some quadratic adjustment costs, it may be that, in practical applications, the optimization approach tends to rule out certain ECM models, e.g. when fitted to data the fitted ECM model may suggest odd lag structures which are hard to rationalize from an economic perspective, even if it is mathematically possible to do so.

financial variables. In practice a combination of both backward and forward looking elements were incorporated into decision rules with the greater weight being given to the "backward effects". Finally, in most instances the parameters of the base model were essentially calibrated by utilizing information in "great ratios", in first moments, and from the attitudes of policy makers and advisers. Even the parameters determined in the second stage of model construction – largely relating to dynamic adjustment – were influenced by attitudes of bank staff about the likely dynamic effects of particular shocks, as well as the evidence on dynamic adjustments in the data through past studies using (say) VARs.

3.2 Fourth Generation Models and Their Characteristics

Today we are seeing the emergence of a fourth generation (4G) of models. Some of the early representatives are TOTEM (Bank of Canada, Binette et al. (2005)), MAS (the Modelling and Simulation model of the Bank of Chile, Medina and Soto (2005)), GEM (the Global Economic Model of the IMF, Laxton and Pesenti (2003)), BEQM (Bank of England Quarterly Model, Harrison et al. 2004), NEMO (Norwegian Economic Model at the Bank of Norway, Brubakk et al. (2005)) and The New Area Wide Model (NAWM) at the European Central Bank.

These new models have two striking general characteristics. First they feature considerable heterogeneity. Thus a variety of types of labour services are available, there are many intermediate goods produced, and there may be a number of final goods produced, rather than the single good of the previous generation of models. This heterogeneity is often associated with monopolistic and monopsonistic behaviour rather than the competitive markets of the 3G models. Second, the degree of intrinsic dynamics has been expanded a great deal, with a large number of constraints upon agents when making decisions, including habit persistence in consumption and labour choices, adjustment costs in investment and labour, capital utilization variations, and wages and prices being adjusted according to various staggered price setting and contractual arrangements. Thus the base model now has much more of the dynamic structure coming from optimal decisions than was the case in the previous generation of models. 4G models do maintain some of the 3G model characteristics though. The models are still relatively small and feature both forward looking and backward looking behaviour. They also still have some features that stem from knowledge about what is needed to fit the data. Prominent among the latter is the distinction between "spenders" and "savers". The latter make optimal decisions about consumption, taking into account that there is some habit persistence built into their utility functions. The former, previously often described as "liquidity constrained consumers", have their consumption determined by current incomes, and their ability to borrow on capital markets is constrained.

It is worth dwelling on the heterogeneity feature. TOTEM is selected as a representative of what is being done in these models. Here there is a continuum of domestic intermediate goods being produced by many firms using a capital/labour composite and a commodity good. Final goods then combine together with domestic goods and imported goods. Different firms produce these differentiated goods and so they hire different types of labour. Imports of the different goods are essentially done by an aggregator who combines these and sells them on to firms. This large degree of heterogeneity might seem to pose difficulties for a macro model, where only aggregated observations are available, but the functional forms describing how the items are combined together, generally CES, ensure that this heterogeneity produces an aggregate demand function for the final good, and the only way in which heterogeneity appears is that a parameter describes the distribution of the different commodities, labour types etc. Sometimes this parameter is treated as evolving stochastically over time, and so is not directly observed, although micro-economic evidence is normally invoked to describe its likely variation. This

heterogeneity is important since it enables monopolistic pricing, and therefore different price adjustment schemes for different firms can be imposed. The micro-macro aggregation procedure is rather clever, and provides an interpretation to some of the aggregate demand functions, but how important it will be to policy discussion remains to be seen. Estimation of parameters in these models still seems to largely involve the calibration strategy that emerged in the third generation models, although often it is projected that there will be more systematic estimation in the future, particularly from a Bayesian perspective.

3.3 Are Fourth Generation Models DSGE Models?

On the surface fourth generation models certainly seem like DSGE models. Euler equations describe agents' decisions, there are extensive cost of adjustment schemes to introduce dynamics, forward looking expectations are standard etc. Many of the shocks that are standard in DSGE models also appear in them. But there seems to be a difference over how observable and unobservable shocks are treated. This difference really traces back to two purposes for shocks – they might be *explanatory* or *exploratory*. In DSGE models they serve both purposes, i.e. they tell one about the model through impulse responses but are also integral to explaining the data. That is true of observable shocks in 4G models as well, but not for unobservable ones. In existing uses of these models shocks such as technology seem to be treated as simply deterministic trends when it comes to assessing whether these models can match the data and are therefore not regarded as being stochastic. Thus Amano et al. (2002, p 11, footnote 11) say when evaluating QPM that "An obvious omission is a productivity shock. In previous work, however, we found, given the current structure and calibration of the shocks, only a small effect arising from productivity shocks". When the models are made operational it may be that a non-deterministic path for unobserved shocks is imposed based on views about its likely evolution, but it is unlikely to be of the simple type often used in DSGE models, i.e. autoregressive structures. This seems a substantial difference to DSGE models. It has the advantage that the models can be solved under the assumption of perfect foresight for observable shocks whereas unobservable shocks can be generated with some statistical model.

There are some exceptions to this blanket statement. TOTEM treats shocks to the inflation target as unobserved and composed of a random walk plus a transitory component and agents need to extract an estimate of the target via Kalman filtering. This will generally mean that they use weighted moving averages of observable variables such as inflation in place of the target. When made operational, however, it would seem that a fixed target was assumed.

A further difficulty with unobservable shocks is that they are defined by the model. A similar situation occurs in a regression context and, although the US approach to econometrics has largely been to treat the error in the regression as a shock with specified properties, there has always been an alternative treatment, particularly in the UK, of regarding this error as a residual, since it is what the model cannot explain. In the latter interpretation there is a serious question about whether (say) a "monetary policy shock" really is that, rather than simply a reflection of the fact that the economic model we are working with is not rich enough to explain observed interest rate decisions. In this latter perspective the interpretation of residuals from the interest rate equation in VAR models as policy shocks often seems misguided.

4. Matching the Data

It is virtually impossible to think that a model which is to be used for policy analysis would not be influenced by data. In particular the need to assign values to the parameters of such models raises the issues of how one does that and how one learns about the accuracy of such choices as well as whether the quantified model is a good representation of the data.

The first item that needs to be addressed is whether it is possible to learn about the values of the parameters of these models. This literature generally goes under the heading of "identification". In extreme cases the model may be unidentified and so one can learn nothing about the parameter values from the data. In most cases however there is "weak identification" in which the data is fairly uninformative about the parameter value and it is also difficult to produce a precise measure of how uninformative it is.

If one thinks that one can obtain some information from the data there are many estimation methods whereby one does this. Estimation approaches involve either formal or informal uses of the data. The latter are often termed "calibration" and often constitute a wide range of procedures – matching of moments, use of opinions and intuition, evidence from previous micro and macroeconomic work etc. Informal methods are rarely uninformed by data. There is a case that they can be highly effective – they can often provide a filter against errors in data and can combine together quite a lot of information in a useful way. The issue shouldn't really be whether informal methods are "bad" estimation methods, but rather whether one performs an adequate evaluation of any model whose parameters have been quantified by such an approach. Indeed, as we will see later, formal methods can also have difficulties and this raises the question of whether they are really superior to informal ones. It might be thought that formal methods have the advantage of replicability, but, as anyone who has done empirical work knows, this is often illusory.

In the sections below we look at issues of identification in DSGE models, the methods proposed to estimate their parameters and measure the degree of uncertainty about their likely values, and, finally, procedures for assessing how well these models match the data.

4.1 Where does Identification Come From?

Let the parameters in the base model be θ and the unknown parameters in B_0, B_1, D, C , and G be η . Since $G = I$ is most common (thus making the standard deviations of the shocks unknown) we will impose that restriction and also ignore the fact that often the shocks η are restricted to be uncorrelated, an assertion that can provide identifying information. It is worth looking at the identification issue in two stages. First, one could ask whether it is possible to recover θ if η is known? If not be then we must have a failure of identification since, when there is identification, Kodde et al. (1990) show that (asymptotically) the MLE of θ can be recovered by utilizing estimates of η . Second, one might ask whether it is possible to estimate η without using any of the restrictions that are imposed between the η due to the fact that $\dim(\theta) < \dim(\eta)$.

To investigate the problems in estimating η , for convenience (and because many DSGE models make this assumption), let us assume that both x_t and u_t are VAR(1)'s with $x_t = \Phi_x x_{t-1} + v_t$

and $u_t = \Phi_u u_{t-1} + e_t$. Then

$$\begin{aligned} E_t(y_{t+1}) &= Py_t + D_0 E_t x_{t+1} + G_0 E_t u_{t+1} \\ &= Py_t + D_0 \Phi_x x_t + G_0 \Phi_u u_t \\ &= Py_t + D_0 \Phi_x x_t + G_0 \Phi_u G_0^+ (y_t - Py_{t-1} - D_0 x_t) \\ &= z_t' \delta \end{aligned}$$

where $z_t = \begin{bmatrix} y_t & y_{t-1} & x_t \end{bmatrix}$. Hence the Euler equations can be re-written as

$$B_0 y_t = B_1 y_{t-1} + D x_t + C z_t' \delta + G u_t,$$

It's clear that to estimate the parameters of these equations we will need to have instruments for the y_t that appear on both the LHS and the RHS of the equations. The latter have coefficients from B_0 attached to them (once we normalize on one of the y_t for the LHS). The former stem from the fact that y_t also appears in $z_t' \delta$, so that instruments are needed for those terms as well. We will assume that we know δ since it can always be estimated by regressing y_{t+1} against y_t, y_{t-1} and x_t .⁶

The New Keynesian Policy Model (NKPM) is a good example of the issues that arise in attempting to estimate η . In its simplest form it has a Phillips curve, an IS curve and an interest rate rule

$$\pi_t = \eta_1 \pi_{t-1} + \eta_2 E_t(\pi_{t+1}) + \eta_3 \xi_t + u_{St} \quad (4.7)$$

$$\xi_t = \eta_4 \xi_{t-1} + \eta_5 E_t(\xi_{t+1}) + \eta_6 (r_t - E_t(\pi_{t+1})) + u_{Dt} \quad (4.8)$$

$$r_t = \eta_7 r_{t-1} + \eta_8 \xi_t + \eta_9 \pi_t + u_{It}, \quad (4.9)$$

where π_t is inflation, ξ_t is demand and r_t is an interest rate. We might have $E_t(\xi_{t+1})$ and $E_t(\pi_{t+1})$ in place of the current values in the policy rule without changing any of the discussion. There are no observable shocks in this system.

Consider the estimation of the parameters of this system assuming there is no serial correlation in the shocks. First, it is clear that the rank of P in (2.4) is likely to be 3 and so r_{t-1}, ξ_{t-1} and π_{t-1} all affect the three variables. Hence three instruments are available to estimate each equation. One of these variables appears lagged in each equation but there are only ever two variables that instruments are needed for in each equation, so that potentially all of the η_j should be capable of being estimated without using information about the covariance matrix of the errors.

Of course there are extra instruments that are not used – namely the residuals from the first equation when estimating the second equation, and the residuals from both the first and second equations when estimating the third. These residuals become instruments owing to the assumption that the shocks are uncorrelated with one another, but this is not due to any economic analysis. As the size of DSGE models rises towards those of 4G models it must become increasingly difficult to rationalize zero correlation between shocks given names such as "mark ups" and "utilization". Against this tendency is the fact that it seems virtually impossible for the assumption to be correct if one only had a two variable system involving demand and supply shocks. The possibility that the assumption is incorrect would mean that basic IV estimation of the NKPM could be more reliable than MLE, as it uses only exclusion restrictions as in

⁶ This assumes shocks are white noise. If they exhibit serial correlation a different estimator is needed.

standard simultaneous equation estimation, whereas MLE generally imposes the uncorrelated shocks assumption.

Now suppose that the NKPM only had forward looking variables, i.e. $\eta_1 = 0, \eta_4 = 0$, a situation sometimes found in theoretical models. Then P has rank one and only r_{t-1} is available as an instrument. Thus in this case it is impossible to estimate all of the remaining η_j .⁷ Now what happens if there is serial correlation in the shocks of the first two equations of the NKPM (a very common assumption)? The expectations are constructed differently but instruments are still needed for them. However now we can transform the equation to eliminate the serial correlation, e.g. the inflation equation with an AR(1) for its shock of the form

$$u_{St} = \rho_S u_{St-1} + e_{St}$$

would become

$$\begin{aligned} \pi_t &= \rho_S \pi_{t-1} + \eta_2 z_t' \delta_1 - \eta_2 \rho_S z_{t-1}' \delta_1 \\ &\quad \eta_3 \xi_t - \eta_3 \rho_S \xi_{t-1} + e_{St}. \end{aligned}$$

Since the same transformation applies to the first two equations this clearly means that there will now be three instruments available to estimate this equation, viz. ξ_{t-1}, r_{t-1} , and π_{t-1} . Hence the assumption that the shocks have an AR structure generates enough instruments for the estimation of ρ, η_2 and η_3 in the inflation equation, and this is also true of the remaining equations. Of course this does not come from the structure of the model but is simply a consequence of an extraneous assumption about shocks.

To take another example that is in the literature, Canova and Sala (2005) look at the following version of the NKPM

$$\begin{aligned} \xi_t &= \frac{h}{1+h} \xi_{t-1} + \frac{1}{1+h} E_t(\xi_{t+1}) + \frac{1}{\varphi} (r_t - E_t(\pi_{t+1})) + v_{1t} \\ \pi_t &= \frac{\omega}{1+\omega\beta} \pi_{t-1} + \frac{\beta}{1+\omega\beta} E_t(\pi_{t+1}) + \\ &\quad \frac{(\varphi+v)(1-\zeta\beta)(1-\zeta)}{(1+\beta\omega)\zeta} \xi_t + v_{2t} \\ r_t &= \phi_r r_{t-1} + (1-\phi_r)(\phi_\pi \pi_{t-1} + \phi_y \xi_{t-1}) + v_{3t} \end{aligned}$$

The coefficients of the LHS variables will be numbered $\eta_j, j = 1, \dots, 9$. Now suppose that the v_{jt} were white noise. Then we can estimate all of the η_j as there are three instruments π_{t-1}, i_{t-1} and ξ_{t-1} – two of these never appear as regressors in each of the first two equations, and only two endogenous variables appear on the RHS of these equations. However, even though η_j are identified, it is immediately obvious that this is not true of θ , since $\eta_6 = \frac{(\varphi+v)(1-\zeta\beta)(1-\zeta)}{(1+\beta\omega)\zeta}$ is the only η_j that involves the two parameters ν and ζ .

⁷ This is true even when the restrictions coming from orthogonal shocks are employed since the number of moment conditions available from using r_{t-1} as an instrument is three while the number from the covariance matrix is six, giving nine moment conditions. But there are ten parameters to be estimated – seven of these are η_j and three are the standard deviations of the shocks. One needs to impose an extra restriction to get identification and often this is $\eta_5 = 1$. General analyses of identification in the NKPM are available in Mavroeidis (2004) and Nason and Smith (2005).

Now in Canova and Sala v_{1t} and v_{2t} are AR(1) processes and, as we might expect from the discussion above, this aids identification a great deal. Elimination of the serial correlation in the second equation introduces terms $\frac{\rho(\varphi+\nu)(1-\zeta\beta)(1-\zeta)}{(1+\beta\omega)\zeta}\xi_{t-1}$ and $\frac{\rho\omega}{1+\omega\beta}\pi_{t-1}$ and creates two extra η_j to use to estimate the three parameters ρ , ν and ζ . Thus identification is once again being achieved by the assumption of the shocks being AR processes, which is not part of the economic model. Of course, as in all the discussion here, the parameters may be very weakly identified and, from Canova and Sala's numerical results, that seems very likely, but it is worth understanding when the model would be exactly unidentified. Notice that one can also derive their result that a lack of identification may depend on whether a systems perspective is taken, simply by looking at how many instruments are available if only various single equations are used in estimation.

As mentioned previously other restrictions may be imposed upon the η_j either by theoretical considerations or as an outcome of the chosen functional form. Thus, as we saw in (2.3), the assumption that the technology shock is permanent and that the production function was Cobb-Douglas meant that an instrument was needed for $\Delta k_t - \Delta l_t$ in order to estimate the share parameter α . This is an example of a restriction between the η_j . Note that in one of the examples in Canova and Sala there is no labour so they effectively need to find an instrument for Δk_t . If Δk_t is close to white noise then there will be no good instruments and so it will be very hard to estimate α (a conclusion they reach from their simulations).⁸

4.2 Estimation Techniques

There are various formal methods of estimation, differentiated largely by how much credence is to be placed upon the complete DSGE model. Single equation method of moments estimators like GMM, which work off the moments coming from Euler equations, utilize the complete system only to the extent of suggesting what would be reasonable instruments. Maximum likelihood methods, which maximize a log likelihood, $L(\theta)$, with respect to the model parameters θ , try to improve on the precision of GMM by using the precise structure of the DSGE model. As has been known for a long time, such efficiency can come at the expense of bias and inconsistency of estimators, unless the complete system is an adequate representation of the data. As Johansen (2005) has pointed out, this is a price of MLE, and it should not be assumed that the DSGE model has that property. Again this calls for a proper examination of the extent to which the DSGE model is capable of capturing the main characteristics of the data.

Bayesian methods have also become increasingly popular. To get point estimates of θ comparable to MLE, one can maximize $L(\theta) + \ln p(\theta)$, where $p(\theta)$ is the prior on θ . The resulting estimate of θ is often referred to as the mode of the posterior. An advantage of the Bayesian method is that there is often information about the range of possible values for θ , either from constraints such as the need to have a steady state or from past knowledge that has accumulated among researchers. Imposing this information upon the MLE is rarely easy. It can be done by penalty functions, but often these make estimation quite difficult. Adding on $\ln p(\theta)$ to the log likelihood generally means that the function being maximized is quite smooth in θ , and so estimation becomes much easier. We think that this advantage has been borne out in practice; the number of parameters being estimated in DSGE models like Smets and Wouters (2003) is quite large, and one suspects that MLE estimation would be quite difficult.

⁸ The interaction between the nature of data and the ability to estimate model parameters was also seen in Gregory et al. (1993) where it was shown that, in a discounted quadratic objective model, the discount factor could not be identified if the forcing variables were $I(1)$.

There is, however, a cost to Bayesian methods. Unlike penalty functions the use of a prior changes the shape of the function being optimized. If $L(\theta)$ is flat in θ then the choice of prior will become very important in determining the estimated parameter values. In DSGE models this seems likely to become an issue. We will illustrate this argument later in an empirical example where what would seem to be a perfectly satisfactory estimate of a parameter by a consistent estimator (OLS) is negative, but the Bayesian estimator is strongly positive, since the prior was set up so that a negative value could not be found.⁹

Another reason why one suspects there may be difficulties in estimating DSGE models is that too much is being asked of the data. The context in which we make this remark is the fact that many DSGE models are estimated treating variables such as the capital stock as unobserved, i.e. there are more unobserved variables in the model than observed variables, so the unobserved variables must be concentrated out of the joint density of the model variables to get the likelihood. This is generally done using the Kalman filter. But that filter does require normality of observations and, given that series like interest rates are rarely normal, this seems immediately to say that the DSGE model likelihood is incorrect. Even if we ignore that problem, and assume that the Kalman filter is an appropriate way of proceeding, it has to be recognized that the presence of more unobserved than observed variables puts great stress upon data sets when it comes to using them to find estimates of the parameters generating the unobservable variables. Strong assumptions may need to be made about variances in order to achieve identification of these parameters, something that does not seem to be appreciated by many of those applying the methods. For example it is not enough to follow Smets and Wouters (2003, p1140) who say "Identification is achieved by assuming that four of the ten shocks follow a white noise process. This allows us to distinguish those shocks from the persistent "technology and preference" shocks and the inflation objective shock".

To see the problem that arises with having an excess of unobservables consider the simplest case where we have one observed variable y_t but two unobserved components y_{1t} and y_{2t} . One of these components (y_{1t}) follows an AR(1) with parameter ρ_1 and innovation variance σ_1^2 , and the other is white noise ($\rho_2 = 0$) with variance σ_2^2 . Then we would have

$$(1 - \rho_1 L)y_t = (1 - \rho_1 L)y_{1t} + (1 - \rho_1 L)y_{2t}$$

and it is clear that, as $\frac{\sigma_2^2}{\sigma_1^2}$ becomes large, it becomes impossible to identify ρ_1 . In this case the likelihood is flat in ρ_1 , and any prior placed on ρ_1 will effectively determine the value of ρ_1 that results. To avoid this situation a prior would need to be placed on the relative variance and not just the values of ρ_1 and ρ_2 as Smets and Wouters argue. To illustrate this we simulated some data from the set up above and then estimated ρ_1 with a beta prior centred at different values. The true value of ρ_1 is .3 and the table below shows the posterior mode for different values of $\frac{\sigma_2^2}{\sigma_1^2}$. It is clear that recovering the true value of ρ_1 is extremely difficult if the type of prior used in many DSGE models is adopted.

⁹ The mode of the posterior is generally used to begin a process of simulating realizations from the posterior density for θ . Often the method used is that set out in Schforheide (2004). One wonders how useful the posteriors being reported are, since being able to characterize a high dimensional density accurately requires huge numbers of realizations from it – the empty-space phenomenon. To illustrate this consider estimating the height at the origin of a multi-dimensional density. Table 4.2 of Silverman (1986) vividly illustrates the fact that, when the density is $N(0, I_d)$, a 90% accuracy for the estimate requires a sample size of 4 when $d = 1$ but one of 842,000 when $d = 10$. There are often far more parameters in DSGE models than ten.

Table 4.1: An example of a too-many-unobservables model estimation. Estimates of ρ_1 and 90% confidence interval

Prior	True σ_2^2/σ_1^2		
	1	2	5
$\rho_1 = 0.85$	0.67 [0.49-0.84]	0.71 [0.53-0.89]	0.80 [0.64-0.94]
$\rho_1 = 0.50$	0.46 [0.32-0.60]	0.48 [0.31-0.62]	0.49 [0.35-0.65]
$\rho_1 = 0.30$	0.28 [0.12-0.41]	0.28 [0.13-0.46]	0.29 [0.12-0.44]

Note: We use a beta prior on ρ_1 , with a standard error 0.1. The true value is $\rho_1 = 0.3$. For σ_1 and σ_2 we use an inverse gamma with a mean 1 and standard error 4 as a prior.

It seems reasonable that one might want to use a variety of estimators rather than a single one. Sometimes the GMM estimator seems to be regarded as inferior, e.g. Lubik and Schorfheide (2005) say about GMM estimation of Euler equations "While potentially robust to mis-specification, this approach suffers from subtle identification problems that can often lead to implausible estimates. Full-information based methods, on the other hand, use the optimal set of instruments embedded in the model's cross-equation restrictions and make identification problems transparent". It is unclear what is meant by the last statement but the first part of the argument is misleading. FIML does indeed use an optimal choice of instruments but GMM can also do this, e.g. in the old simultaneous equations literature Brundy and Jorgenson (1971) suggested the FIVE and LIVE estimators. These estimators used as instruments the predictions from the derived reduced form after imposing either all the restrictions on it (FIVE) or the sub-set of them stemming from the equation being estimated as instruments (since any set of weights attached to the instruments produce consistent estimators we are free to choose these weights however we wish) and Fuhrer and Olivei (2004) essentially applied these estimators when estimating NKPM systems. Because MLE imposes restrictions on the parameters θ it is sensitive to mis-specification and it is on this dimension, rather than in the choice of instruments, that it can be inferior to single-equation GMM. It would seem to be a good idea to always estimate the Euler equations with GMM using instruments constructed as predictions from the complete system estimated by Bayesian or ML estimators, since then we have held the instruments constant and any differences are due to the cross-equation restrictions. If the differences are very pronounced one might be concerned about the possibility that specification errors have affected the ML and Bayesian estimators. A Hausman test might be constructed which compares the MLE and single-equation GMM estimates. It possibly should be said that it is not always easy to apply GMM as there may be too many unobserved variables in the Euler equations to make for easy estimation – this occurs in the system we analyse later.

4.3 Evaluation Issues

There is an old literature on the evaluation of DSGE models and an emerging literature on the evaluation of models that are closer to policy models in size and complexity. Two recent studies stand out – an evaluation of QPM by Amano et al (2002) and an analysis of the Smets and Wouters model by del Negro et al. (2004). The former is a very comprehensive examination of the QPM model and its characteristics in relation to the data, whereas the latter is more concerned with the development of some new evaluation techniques, although it also makes some points that are relevant for the analysis of any 4G model. In the documentation of BEQM – Harrison et al. (2005) – there is also some mention of the ability of the model to match the data, but this is a little sketchy.

Evaluation really has two dimensions to it. One is largely focussed upon the operating characteristics of the model and whether these are "sensible". The other is more about the ability of the model to match the data along a variety of dimensions. The two themes are not really independent but it is useful to make the distinction. Thus it might be that, while a model could produce reasonable impulse responses, it may not produce a close match to the data, and conversely.

4.3.1 Operating Features

Standard questions that are often asked about the operating features of the model are whether the impulse responses to selected shocks are reasonable and what the relative importance of various shocks are to the explanation of (say) output growth. Although the latter is often answered by recourse to variance decompositions perhaps a better question to ask is how important the assumptions made about the dynamics of shocks are to the solutions, as it seems crucial to know how much of the operating characteristics and fit to data comes from the economics and how much from exogenous assumptions. This concern stems back at least to Cogley and Nason (1993), who argued that standard RBC models produced weak dynamics if shocks were not highly serially correlated. In terms of (2.4) one can envisage computing the one-step ahead forecast for y_t, \hat{y}_t , as

$$\hat{y}_t = P y_{t-1} + D_0 x_t + G_0^+ \Phi (y_{t-1} - D_0 x_{t-1}), \quad (4.10)$$

and seeing the effect as Φ is varied. An alternative would be to re-estimate P, D_0 etc for the different values of Φ . We use the first of these in the later empirical work, i.e. we set $\Phi = 0$ since it seems to more directly answer the question we are asking.

The appropriate strategy for assessing operating characteristics depends on whether the model parameters have been formally or informally quantified. If done informally researchers such as Amano et al. (2002) and Canova (1994) have asked the question of whether there is a set of such parameters that would be capable of generating some of the outcomes seen in the data, e.g. ratios ϕ such as (say) the consumption-income ratio. Such a ratio would be a function of the model parameters θ . The existing value used for θ in the model, θ^* , is then taken as one element in a set and a search is conducted over the set to see what sort of variation would occur in the resulting value of ϕ . If it is hard to reproduce the value of ϕ observed in the data, $\hat{\phi}$, then the model might be regarded as suspect. In this approach the estimate of θ from the data is held fixed and the possible values of model parameters are varied to trace out a range of values of ϕ . An efficient way of doing this is a pseudo-Bayesian approach in which the range for θ is described by a multivariate density, and then the induced density for ϕ can be determined. If the observed value $\hat{\phi}$ lies too far in the tails of the density of the induced value of ϕ , one would regard the model as inadequately explaining the feature summarized by ϕ .

The second approach treats the parameter values entered into the model, θ^* , as constant and asks whether the estimate $\hat{\phi}$ is close to the value $\phi^* = \phi(\theta^*)$ implied by the model. This is simply an encompassing test of the hypothesis that $\phi = \phi^*$. One needs to make these tests robust as one does not know whether the model is the DGP. If the value of θ used in the model has been estimated, but not in a way that can be easily replicated, these tests can be regarded as (asymptotically) conservative tests of the hypothesis, i.e. if the hypothesis is rejected it would be rejected even more strongly if one allowed for that estimation error – see Breunig et al. (2003) for this use.

4.3.2 How Well Does the Model Track the Data?

Although the procedures just discussed can be thought of as performing some data matching the focus is often upon a limited range of items such as great ratios. It seems desirable to do more than that and, in particular, to emphasize system properties.

In the first and second generation of models a primary way of assessing the quality of models was via historical simulation of them using a set of observed values of exogenous variables. The maxim among the proprietors of such models was "simulate early and simulate often", as that enabled the system properties to be viewed and was a complement to single equation tests of adequacy such as serial correlation. It seems important that we see such model tracking exercises for DSGE models, as the plots of the paths are often very revealing about model performance, far more than might be found from just an examination of a few serial correlation coefficients and bivariate correlations, which have been the standard way of looking at DSGE output to date.¹⁰ It is not that one should avoid computing moments for comparison, but it seems to have been overdone, in comparison to tests that focus more on the uses of these models such as forecasting (which is effectively what the tracking exercise is about).

Now there is a problem with producing such exercises for DSGE models. Let y_t^* be the implied value of y_t produced by the DSGE and let y_t^D be realizations of y_t . In (4.10) we effectively used y_{t-1}^D in place of y_{t-1} to generate the forecasts, i.e. we identified y_t^* as y_t^D , in which case the model perfectly reproduces the data through the values given to the shocks. Thus there is no residual. But, if we solve for the moments of y_t^* implied by the model, these may not match up those from the data, and so the assumption $y_{t-1}^* = y_{t-1}^D$ is implausible. Some way must therefore be found that will allow for a residual or a wedge between y_t^* and y_t^D . Altug (1989) pioneered one way of doing this by writing $y_t = y_t^* + \eta_t$ and then assuming that the η_t were i.i.d. and uncorrelated with model shocks. One can then estimate the $var(\eta_t)$ and extract estimates of y_t^* using Kalman filtering methods. Ireland (2004) has a generalization of this where η_t can be serially correlated. But there seems no reason to think that these residuals should be uncorrelated with model shocks and it is easy to construct cases where they would not be.

An alternative approach was developed by Watson (1993), in which he asked what was the smallest η_t that one needed to reconcile the DSGE characteristics with the same characteristics in the data. Thus when y_t is a single variable and both y_t^* and y_t are *i.i.d.* one can show that the smallest variance of η_t will be $(var(y_t^*) - var(y_t))^2$, and the values of y_t^* which are consistent with this minimal variance will be equal to $\left[\frac{var(y_t^*)}{var(y_t)} \right]^{1/2}$.

¹⁰ One problem with such moment comparisons is when parameters are estimated from the data and these involve moments of shocks. In a regression model this would mean that the variance of the regression error (shock) can be chosen to perfectly match the variance of the variable being explained. Thus the comparison of moments is often best when parameters have not been estimated.

If the data and model are not *i.i.d.* then one needs to somehow solve the same problem allowing for the serial correlation. Watson's suggestion was to find the shock that would minimize the gap between the spectra of y_t^* and y_t . He then showed that the value of y_t^* could be reconstructed as $y_t^* = \Xi(L)y_t$, where $\Xi(L)$ has both backward and forward elements. This creates a difficulty since, if one was asking what the forecast implications of the DSGE model were, one would want to only allow backward lags. Nevertheless, it is clear that it will be useful to look at the y_t^* constructed as in Watson and to then replace y_{t-1} with y_{t-1}^* in (4.10). In many ways the problem is the same as dynamic factor extraction when one wishes to only have a one-sided window. Clearly more work needs to be done on this issue.

Watson illustrates the technique with a basic RBC model. Oddly enough it does not seem to have been used much, although it is obviously a very appealing way of getting some feel for how well the DSGE model is performing. It may also be worth attempting to develop a method that defines the addition of a vector of shock that has some fixed properties, e.g. it is a VAR(1) as in Ireland, and which minimizes the gap between a finite number of model and data autocovariances rather than the spectra.

4.3.3 How Well Does the Model Match Selected Characteristics of the Data

We might think of matching tests as answering four questions:

1. Do variables in the model that have deterministic trends co-trend? Essentially this would ask whether the mean responses in the growth rate for deterministically trending endogenous variables agree with the values projected by using the long-run growth of the deterministically trending exogenous variables.
2. Do variables that are I(1) co-integrate and, if so, are the co-integrating vectors those implied by the model? This is an important question since a failure will mean that forecasts made with the model are being attracted to the wrong point as the horizon lengthens. Moreover, since the co-integrating vectors depend on many parameters of the model, the dynamic responses can be affected by incorrect information on this aspect.
3. Are the dynamic responses consistent with the data?
4. Is the covariance matrix of the implied errors in reduced form representations of the DSGE model consistent with that in the data?

We discuss each of these in turn.

Co-Trending Co-trending is something that can be checked if the model that has been quantified actually produces an estimate of (say) GDP growth. Often, however, data is de-meant, so that all growth variables have zero mean by construction, meaning that there is no specific information in the model about what the means are. To give an example of this that occurs quite a bit, output is scaled by the level of technology, so that the actual growth rate in output is the sum of the labour force growth rate plus the rate of technology change. If actual growth is de-meant then this means that the technology process would be effectively specified as having a zero mean growth. In this case there is no prediction of the model about the growth rate in output. However, we could invert the equation and ask what growth rate in technology would be needed to produce the observed growth rate in output, and we use this approach in our example later.

Integration and Co-integration To date DSGE modelling has not dealt very well with the fact that many macroeconomic variables are integrated, i.e. possess stochastic permanent stochastic components. Often these have been removed by filters such as HP. But this makes little sense since it is reducing the role of technology shocks a great deal. Co-integration is rarely tested. In complex models the co-integrating vectors are often quite complex but can be worked out, e.g. see a simple example in Kapetanios et al. (2005). The standard approach in many DSGE models of having technology as the sole I(1) unobserved shock means that there should be many co-integrating vectors. In practice it is hard to find more than a few among most macroeconomic variables and the only way that the DSGE model and data would then be reconciled is for the shocks in the DSGE model to be I(1), i.e. if they are treated as AR(1) processes the AR parameter will be estimated as very close to unity. This is a very common outcome when these models are estimated, e.g. the Euro-Area model estimated in Adolfson et al. (2005) has a unit root imposed on domestic technology shocks but four other shocks have estimated AR parameters greater than .983. It is hard to believe that these are not unit root processes but, with Bayesian methods that effectively impose a prior that says there is a zero probability of getting a unit root in these shocks, one would never observe that.

Co-integration also implies that the appropriate model for comparisons with the data is a VECM and not a VAR, and a failure to recognize this fact can result in some inaccuracies – a point shown very well in del Negro et al. (2004). If one is to take co-integration issues seriously however, a problem that arises in any comparisons of the model with the data is what co-integrating vectors to use when fitting the VECM to the data. Del Negro et al. utilize the co-integrating vectors implied by the model but that seems to load the comparison in favour of the model in the event that the co-integrating vectors do not agree with those of the data.¹¹ Effectively, this problem emphasizes that a test of whether the co-integrating vectors are compatible with the data should precede any investigation of dynamics through fitted VECMs.

As we have just said we could compare the VECMs implied by the model and data. This does provide a test of dynamic structure but, in the event that it fails, may not be very informative. Often a comparison of estimated and model-implied impulse responses can provide a more useful packaging of dynamic information. This comparison can only rarely be done with the impulse responses to economic shocks since it is hard to identify these shocks from a VECM without using the model. There are some exceptions to this. A technology shock implied by the data can generally be isolated by simply using long-run restrictions, and so its impulse response function can be found without imposing precise model information, but mostly one cannot extract unknown economic shocks without the model restrictions.

Dynamic Response Comparisons One can formally test whether the model VECM matches the data using standard testing methods. Of course such tests are either accept/reject or, in Bayesian terms, low or high probability events. Del Negro et al. (2004) suggest that one should think about the two models as being polar cases and to then connect the two up with a parameter λ that varies from zero to infinity, such that at one extreme one gets the model and, at the other, the data. Of course one needs some criterion to determine what the value of λ is. If one just used the likelihood one would always take that value of λ which maximized the likelihood, i.e. the VECM based on the data. So they place a prior upon λ which ensures that there is a trade-off between the data and the prior. Obviously the answer one gets is crucially dependent on the prior chosen, and it was constructed in order to get the model and data VECMs to be at

¹¹ In fact they eventually use a VAR rather than a VECM because the model-implied co-integrating restrictions seem incorrect based on graphical evidence. This seems an odd response unless there really is no co-integration (rather than just a failure of the model-implied vectors to produce co-integration) and that was never demonstrated.

opposite ends of the range of λ . In some ways the method is reminiscent of ridge regression, in which OLS is at one end of the ridge parameter range, and values of zero for the parameters at the other. In any case if one follows the method one would end up with a value of λ that can be used to determine what value should be placed on the parameters of the VECM. In their work, using an intermediate value of λ seemed to produce better forecasts.

Matching Covariances of VAR Residuals The DSGE model also makes predictions about the covariance matrix of the errors in the VECM. Specifically these would be $G_0 \text{var}(u_t) G_0'$. Because $\text{var}(u_t)$ is generally assumed to be diagonal there are strong second moment restrictions imposed on the VECM. Again, these have rarely been tested in the literature. A useful way of viewing the compatibility of the assumptions about model shocks with the data is to form $G_0^{-1} V_D (G_0')^{-1}$, where V_D is the estimated covariance matrix of the residuals from fitting a VECM to the data. This is an estimate of what the covariance matrix of the shocks would have to be if one was to maintain that the initial impulse response was that implied by the model (G_0). This is informative in its own right but it also tells us whether it would be reasonable to proceed to estimate the DSGE model under the assumption of uncorrelated shocks.

4.4 An Example: A Simple, Structural Open Economy Model

4.4.1 The Model

We wish to illustrate some of the ideas above using a relatively simple DSGE model. For this purpose we use the model in Lubik and Schorfheide (2006) (LS), which is a simplified version of Gali and Monacelli (2002).

The model has four equations. An open economy IS curve

$$\begin{aligned} \tilde{y}_t &= E_t \tilde{y}_{t+1} - [\tau + \alpha(2 - \alpha)(1 - \tau)](R_t - E_t \pi_{t+1}) - \alpha[\tau + \alpha(2 - \alpha)(1 - \tau)]\rho_q \Delta q_t \\ &\quad + \rho_A dA_t - \alpha(2 - \alpha) \frac{1 - \tau}{\tau} (1 - \rho_{y^*}) \tilde{y}_t^*, \\ 0 &< \alpha < 1, \tau^{-1} > 0, \end{aligned} \quad (4.11)$$

an open economy Phillips curve

$$\begin{aligned} \pi_t &= \beta E_t \pi_{t+1} - \alpha(1 - \beta\rho_q) \Delta q_t + \frac{\kappa}{\tau + \alpha(2 - \alpha)(1 - \tau)} \tilde{y}_t \\ &\quad + \frac{\kappa\alpha(2 - \alpha)(1 - \tau)}{\tau[\tau + \alpha(2 - \alpha)(1 - \tau)]} \tilde{y}_t^* \end{aligned} \quad (4.12)$$

an exchange rate equation

$$\Delta e_t = \pi_t - (1 - \alpha) \Delta q_t - \pi_t^*, \quad (4.13)$$

and the policy rule

$$R_t = \rho_R R_{t-1} + (1 - \rho_R) [\psi_1 \pi_t + \psi_2 \tilde{y}_t + \psi_3 \Delta e_t] + \varepsilon_t^R. \quad (4.14)$$

In the equations above \tilde{y}_t is the difference between the log of the level of output (y_t) and the log level of technology (A_t), y_t^* is the log of world output, R_t is the interest rate, π_t is the inflation rate, π_t^* is the world inflation rate, q_t is the log of the terms of trade, and e_t is the nominal exchange rate.

The terms of trade should theoretically be endogenous but this created problems in their estimation and so was assumed to be an observed exogenous shock evolving as

$$\Delta q_t = \rho_q \Delta q_{t-1} + \varepsilon_t^q. \quad (4.15)$$

The variables $dA_t, \tilde{y}_t^*, \tilde{\pi}_t^*, \varepsilon_t^R$ are unobservable shocks, being the growth rate in technology, the world level of output, world inflation and the monetary policy shock respectively. The first three follow AR(1) processes:

$$dA_t = \rho_a dA_{t-1} + \varepsilon_t^a \quad (4.16)$$

$$\tilde{y}_t^* = \rho_{y^*} \tilde{y}_{t-1}^* + \varepsilon_t^{y^*}, \quad (4.17)$$

$$\pi_t^* = \rho_{\pi^*} \pi_{t-1}^* + \varepsilon_t^{\pi^*}, \quad (4.18)$$

with $\rho_a, \rho_{y^*}, \rho_{\pi^*}$ assumed to lie between zero and unity and $\varepsilon_t^a, \varepsilon_t^{y^*}$, and $\varepsilon_t^{\pi^*}$ being white noise processes. The monetary policy shock ε_t^R is taken to be white noise. Since technology is an $I(1)$ process, output will also be, and so the model is made stationary by using $\tilde{y}_t = y_t - \ln A_t$ and $\tilde{y}_t^* = y_t^* - \ln A_t$, where $\ln A_t = \ln A_{t-1} + dA_t$ is the log level of technology. LS estimated the time preference parameter as a function of a real interest rate of 2.5, i.e. $\beta = .99$

4.4.2 Estimation of the Model

The model is estimated on the UK data provided on Schorfheide's web page. This contains series on the quarterly real output growth Δy_t , annualized quarterly inflation $\pi 4_t$, annualized nominal interest rate $R 4_t$, quarterly exchange rate change Δe , and terms of trade growth Δq . The series were de-meant and the data are related to the model variables as follows

$$[\Delta y_t, \pi 4_t, R 4_t, \Delta e, \Delta q] = [\Delta \tilde{y}_t + dA_t, 4 \times \pi_t, 4 \times R_t, \Delta e, \Delta q].$$

LS estimate the model by Bayesian methods. Table 4.2 contains estimates of the parameters.¹²

There is some difference between our Bayesian estimates and LS's. One source of the difference is that we just used 20,000 draws from a Metropolis algorithm and they used one million. All results presented below when evaluating the model match to the data are obtained based on LS's mean posterior estimates reported in their paper, except when they don't provide an estimate, whereupon we use our estimates in Table 4.2

Replicating LS's results, we performed a robustness check on their conclusion about whether the policy rule (4.14) should include a reaction to the change in exchange rate Δe . We considered different prior mean values for the parameter ψ_3 as well as different prior distributions, including a flat one. The posterior of ψ_3 always resulted in a distribution with mode between 0.10 and .20, and different from zero.

Now if one just fitted (4.15) to the data by OLS one would get an estimate of ρ_q of -.18, so that the restriction used in LS's prior that ρ_q lies between zero and unity is clearly counterfactual. It should be noted that the standard error on ρ_q makes the OLS estimate of ρ_q significantly different from zero, so imposing an incorrect prior has the potential to distort estimates. To assess the consequences of this, we therefore used a t density centred on zero as the prior. However, the

¹² To estimate the model we employ DYNARE version 3.042, by S. Adjemian, M. Juillard and O. Kamenik.

Table 4.2: Prior distributions and parameter estimation results

	Prior distribution			LS' original estimates		Our estimates	
	Density	Mean	std	Mean	90% Interval	Mean	90% Interval
ψ_1	Gamma	1.5	0.25	1.17	[0.92,1.41]	1.16	[0.97, 1.33]
ψ_2	Gamma	0.25	0.13	0.4	[0.15,0.63]	0.4	[0.18, 0.61]
ψ_3	Gamma	0.25	0.13	0.12	[0.07,0.18]	0.12	[0.07, 0.17]
ρ_R	Beta	0.5	0.2	0.68	[0.60, 0.77]	0.69	[0.63, 0.77]
α	Beta	0.2	0.05	0.12	[0.06, 0.18]	0.12	[0.07, 0.18]
r	-	2.5	calibrated	2.46	[0.90, 3.97]	-	-
κ	Gamma	0.5	0.1	1.93	[1.21, 2.65]	1.97	[1.18, 2.67]
τ	Gamma	0.5	0.2	0.52	[0.34, 0.69]	0.44	[0.6, 0.64]
ρ_q	Normal	-0.2	0.2	0.09*	[0.01, 0.16]	-0.17	[-0.29, -0.05]
ρ_A	Beta	0.2	0.05	0.2	[0.06, 0.33]	0.18	[0.11, 0.24]
ρ_{y^*}	Beta	0.9	0.05	0.97	[0.95, 0.99]	0.97	[0.96, 0.99]
ρ_{π^*}	Beta	0.8	0.1	0.37	[0.27, 0.48]	0.39	[0.29, 0.50]
σ_R	InvGamma	0.5	4	0.36	[0.29, 0.43]	0.33	[0.26, 0.39]
σ_g	InvGamma	1.5	4	1.39	[1.21, 1.57]	1.33	[1.17, 1.50]
σ_A	InvGamma	1	4	0.67	[0.58, 0.77]	0.57	[0.48, 0.65]
σ_{y^*}	InvGamma	1.5	4	1.18	[0.70, 1.65]	0.83	[0.32, 1.49]
σ_{π^*}	InvGamma	0.55	4	3.23	[2.81, 3.64]	3.31	[2.90, 3.73]

Note: * LS use different prior - $\beta(0.2, 0.1)$

answers were much the same. Nevertheless, the example raises the question of whether one should try to get more efficient estimates of parameters such as ρ_q by using the model, as is attempted here, when they can be estimated independently of the model. Most of those who generate data on observable shocks for the 3G models in use seem to choose the DGP for such a series by estimating the parameters independently of the model, and we regard this as a sensible strategy.

There is also a possibility that prior selection can be done to produce outcomes that look good from a theoretical perspective even though the data might suggest that the "nice" results are more the consequence of the selected prior than the data. Since MLE does not impose such information it would seem obvious that one would always want to look at the MLEs of the parameters as well as any Bayesian estimates. A useful feature of DYNARE is that this is possible. Hence we proceed to such a comparison in the context of the current model. To start the iterations of the MLE we used the prior mean values of the Bayesian estimation procedure, but later other values were tried. Table 4.3 contains the results.

First, it should be said that the MLE of ψ_3 appears to be much the same as under Bayesian estimation methods. Second, it is noticeable that ψ_2 has the "wrong sign" for a monetary rule, and the MLE estimates of κ and σ_{y^*} look odd, whereas they are much more "reasonable" for the Bayesian results. Is this an example of the "dilemma of absurd parameter estimates" (An and Schorfheide (2005)) found when applying MLE to DSGE models? To answer that we need to ask two questions about any ML estimates. The first is whether they are perhaps giving us

some information about the "absurdity of the DSGE model". The second is what our response should be to them.

Knowing that the MLE gives "absurd parameter estimates" can often be an indicator of the fact that there are some specification problems with the model. Consequently, just suppressing these is not necessarily a wise move. It has always been the case that "wrong signs and magnitudes" can be an important part of the discovery process concerning the adequacy of a model, and so an approach which discards this information seems odd. In many ways the utility of such an action depends on why one is performing the quantitative work. The Bayesian approach is obviously attractive if one is simply seeking to assign some values to the parameters of the DSGE model and it is the only model that we are interested in entertaining. However, if we are not sure that this will be the final model to be used, we would presumably want to utilize a prior that did not exclude the possibility of a negative value for ψ_2 . Making the prior on ψ_2 be normal but centred on .1 we find that the modal estimate shrinks to .03, so that the sample certainly points towards a very weak effect of output upon interest rates in this particular specification of the interest rate rule. In general one wonders how often it is the inadequacy of the DSGE model that leads to "absurd parameter estimates" when estimated by MLE.

Moving to an appropriate response it has often been the case that those using MLE in instances like this have fixed the values of parameters that seem absurd. It is often said that Bayesian methods are a more sophisticated version of this strategy, since a range of possible values are entertained. But, if there is some identification problem associated with the parameter, it is really only the mode of the prior that matters. In that case Bayesian methods are akin to just fixing a coefficient. This is what seems to be happening with κ . The posterior and prior for this parameter turn out to be virtually identical, and so the MLE equivalent of what the Bayesian method is doing would be to constrain κ to be the modal value of the prior. When this is done the MLE estimates are now quite sensible – σ_{y^*} is 1.52 – and only a small negative value of ψ_2 (-.02) is left as being "absurd".

4.4.3 Matching the Data

We start the evaluation exercise by asking if the model can produce co-trending behavior. However, because the data is all de-meanned there is no such information provided in the estimates. The best we can do is to note that, if Δy_t had a mean of μ_y before demeaning, and dA_t^* was technology growth with mean μ_A , then we would have

$$\begin{aligned}\Delta y_t &= dA_t + \Delta \tilde{y}_t \\ &= dA_t + \frac{\mu_A}{1 - \rho_A} + \Delta \tilde{y}_t\end{aligned}$$

so that $\mu_A = (1 - \rho_A)\mu_y$. Given that $\rho_A = .53$ and $\mu_y = .68$ we would have $\mu_A = .32$. The standard deviation of dA_t must exceed .45 so that the possibility of technical regress is extraordinarily high and does not seem to be a realistic description of technology. Of course it may be that $E(\Delta y_t)$ is also influenced by non-zero means for terms of trade growth and world output growth. The former is quite small so it would be unlikely to influence μ_y much, but the latter could be substantial, although the long-run growth would need to be in technology not shared by the UK, and so it is not easy to believe that this would be substantial.

Technically there are no co-integration implications of this model as only one observable variable, y_t , is $I(1)$. But, since ρ_{y^*} is very close to unity, it would appear that y_t^* is $I(1)$, and the

Table 4.3: Maximum likelihood estimation

Parameter	Initial condition	Range	Estimate	std
ψ_1	1.50	[0 ∞]	1.73	0.38
ψ_2	0.25	$[-\infty \infty]$	-0.06	0.05
ψ_3	0.25	[0 ∞]	0.19	0.05
ρ_R	0.50	[0 1]	0.53	0.18
α	0.20	[0 1]	0.01	0.01
κ	0.50	[0 ∞]	15.64	13.19
τ	0.50	[0 1]	0.29	0.16
ρ_q	0.40	[-1 1]	-0.21	0.10
ρ_A	0.20	[0 1]	0.96	0.03
ρ_{y^*}	0.90	[0 1]	0.98	0.02
ρ_{π^*}	0.70	[0 1]	0.04	0.08
σ_R	0.50	[0 ∞]	0.58	0.21
σ_g	1.50	[0 ∞]	1.32	0.10
σ_A	1.50	[0 ∞]	0.03	0.02
σ_{y^*}	1.50	[0 ∞]	17.0	32.1
σ_{π^*}	0.55	[0 ∞]	3.16	0.54

presence of two $I(1)$ shocks A_t and y_t^* among the two $I(1)$ variables y_t and y_t^* means that there would be no co-integration between UK and foreign output.

A further relationship that also involves $I(1)$ variables is the levels version of (4.13):

$$e_t = p_t - (1 - \alpha)q_t - p_t^*. \quad (4.19)$$

Since p_t^* is $I(1)$ there is no co-integration between e_t , q_t and p_t , but, if p_t^* was observable, the real exchange rate should be co-integrated with q_t . Of course the relation would become an identity in this case which would be immediately counter-factual. Mostly this has been avoided in open economy DSGE models through the addition of a risk premium to this equation and the properties of it will determine if there is co-integration between the real exchange rate and the terms of trade. Risk premia that are close to unit root processes seem to bedevil the literature, e.g. in Adolfson et al. (2005) the risk premium shock has an AR(1) coefficient of .995.

How good is the model at producing the dynamic responses seen in the data. Tables 4–7 provide some evidence on how it does at matching univariate moments and dynamics.¹³ These show that the model fails on many dimensions – output is too volatile, there is effectively no serial correlation in output growth, and too much in exchange rate changes. *Prima facie*, given the difficulties that researchers have faced in explaining exchange rate variation, what is surprising is how well the model and data match on this dimension. But it is simple to explain why. The world inflation shock is actually a "residual" since it only appears in (4.19) and therefore its volatility can be adjusted so that the model produces an exchange rate volatility that matches the data. It's hard to see this as an "explanation". To dig deeper one might ask whether a standard

¹³ Note the peak in autocorrelation for the fourth lag of inflation. Despite what Lubik and Schorfheide (2006) state it seems as if the inflation rate for the UK may be seasonally unadjusted.

deviation of quarterly world inflation of 3.7 is realistic. Over this period most estimates would suggest something closer to 1.6.

Table 4.4: Descriptive statistics of actual data

	mean	std	Δy	π_4	R_4	Δe	Δq
Δy	0.68	0.56	1				
π_4	3.85	3.47	-0.15	1			
R_4	8.45	3.12	-0.22	0.53	1		
Δe	0.16	3.29	0.09	-0.18	0.11	1	
Δq	-0.04	1.35	-0.12	0.04	-0.02	-0.47	1

Table 4.5: Implied descriptive statistics for DSGE model variables

	mean	std	Δy	π_4	R_4	Δe	Δq	ΔA	y^*	π^*
Δy	0	0.72	1							
π_4	0	6.49	0.01	1						
R_4	0	4.95	-0.01	0.81	1					
Δe	0	4.00	-0.10	0.21	0.36	1				
Δq	0	1.35	0.07	-0.01	-0.02	-0.31	1			
ΔA	0	0.57	0.87	0.07	0.03	0.01	0.00	1		
y^*	0	3.00	-0.05	0.87	0.98	0.38	0.00	0.00	1	
π^*	0	3.78	0.09	0.20	-0.02	-0.86	0.00	0.01	-0.03	1

Turning to systems dynamics, Tables 9 and 10 compare the VARX(1) equations estimated from data with those implied by the model. The model-implied results presented below were obtained by simulating 100,000 pseudo-data points from the estimated DSGE model and then fitting the same VARX(1) as was applied to the actual data. It is readily apparent that the match between the two is quite poor.

The same can be said for the covariance matrices of the fitted VARX(1) models in Tables 9 and 10. As was mentioned earlier a way to summarize these differences is to directly compare the diagonal elements (the variances) and to determine what the correlation of the implied shocks would be from the model. Table 12 does this, with the ratio of the implied to observed standard deviations on the diagonal. Since it was assumed in the estimation that the shocks are uncorrelated the off-diagonal elements shed light on the validity of that assumption. Now there are eighty observations so that the standard error of the correlations under the assumption that they are zero would be asymptotically $\frac{1}{\sqrt{80}} = .11$, so that many of these correlations would seem to depart significantly from zero.

As we mentioned earlier a visual impression of the tracking performance of the model is useful and indeed one can see from the figures that the tracking of output growth and inflation seems to be reasonable, although the relative volatilities are very different.¹⁴ The dotted lines show

¹⁴ The standard deviation of the "explained" part of output growth is .37, i.e. some 34% less than the standard deviation of the data. This is not unexpected given that it is effectively a one-step forecast from an autoregressive

Table 4.6: Autocorrelation of actual time series

	Lag				
	1	2	3	4	5
Δy	0.37	0.26	0.33	0.26	0.23
π_4	0.07	0.33	-0.05	0.65	-0.13
R_4	0.95	0.89	0.82	0.74	0.67
Δe	0.15	-0.10	-0.03	-0.02	-0.15
Δq	-0.17	-0.02	0.05	0.05	-0.15

Table 4.7: Implied autocorrelation of DSGE model variables

	Lag				
	1	2	3	4	5
Δy	0.03	0.02	-0.01	0	0
π_4	0.78	0.69	0.65	0.63	0.61
R_4	0.97	0.95	0.92	0.90	0.87
Δe	0.46	0.28	0.20	0.16	0.14
Δq	-0.17	0.01	0	0	0
ΔA	0.19	0.04	0	0	0
y^*	0.97	0.94	0.92	0.89	0.87
π^*	0.44	0.19	0.09	0.05	0.03

what happens when shocks are taken to have zero autocorrelations (so that the dynamics are effectively much more restrictive) and it is obvious that the assumption that shocks are serially correlated is very influential in explaining the data. Without it the tracking performance of the model would be very poor. The particularly poor description of the terms of trade clearly arises because the series displays mean reversion ($\rho_q < 0$) but the assumption used in Bayesian estimation is that it does not.

Rather than looking just at growth rates etc. we decided to examine the log levels of output, prices, the nominal exchange rate and the terms of trade, as sometimes this comparison can pick up deficiencies more clearly. We have added back into the series displayed in the previous figures the sample means, and then cumulated these adjusted variables. Since we added the same values to both model and data before cumulating this has no effect on the final outcomes, but provides a picture that is more recognizable. It is pretty clear from these figures that the model fails to capture the early 1990s recession very well, something that can be seen in the growth rate information, but not as clearly. There are also difficulties in tracking the price level and the movements in sterling, although it has to be said that the 1992 crisis leading to the departure of sterling from the ERM would scarcely be explained by a model like this. The

model. In contrast the implied standard deviation for this quantity from the model is some 60% greater. This illustrates the point we made earlier that we would want the estimates of the output growth implied by the model to be constrained so that they produced a volatility that was some 60% higher than what is in the data.

Table 4.8: VARX(1) model parameter estimates from data

	Δy_t	$\pi 4_t$	$R 4_t$	Δe_t
Δy_{t-1}	0.3214 (3.0045)	1.2365 (2.0705)	0.0931 (0.5367)	-1.5965 (-2.8895)
$\pi 4_{t-1}$	-0.0173 (-0.8235)	-0.3367 (-2.8638)	0.0630 (1.8462)	-0.0656 (-0.6036)
$R 4_{t-1}$	-0.0278 (-1.1763)	0.8366 (6.3329)	0.9416 (24.5427)	0.0327 (0.2682)
Δe_{t-1}	-0.0035 (-0.1905)	-0.0965 (-0.9315)	-0.0195 (-0.6486)	0.1148 (1.1972)
Δq_t	-0.0313 (-0.7131)	0.2329 (0.9490)	-0.0574 (-0.8061)	-1.0673 (-4.7003)
constant	-0.0091 (-0.1574)	-0.0285 (-0.0884)	-0.0786 (-0.837)	-0.1070 (-0.3579)

Note: *t*-statistics in parentheses

Table 4.9: Implied VARX(1) parameter estimates from DSGE model

	Δy_t	$\pi 4_t$	$R 4_t$	Δe_t
Δy_{t-1}	0.0400 (3.97)	-0.5280 (-9.75)	-0.1283 (-9.59)	-0.1201 (-2.58)
$\pi 4_{t-1}$	-0.0423 (-20.87)	0.4583 (43.13)	0.1104 (42.03)	0.0047 (0.52)
$R 4_{t-1}$	0.0510 (18.38)	0.5230 (35.95)	0.8503 (236.33)	0.1500 (12.01)
Δe_{t-1}	0.0026 (1.28)	-0.0004 (-0.03)	0.0339 (12.71)	0.4055 (43.80)
Δq_t	0.0363 (6.56)	-0.0846 (-2.91)	-0.1153 (-16.07)	-0.9887 (-39.68)
constant	-0.0083 (-1.10)	-0.0661 (-1.67)	-0.0023 (-0.23)	0.0281 (0.83)

Note: *t*-statistics in parentheses

strong subsequent appreciation of sterling is also missed by the model but this was also a feature that was badly forecast by most commentators.

Table 4.10: Covariance matrix of residuals of VARX(1) model from data

Equation	Δy_t	$\pi 4_t$	$R 4_t$	Δe_t
Δy_t	0.2478			
$\pi 4_t$	-0.1443	7.7209		
$R 4_t$	0.0826	0.7506	0.6512	
Δe_t	0.1427	-1.6908	0.2320	6.6090

Table 4.11: Implied covariance matrix of VARX(1) residuals from DSGE model

Equation	Δy_t	$\pi 4_t$	$R 4_t$	Δe_t
Δy_t	0.5054			
$\pi 4_t$	0.1832	13.884		
$R 4_t$	-0.1424	1.2314	0.8493	
Δe_t	-0.3453	-0.9621	0.3441	10.2325

5. Operational Issues

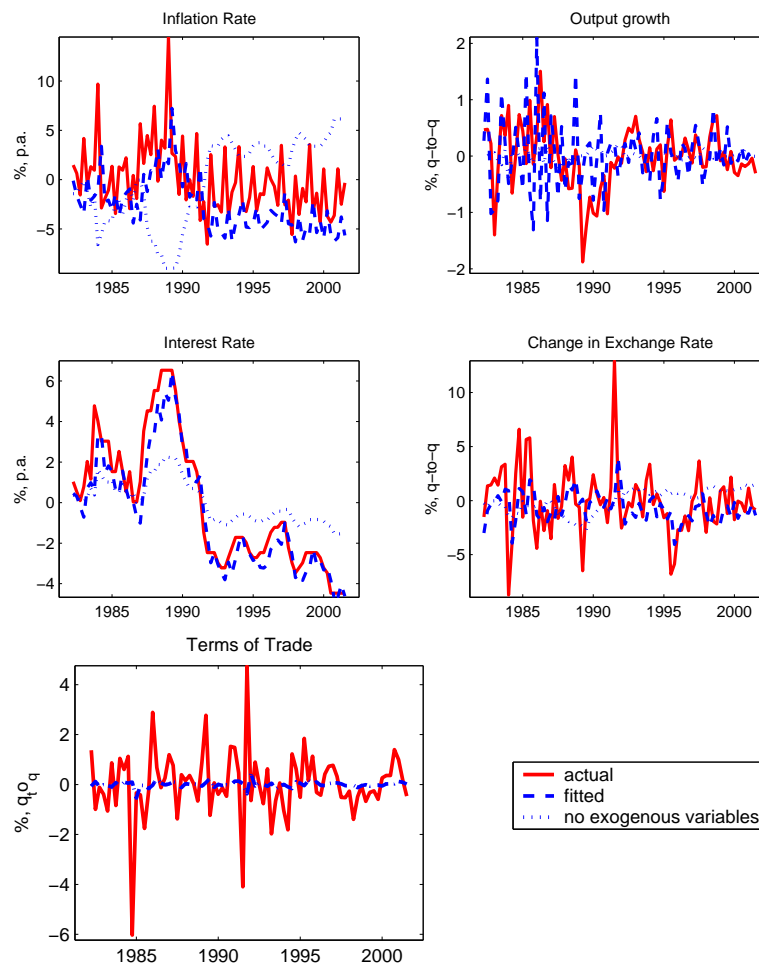
We have spent a good deal of time discussing the issues relating to DSGE models as they are used in the academic literature. One reason for this is that there is much more work that has been done on the latter class of models than on 3G or 4G models. Thus the information provided about the estimation and fit of models such as FRBUS, QPM and even BEQM is minimal. From what literature there is on the properties of 3G and 4G models it is clear that many of the problems coming up with simpler DSGE models also occur in the more complex policy models. Consequently, understanding these issues in the simpler context seems important before one comes to consider the extra complications that arise if one wants to migrate DSGE models into a central role in policy decisions. Of course DSGE (certainly CGE) models have often been used as auxiliary models, i.e. either designed to answer a limited range of specific questions at a particular point in time or as a check on conclusions reached with the primary model, but here we want to examine the difficulties that one has to face if they are to become the primary model.

5.1 Difficult and Poor Data

DSGE models in academic use range from NKPM type systems – which eschew a description of production – to RBC type models where it becomes central. 4G models tend to blend both of these aspects and such a synthesis alone makes them more complex. RBC models also have a significant difference to many NKPMs in the fact that stocks of capital and household wealth become key elements in them. NKPMs tend to resemble VAR studies, being composed largely of flow variables, and rarely include stocks (of course all models have implied stock behaviour but there is no stock-flow consistency built into most NKPMs). In 3G and 4G models the role of stocks is central to both. This has the implication that, at some point, stocks will need to be measured. Treating them as unobserved and trying to estimate them via filters, as in Smets and Wouters (2003), is rarely going to be an option when policy decisions are to be made.

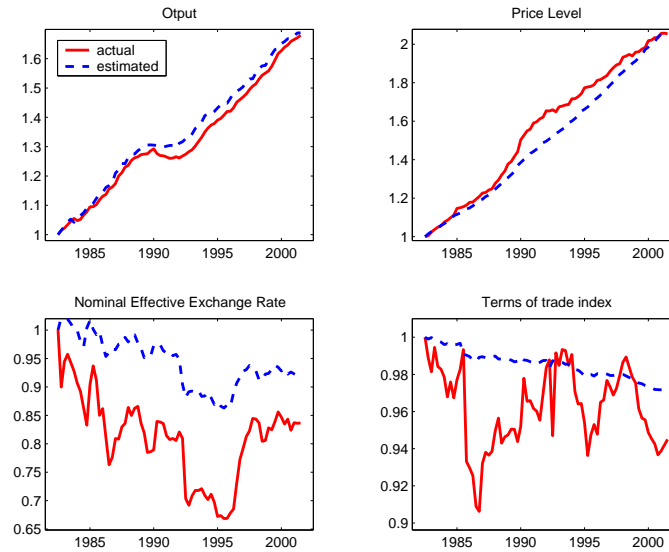
Table 4.12: Correlation matrix and variance ratios of model shocks needed to match data

Equation	Δy_t	$\pi 4_t$	$R4_t$	Δe_t
Δy_t	1.55			
$\pi 4_t$	0.42	1.28		
$R4_t$	-0.18	0.05	1.59	
Δe_t	-0.30	-0.27	0.21	1.23

Figure 4.1: Estimation results and the model predictions

Another difficulty is that all models which derive from a DSGE philosophy are guided by the idea that there is a steady state to the economy. Generally this involves assuming that ratios of certain variables are constants. But often a glance at the data doesn't seem to support this idea. One can see this difficulty in accounts of FPS and BEQM, where certain ratios seem to have changed a great deal over a sample period, and the users of the model are forced to make ad-hoc decisions about what values to allocate to the ratios. It may be reasonable to impose the last value when one uses the model to make forecasts, rather than using the average over a sample period. This is where informal methods of estimation probably pay off. Trying to estimate these models by ML or Bayesian methods seems a little odd when the data just doesn't

Figure 4.2: Estimation results in levels



reflect the specification of the model. This problem is likely to be particularly acute in transition economies. Indeed, the Negro et al. (2004) investigation suggested that some of the problems with the Smets-Wouters model came from the fact that the balanced growth restrictions imposed on the model may have led to poor forecast performance.

5.2 If the Base Model Can't Fit the Data What Do We Do?

There are many ways in which this model might not fit the data. In some instances we might be able to fit the data better by varying the model parameters, e.g. if we have some difficulties in getting cointegration between variables we might be prepared to change the model parameters to ensure that this holds even if the fit to the dynamics becomes poorer. Because many DSGE models have been fitted using data from which permanent components have been removed there has been an excessive amount of attention paid to choosing parameters that attempt to match dynamics.

If it is the case that the long-run properties look satisfactory, and one does not want to change the parameters of the base model to affect the dynamics even more, it needs to be asked how we would proceed if it was desired to get a better fit to the dynamics. To analyse this, let the base or core model in (2.4) be re-written as

$$\Delta y_t^* = (P - I)y_{t-1}^* + D_0 x_{t-1} + D_0 \Delta x_t + G_0 u_t, \quad (5.20)$$

while the data comes from

$$\Delta y_t = (F_1 + F_2 - I)y_{t-1} + F_2 \Delta y_{t-1} + N x_{t-1} + N \Delta x_t + \nu_t,$$

where both the shocks u_t and the errors ν_t are taken to be white noise. Then

$$\begin{aligned} \Delta \zeta_t = & \{(F_1 + F_2 - I)y_{t-1} + N x_{t-1}\} - \{(P - I)y_{t-1}^* + D_0 x_{t-1}\} \\ & + F_2 \Delta y_{t-1} + N \Delta x_t - D_0 \Delta x_t + \nu_t - G_0 u_t, \end{aligned} \quad (5.21)$$

where $\zeta_t = y_t - y_t^*$. Hence, after recognizing the presence of co-integration one would get

$$\begin{aligned}\Delta\zeta_t &= \alpha_D\psi_{t-1}^D - \alpha_M\psi_{t-1}^M + F_2\Delta y_{t-1} + (N - D_0)\Delta x_t \\ &\quad + \nu_t - G_0u_t,\end{aligned}$$

where "D" refers to data and "M" to model. In this expression we have distinguished the cointegrating errors ψ_t^D and ψ_t^M . Provided the long-run implications of the model are the same as the data, both of ψ_t^D and ψ_t^M would be $I(0)$; otherwise the difference between y_t and y_t^* would be $I(1)$. This points to the crucial importance of investigating the long-run implications of the model first.

Now, for the data to agree with the model, $F_2 = 0$, $N = D_0$, $\psi_{t-1}^D = \psi_{t-1}^M$. Thereupon, adding the extra regressors ψ_{t-1}^D , Δy_{t-1} , and Δx_t into the ECM generated by the model would provide a way of matching the data. The forecast from the model would then be adjusted with predicted ζ_t to get the final forecast. A more formal way of getting this same result would be to follow the approach used in 3G models in which the extra variables were introduced to match the dynamics, by thinking of adjusting y_t to meet a target y_t^* when there are adjustment costs (possibly polynomial).

The procedure above is essentially the core/non-core distinction used in working with BEQM by the Bank of England. A comprehensive description of this is available in Alvarez-Lois et al. (2005). Non-core variables such as Δy_{t-1} (or variables that are not even in the core model) are introduced in this supplementary equation with the aim of modifying the core model projections in an operational environment. The main difficulty in performing this analysis resides in the construction of y_t^* . If there are no unobservable shocks then we can construct it simply by solving (2.4) for y_t^* with a series on x_t . But when there are unobservable shocks it will generally be the case that we need to construct y_t^* from the data on y_t , which means that the distinction between data and model becomes blurred. It also has the effect that, if we were doing a dynamic simulation, the model solution y_t^* would be effectively made a function of the past and current values of the complete model predictions of y_t , i.e. there would be feedback of the non-core part of the model into the core model. The Bank of England argued very strongly against this outcome, maintaining that the core model solution has to be protected from the non-core variables. Within BEQM this strategy seemed to work well, but it should be said that there do not seem to be any unobservable shocks in that model in some of its operational uses. Whilst unobservable shocks are experimented with, these basically provide information on the responses of the model, and do not appear in simulations of the historical path. Of course, in a typical forecast environment, a path for the unobservable shocks might be provided, and so this would circumvent the feedback problem.

5.3 Flexibility

Any model used for policy has to be flexible. It needs to be adaptable so as to meet policy makers' changing preferences and to be able to incorporate their opinions and attitudes. This constraint goes to the nub of what a policy model is about. Do we want an economic story, or the ability to effectively utilize a lot of numbers? 2G models handled many pieces of information. They generally had a story, but often it was a bit confused, and it was left to their owners to make sense of it. Later generation models have a better story but perhaps less flexibility, and one needs methods to produce the desired degree of flexibility. The core/non-core distinction used in BEQM seems a good start, although it appears to be more directed towards augmenting the core model with opinions that the MPC and their advisers have, meaning that the number

of extra series being incorporated into the complete model projections seems to be relatively small.

An alternative set of auxiliary models that have become popular in recent times, and which summarize the information in many series, are factor models. It would seem worthwhile incorporating this modelling approach into the policy selection process in a more formal way than simply as a statistical model providing auxiliary information. The simplest way would appear to be via the core/non-core approach above but, as the derivation of the estimating equations shows, the error terms in such equations include the shocks of the core model, and so the question that might arise is whether the constructed factors are uncorrelated with these. Much of the literature that examines this type of set-up, i.e. a VAR augmented by factors, e.g. Forni et al. (2003), make this assumption. But it does not seem easy to believe this of factors that are constructed from a good deal of survey information, and these are often the most important determinants of movements in the measured factor. Whether this matters needs to be looked at more carefully, as we are not interested in estimating the coefficients attached to the factors themselves, but rather their contribution to a forecast. Consequently, to the extent that they can capture shocks that are unobservable, they may well provide important non-core information that can be efficiently exploited.

6. Conclusion

Macroeconometric modelling has seen a steady progression away from single equation data-dominated approaches towards relatively complete, albeit small, systems. This trend has been driven by the desire to have a strong economic perspective which links much more closely with models developed by macro theorists. It does not seem likely that this movement will be reversed. Indeed, a number of central banks have already adopted variants of 4G models as the core model in their forecasting and policy assessment tasks.

It is notable that the more information that a central bank needs to provide in terms of forecasts and explanations of policy decisions, the more there is a tendency to utilize 3G and 4G models. Thus a central bank like the Reserve Bank of Australia, which has to publish little in the way of forecasts, still works with a small 2G model – see Stone et al. (2005). Since it seems unlikely that the future will see a decline in the amount of information and explanation that is demanded, DSGE type models are almost certainly "here to stay".

Given this fact it is imperative that procedures be developed for evaluating these models as well as the processes for constructing them, just as happened with the first and second generation of models. Our feeling is that too little has been done in this regard. In particular, it is still very hard to determine how well the models track the data. Estimation of the parameters has also often been poorly described, although it has to be said that the recent enthusiasm for Bayesian methods may well be going too far in the direction of formal estimation. Often this method seems to conceal as much as it reveals. Its implementation is not entirely straightforward either. Even apart from all the difficulties facing choice of priors, the need to construct posterior densities by simulation methods often requires a considerable amount of experimentation with tuning parameters, such as the scaling factor in proposal densities in a Metropolis-Hastings algorithm. Sometimes the answers seem to be quite sensitive to these choices. Some of the priors that we used with the Lubik and Schorfheide model seemed to give odd simulated posteriors at times, and adjustments needed to be made to produce "sensible outcomes". In many instances we found that the mode of the posterior obtained by maximizing the $L(\theta) + \ln p(\theta)$ did not

coincide with the mode of the simulated posteriors even with 1 million simulations. This means that replication of Bayesian results is not straightforward and it is unclear whether what has been presented in published studies is in fact a good representation of the "true" posteriors. One needs to have more scepticism about the Bayesian results that have been obtained.

Even if we accept that the models are performing adequately in their explanation of outcomes we still have to recognize that there are many sources of information and some of these have no obvious correspondence to variables in a DSGE model, e.g. confidence measures. Yet these are rarely ignored when making policy decisions. How we adapt these models to make them "policy friendly" is a fascinating question, but it is perhaps the biggest question of them all. Good work has begun on this topic but more attention will need to be paid to it in the years to come.

References

- ADOLFSON, M., LASSEN, S., LINDE, J., AND VILLANI, M. (2005): "Forecasting Performance of an Open Economy Dynamic Stochastic General Equilibrium Model." *Sveriges Riksbank Working Paper*, (190).
- ALTUG, S. (1989): "Time to Build and Aggregate Fluctuations: Some New Evidence." *International Economic Review*, (30):889–920.
- ALVAREZ-LOIS, P., HARRISON, R., PISCATELLI, L., AND SCOTT, A. (2005): "Taking DSGE Models to the Policy Environment." In *IRFMP/IMP Conference on DSGE Modelling at Policymaking Institutions: Progress and Prospects*, Washington, DC.
- AMANO, R., MCPHAIL, K., PIORO, H., AND RENNISON, A. (2002): "Evaluating the Quarterly Projection Model: A Preliminary Investigation." *Bank of Canada Working Paper*, (2002-20).
- AN, S. AND SCHORFHEIDE, F. (2005): "Bayesian Analysis of DSGE Models." *mimeo*, University of Pennsylvania.
- BINDER, M. AND PESARAN, M. (1995): *Handbook of Applied Econometrics: Macroeconomics*, chapter Multivariate Rational Expectations Models and Macroeconomic Modelling: A Review and Some New Results. Basil Blackwell.
- BINETTE, A., MURCHISON, S., PERRIER, P., AND RENNISON, A. (2005): "An Introduction to TOTEM." *mimeo*, Bank of Canada.
- BLACK, R., LAXTON, D., ROSE, D., AND TETLOW, R. (1994): "The Bank of Canada's New Quarterly Projection Model Part 1: The Steady-State Model: SSQPM." *Bank of Canada Technical Report*, (72).
- BREUNIG, R., NAJARIAN, S., AND PAGAN, A. (2003): "Specification Testing of Markov Switching Models." *Oxford Bulletin of Economics and Statistics*, (65):703–725.
- BRUBAKK, L., HUSEBO, T., MCCAW, S., MUIR, D., NAUG, B., OLSEN, K., ROISLAND, O., SVEEN, T., AND WILHELMSSEN, B. (2005): "An Overview of the Norwegian Economy Model (NEMO)." *mimeo*, Bank of Norway.
- BRUNDY, J. AND JORGENSON, D. (1971): "Efficient Estimation of Simultaneous Equations by Instrumental Variables." *Review of Economics and Statistics*, (53):207–234.
- CANOVA, F. (1994): "Statistical Inference in Calibrated Models." *Journal of Applied Econometrics*, (9):S123–S144.
- CANOVA, F. AND SALA, L. (2005): "Learning about the Parameters and the Dynamics of DSGE Models: Identification and Estimation." In *Model Evaluation Conference*, Oslo.
- CLEMENTS, M. AND HENDRY, D. (1999): *Forecasting Non-stationary Economic Time Series*. MIT Press.
- COGLEY, T. AND NASON, J. (1993): "Impulse Dynamics and Propagation Mechanisms in a Real Business Cycle Model." *Economics Letters*, (43):77–81.
- COLLETI, D., HUNT, B., ROSE, D., AND TETLOW, R. (1996): "The Bank of Canada's New Quarterly Projection Model, Part 3: The dynamic model: QPM." *Bank of Canada Technical Report*, (75).

- DAVIDSON, J., HENDRY, D., SRBA, F., AND YEO, S. (1978): “Econometric Modelling of the Aggregate Time-Series Relationship Between Consumers’ Expenditure and Income in the United States.” *Economic Journal*, (88):661–692.
- FORNI, M., LIPPI, M., AND REICHLIN, L. (2003): “Opening the Black Box: Structural Factor Models versus Structural VARs.” *mimeo*, ECARES, University Libre de Bruxelles.
- FUHRER, J. AND OLIVEI, G. (2004): “Estimating Forward-Looking Euler Equations with GMM and Maximum Likelihood Estimators: An Optimal Instruments Approach.” In *Conference on Models and Monetary Policy*, Federal Reserve Bank Board of Governors.
- GALI, J. AND MONACHELLI, T. (2002): “Monetary Policy and Exchange Rate Volatility in a Small Open Economy.” *NBER Working Paper*, (8850).
- GRAMLICH, E. (2004): “Remarks.” In *Conference on Models and Monetary Policy*, Federal Reserve Bank Board of Governors.
- GREGORY, A., PAGAN, A., AND SMITH, G. (1993): *Models, Methods and Applications of Econometrics*, chapter Estimating Linear Quadratic Models with Integrated Processes. Blackwell.
- HARRISON, R., NIKOLOV, K., QUINN, M., RAMSAY, G., SCOTT, A., AND THOMAS, R. (2005): “The Bank of England Quarterly Model.” *Bank of England*.
- HELLIWELL, J., SPARKS, G., GORBET, F., H.T.SHAPIRO, STEWART, I., AND STEPHENSON, D. (1971): “The Structure of RDX2.” *Bank of Canada Staff Study*, (7).
- IRELAND, P. (2004): “A Method for Taking Models to the Data.” *Journal of Economic Dynamics and Control*, (28):1205–1226.
- JOHANSEN, S. (2005): “What is the Price of Maximum Likelihood.” In *Model Evaluation Conference*, Oslo.
- KAPETANIOS, G., PAGAN, A., AND SCOTT, A. (2005): “Making a Match: Combining Theory and Evidence in Policy-oriented Macroeconomic Modelling.” *Journal of Econometrics*, (forthcoming).
- KODDE, D., PALM, F., AND PFANN, G. (1990): “Asymptotic Least Squares Estimation: Efficiency Considerations and Applications.” *Journal of Applied Econometrics*, (5): 229–243.
- LAXTON, D. AND PESENTI, P. (2003): “Monetary Policy Rules for Small, Open, Emerging Economies.” *Journal of Monetary Economics*, (50):1109–1146.
- LUBIK, T. (2005): “A Simple Structural and Empirical Model of the Antipodean Transmission Mechanism.” In *DSGE Workshop, reserve Bank of New Zealand*.
- LUBIK, T. AND SCHORFHEIDE, F. (2005): *NBER Macroeconomics Annual*, chapter A Bayesian Look at New Open Economy Macroeconomics. NBER.
- LUBIK, T. AND SCHORFHEIDE, F. (2006): “Do Central Banks Respond to Exchange Rate Movements? A Structural Investigation.” *Journal of Monetary Economics*, (forthcoming).
- MAVROEIDIS, S. (2004): “Weak Identification of Forward-Looking Models in Monetary Economics.” *Oxford Bulletin of Economics and Statistics*, (66):609–635.

- MEDINA, J. AND SOTO, C. (2005): "Model for Analysis and Simulations (MAS)." In *Conference on Dynamic Stochastic General Equilibrium (DSGE) Models*, Reserve Bank of New Zealand.
- MURPHY, C. (1988): *Macroeconomic Modelling in Australia*, chapter An Overview of the Murphy Model, pages 61–68. Australian Economic Papers.
- NASON, J. AND SMITH, G. (2005): "Identifying the New Keynesian Phillips Curve." *Federal Reserve Bank of Atlanta Working Paper*, (2005-1).
- NEGRO, M. D., SCHORFHEIDE, F., SMETS, F., AND WOUTERS, R. (2004): "On the Fit and Forecasting Performance of New Keynesian Models." *Federal Reserve Bank of Atlanta Working Paper*, (2004-37).
- NICKELL, S. (1985): "Error Correction, Partial Adjustment and All That." *Oxford Bulletin of Economics and Statistics*, (47):119–130.
- POWEL, A. AND MURPHY, C. (1995): *Inside a Modern Macroeconometric Model*, number 428 of *Lecture Notes on Economics and Mathematical Systems*. Springer.
- SILVERMAN, B. (1986): *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, New York.
- SMETS, F. AND WOUTERS, R. (2003): "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area." *Journal of the European Economic Association*, (1): 1123–1175.
- STONE, A., WHEATLEY, T., AND WILKINSON, L. (2005): "A Small Model of the Australian Macroeconomy: An Update." *Reserve Bank of Australia Discussion Paper*, (2005-11).
- WATSON, M. (1993): "Measures of Fit for Calibrated Models." *Journal of Political Economy*, (101):1011–1041.

CNB WORKING PAPER SERIES

6/2006	Martin Fukač: Adrian Pagan	<i>Issues in adopting DSGE models for use in the policy process</i>
5/2006	Martin Fukač:	<i>New Keynesian model dynamics under heterogeneous expectations and adaptive learning</i>
4/2006	Kamil Dybczak: Vladislav Flek Dana Hájková Jaromír Hurník	<i>Supply-side performance and structure in the Czech Republic (1995–2005)</i>
3/2006	Aleš Krejdl:	<i>Fiscal sustainability – definition, indicators and assessment of Czech public finance sustainability</i>
2/2006	Kamil Dybczak:	<i>Generational accounts in the Czech Republic</i>
1/2006	Ian Babetskii:	<i>Aggregate wage flexibility in selected new EU member states</i>
<hr/>		
14/2005	Stephen G. Cecchetti:	<i>The brave new world of central banking: The policy challenges posed by asset price booms and busts</i>
13/2005	Robert F. Engle: Jose Gonzalo Rangel	<i>The spline GARCH model for unconditional volatility and its global macroeconomic causes</i>
12/2005	Jaromír Beneš: Tibor Hlédik Michael Kumhof David Vávra	<i>An economy in transition and DSGE: What the Czech national bank's new projection model needs</i>
11/2005	Marek Hlaváček: Michael Koňák Josef Čada	<i>The application of structured feedforward neural networks to the modelling of daily series of currency in circulation</i>
10/2005	Ondřej Kameník:	<i>Solving SDGE models: A new algorithm for the sylvester equation</i>
9/2005	Roman Šustek:	<i>Plant-level nonconvexities and the monetary transmission mechanism</i>
8/2005	Roman Horváth:	<i>Exchange rate variability, pressures and optimum currency area criteria: Implications for the central and eastern european countries</i>
7/2005	Balázs Égert: Luboš Komárek	<i>Foreign exchange interventions and interest rate policy in the Czech Republic: Hand in glove?</i>
6/2005	Anca Podpiera: Jiří Podpiera	<i>Deteriorating cost efficiency in commercial banks signals an increasing risk of failure</i>
5/2005	Luboš Komárek: Martin Melecký	<i>The behavioural equilibrium exchange rate of the Czech koruna</i>
4/2005	Kateřina Arnoštová: Jaromír Hurník	<i>The monetary transmission mechanism in the Czech Republic (evidence from VAR analysis)</i>
3/2005	Vladimír Benáček: Jiří Podpiera Ladislav Prokop	<i>Determining factors of Czech foreign trade: A cross-section time series perspective</i>
2/2005	Kamil Galuščák: Daniel Münich	<i>Structural and cyclical unemployment: What can we derive from the matching function?</i>
1/2005	Ivan Babouček: Martin Jančar	<i>Effects of macroeconomic shocks to the quality of the aggregate loan portfolio</i>

10/2004	Aleš Bulíř: Kateřina Šmídková	<i>Exchange rates in the new EU accession countries: What have we learned from the forerunners</i>
9/2004	Martin Cincibuch: Jiří Podpiera	<i>Beyond Balassa-Samuelson: Real appreciation in tradables in transition countries</i>
8/2004	Jaromír Beneš: David Vávra	<i>Eigenvalue decomposition of time series with application to the Czech business cycle</i>
7/2004	Vladislav Flek, ed.:	<i>Anatomy of the Czech labour market: From over-employment to under-employment in ten years?</i>
6/2004	Narcisa Kadlčáková: Joerg Keplinger	<i>Credit risk and bank lending in the Czech Republic</i>
5/2004	Petr Král:	<i>Identification and measurement of relationships concerning inflow of FDI: The case of the Czech Republic</i>
4/2004	Jiří Podpiera:	<i>Consumers, consumer prices and the Czech business cycle identification</i>
3/2004	Anca Pruteanu:	<i>The role of banks in the Czech monetary policy transmission mechanism</i>
2/2004	Ian Babetskii:	<i>EU enlargement and endogeneity of some OCA criteria: Evidence from the CEECs</i>
1/2004	Alexis Derviz: Jiří Podpiera	<i>Predicting bank CAMELS and S&P ratings: The case of the Czech Republic</i>

12/2003	Tibor Hlédik:	<i>Modelling the second-round effects of supply-side shocks on inflation</i>
11/2003	Luboš Komárek: Zdeněk Čech Roman Horváth	<i>ERM II membership – the view of the accession countries</i>
10/2003	Luboš Komárek: Zdeněk Čech Roman Horváth	<i>Optimum currency area indices – how close is the Czech Republic to the eurozone?</i>
9/2003	Alexis Derviz: Narcisa Kadlčáková Lucie Kobzová	<i>Credit risk, systemic uncertainties and economic capital requirements for an artificial bank loan portfolio</i>
8/2003	Tomáš Holub: Martin Čihák	<i>Price convergence: What can the Balassa–Samuelson model tell us?</i>
7/2003	Vladimír Bezděk: Kamil Dybczak Aleš Krejdl	<i>Czech fiscal policy: Introductory analysis</i>
6/2003	Alexis Derviz:	<i>FOREX microstructure, invisible price determinants, and the central bank's understanding of exchange rate formation</i>
5/2003	Aleš Bulíř:	<i>Some exchange rates are more stable than others: Short-run evidence from transition countries</i>
4/2003	Alexis Derviz:	<i>Components of the Czech koruna risk premium in a multiple-dealer FX market</i>
3/2003	Vladimír Benáček: Ladislav Prokop Jan Á. Víšek	<i>Determining factors of the Czech foreign trade balance: Structural issues in trade creation</i>

2/2003	Martin Čihák: Tomáš Holub	<i>Price convergence to the EU: What do the 1999 ICP data tell us?</i>
1/2003	Kamil Galuščák: Daniel Münich	<i>Microfoundations of the wage inflation in the Czech Republic</i>
4/2002	Vladislav Flek: Lenka Marková Jiří Podpiera	<i>Sectoral productivity and real exchange rate appreciation: Much ado about nothing?</i>
3/2002	Kateřina Šmídková: Ray Barrell Dawn Holland	<i>Estimates of fundamental real exchange rates for the five EU pre-accession countries</i>
2/2002	Martin Hlušek:	<i>Estimating market probabilities of future interest rate changes</i>
1/2002	Viktor Kotlán:	<i>Monetary policy and the term spread in a macro model of a small open economy</i>

CNB RESEARCH AND POLICY NOTES

5/2005	Jan Stráský:	<i>Optimal forward-looking policy rules in the quarterly projection model of the Czech National Bank</i>
4/2005	Vít Bárta:	<i>Fulfilment of the Maastricht inflation criterion by the Czech Republic: Potential costs and policy options</i>
3/2005	Helena Šůvová: Eva Kozelková David Zeman Jaroslava Bauerová	<i>Eligibility of external credit assessment institutions</i>
2/2005	Martin Čihák: Jaroslav Heřmánek	<i>Stress testing the Czech banking system: Where are we? Where are we going?</i>
1/2005	David Navrátil: Viktor Kotlán	<i>The CNB's policy decisions – Are they priced in by the markets?</i>
4/2004	Aleš Bulíř:	<i>External and fiscal sustainability of the Czech economy: A quick look through the IMF's night-vision goggles</i>
3/2004	Martin Čihák:	<i>Designing stress tests for the Czech banking system</i>
2/2004	Martin Čihák:	<i>Stress testing: A review of key concepts</i>
1/2004	Tomáš Holub:	<i>Foreign exchange interventions under inflation targeting: The Czech experience</i>
2/2003	Kateřina Šmídková:	<i>Targeting inflation under uncertainty: Policy makers' perspective</i>
1/2003	Michal Skořepa: Viktor Kotlán	<i>Inflation targeting: To forecast or to simulate?</i>

CNB ECONOMIC RESEARCH BULLETIN

August 2006	<i>Research priorities and central banks</i>
November 2005	<i>Financial stability</i>
May 2005	<i>Potential output</i>
October 2004	<i>Fiscal issues</i>
May 2004	<i>Inflation targeting</i>
December 2003	<i>Equilibrium exchange rate</i>

Czech National Bank
Economic Research Department
Na Příkopě 28, 115 03 Praha 1
Czech Republic
phone: +420 2 244 12 321
fax: +420 2 244 14 278
<http://www.cnb.cz>
e-mail: research@cnb.cz