.

To my parents, brother, grandparents, and my wife Lenka.

Acknowledgements

I thank all the people influencing the work that follows. In particular, I am much indebted to my supervisor Michal Kejak, Adrian Pagan, Sergey Slobodyan, Kristoffer Nimark, Peter Sinclair, Evžen Kočenda, Gunnar Bardsen, Mardi Dungey, Sophocles Mavroeides, Jim Nason, Jan Kmenta, and Peter Tóth for contributing ideas and providing me with valuable insights, suggestions, and comments throughout the work on the dissertation.

I am also grateful for being a student of Osvald Vasiček, Michal Kejak, Doug Laxton, and Adrian Pagan, who influenced me the most on the journey of exploring economics, econometrics, and policy oriented research. They taught me a passion for the profession.

Contents

Preface

This dissertation addresses three topics: (i) a recent boom in DSGE model popularity and their use for monetary policy, (ii) imperfect knowledge and expectations heterogeneity in New Keynesian policy models, and (iii) the proximity of surveyed inflation expectations to market expectations.

In the first chapter, co-authored with Adrian Pagan, we discuss three basic issues in adopting DSGE models for policy making: model design, matching the data, and operational requirements. We begin with a general discussion of the structure of dynamic stochastic general equilibrium (DSGE) models where we investigate issues such as (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 first chapter 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 chapter is devoted to a variety of topics involving the estimation and evaluation of DSGE models.

In the second chapter, I examine the role of monetary policy in a heterogeneous expectations environment. I use a New Keynesian business cycle model as the experiment laboratory. I assume that a central bank, and private economic agents (households and producing firms) have imperfect and heterogeneous information about the economy, and as a consequence, they disagree in their views on its future development. All agents are adaptively learning. Measured by the central bank's expected loss, the two major findings are: (i) policy, which is efficient under homogeneous expectations is not efficient under the heterogeneous expectations; (ii) in the short- and mid- run, too, inflation responsive policy increases the volatility of inflation and output, but in the long-run, inflation responsive policy improves economic volatility.

In the third chapter, I develop a simple econometric tool to test for the similarity between surveyed inflation expectations and market expectations on the Czech Interbank money market. Monthly data on the inflation expectations of financial analysts in the Czech Republic exhibit a tendency for permanent bias and ineffectiveness. But the surveyed data seem to be a good predictor for the 1-year Prague interbank offer rate. Using the methodology based on an adjusted Fisher rule – relating inflation expectations, nominal and real interest rates –, I find that the difference between the surveyed and market expectations is not statistically significant.

Chapter 1

Issues in Adopting DSGE Models for Use in the Policy Process

This chapter is co-authored with Adrian Pagan and has been published as CAMA Working paper No. 9/2006. An updated version has also been published as the Czech National Bank Working Paper No. 6/2006.

1.1 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 U.K. 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 shall 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.

1.2 Some preliminaries

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

$$B_0 y_t = B_1 y_{t-1} + D x_t + C E_t y_{t+1} + G u_t, \tag{1.1}$$

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 side step that issue until later. By observable we shall mean that the shocks can be recovered from a statistical model. By unobservable we shall mainly mean that the shocks are defined by the economic model. Later, however, we shall 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 shall call tracking shocks.

We shall 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

$$y_t = \alpha k_t + (1-\alpha) l_t + u_{1t}, \tag{1.2}$$

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

$$\Delta y_t = \alpha \Delta k_t + (1-\alpha)\Delta l_t + e_{1t}, \tag{1.3}$$

where $\Delta u_{1t} = e_{1t}$. It is the fact that u_{1t} is a structural shock that enables us to eliminate it by the differencing of the variables in the structural relation. One would then just re-write the system in (1.1) 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.

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

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 $\left[\begin{array}{cc} B_0^p - B_1^p - C^p & -D^p \end{array} \right]$ 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 relations. Parameters such as B_0, B_1, D, and C enter the long-run, but B_2 does not. Accordingly, while 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 $\xi_t = y_t - P y_{t-1}$, which is then substituted to obtain

$$\begin{aligned} B_0(\xi_t + P y_{t-1}) &= B_1 y_{t-1} + D x_t + C E_t(\xi_{t+1} + P y_t) + G u_t \\ &= B_1 y_{t-1} + D x_t + C E_t(\xi_{t+1} + P(\xi_t + P y_{t-1})) + G u_t \\ &= B_1 y_{t-1} + D x_t + C E_t(\xi_{t+1}) + C P \xi_t + C P^2 y_{t-1} + G u_t, \end{aligned}$$

so that we need $B_0 P - B_1 - C P^2 = 0$ to eliminate the y_{t-1} term and to produce

$$B_0 \xi_t = C E_t(\xi_{t+1}) + D x_t + C P \xi_t + G u_t.$$

This then implies

$$\begin{aligned} \xi_t &= (B_0 - CP)^{-1} C E_t(\xi_{t+1}) + (B_0 - CP)^{-1} D x_t + (B_0 - CP)^{-1} G u_t \\ &= \Pi_1 E_t \xi_{t+1} + \Pi_2 x_t + \Pi_3 u_t, \end{aligned}$$

and the solution to the latter would be

$$\xi_t = \sum \Pi_1^j E_t(\Pi_2 x_{t+j} + \Pi_3 u_{t+j}),$$

Thus, the solution is of the form

$$y_t = P y_{t-1} + \sum \Pi_1^j (\Pi_2 E_t x_{t+j} + \Pi_3 E_t u_{t+j}),$$

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

$$y_t = Py_{t-1} + D_0 x_t + G_0 u_t. \tag{1.4}$$

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.

Now, 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 a 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 see later, there seems 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 (1.4) can be written as

$$\begin{aligned} \Delta y_t &= (P - I)y_{t-1} + D_0 x_{t-1} + D_0 \Delta x_t + G_0 u_t \\ &= \Psi z_{t-1} + D_0 \Delta x_t + G_0 u_t \\ &= \alpha_1 \beta' z_{t-1} + D_0 \Delta x_t + G_0 u_t, \end{aligned} \tag{1.5}$$

where $z_t = \begin{bmatrix} y_t \\ x_t \end{bmatrix}$, $\Psi = \begin{bmatrix} P - I & D_0 \end{bmatrix}$ 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 (1.5), 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 not only depends upon the mean of the growth rate in x_t, but also upon the complete set of parameters in the model. Shifts in μ_y, which will often be the source of forecast failures, can occur either because of changes in μ_x or 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-meaned" 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 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 become more complex is if the observed data is an I(1) process such 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

$$\Delta y_t^0 = \Delta y_t + \Delta a_t. \tag{1.6}$$

Afterwards, (1.6) combines with (1.4) 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.

1.3 Base Policy Model Design

1.3.1 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, 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, Sparks, Gorbet, Shapiro, Stewart and Stephenson (1971) - and the MPS model[2] at the Fed, introduced much stronger supply side features and also moved towards deriving some of

[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.

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. However, 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, Hendry, Srba and Yeo (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 dis-satisfaction 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. 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 shall 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. 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.[3]

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, Laxton, Rose and Tetlow (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, Hunt, Rose and Tetlow (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.[4] In

[3]The model was more fully described in Powel and Murphy (1995).

[4]Although there is an isomorphism between ECMs and the strategy of maximizing some linear

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 a 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 financial variables. In practice, a combination of both backward and forward looking elements were incorporated into the 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 the 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.

1.3.2 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, Murchison, Perrier and Rennison (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, Husebo, McCaw, Muir, Naug, Olsen, Roisland, Sveen and Wilhelmsen (2005)); and The New Area Wide Model (NAWM) at the European Central Bank.

These new models have two striking general characteristics. First, they feature con-

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.

siderable 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 with 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.

1.3.3 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 U.K., 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.

1.4 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, and evidence from previous micro and macroeconometric work. 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 should not 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 shall 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 the 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, the procedures for assessing how well these models match the data.

1.4.1 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 shall 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 then we must have a failure of identification since, when there is identification, Kodde, Palm and Pfann (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 $z_t = \left[\begin{array}{ccc} y_t & y_{t-1} & x_t \end{array} \right]$. 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 is clear that to estimate the parameters of these equations we shall 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 shall 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.[5]

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:

$$\begin{aligned} \pi_t &= \eta_1 \pi_{t-1} + \eta_2 E_t(\pi_{t+1}) + \eta_3 \xi_t + u_{St} && (1.7) \\ \xi_t &= \eta_4 \xi_{t-1} + \eta_5 E_t(\xi_{t+1}) + \eta_6 (r_t - E_t(\pi_{t+1})) + u_{Dt} && (1.8) \\ r_t &= \eta_7 r_{t-1} + \eta_8 \xi_t + \eta_9 \pi_t + u_{It}, && (1.9) \end{aligned}$$

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 (1.4) 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 increases 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 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.[6] Now what happens if there is serial correlation in the shocks of the first two

[5] This assumes shocks are white noise. If they exhibit serial correlation, a different estimator is needed.

[6] 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. However, 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

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 $\eta_j, j = 1, ..., 9$. 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.

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. The elimination of the serial correlation in the second equation introduces terms $\frac{\rho(\varphi+v)(1-\zeta\beta)(1-\zeta)}{(1+\beta\omega)\zeta}\xi_{t-1}$ 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 the 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.

in the NKPM are available in Mavroeidis (2004) and Nason and Smith (2005).

As mentioned previously, other restrictions may be imposed upon the η_j by either theoretical considerations or as an outcome of the chosen functional form. Thus, as we saw in (1.3), the assumption that the technology shock is permanent and that the production function is 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).[7]

1.4.2 Estimation Techniques

There are various formal methods of estimation, differentiated largely by the extent of 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 the MLE estimation would be quite difficult.

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 shall 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

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

strongly positive since the prior was set up so that a negative value could not be found.[8]

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, p 1140) 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 centered 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.

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 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

[8]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 1.1: An example of a too-many-unobservables model estimation. Estimates of ρ_1 and 90% confidence interval

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

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 are used as instruments to form 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 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.

1.4.3 Evaluation Issues

There is 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, Nikolov, Quinn, Ramsay, Scott and Thomas (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.

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 (1.4), one can envisage computing the one-step ahead forecast for $y_t, \hat{y}_t$, as

$$\hat{y}_t = Py_{t-1} + D_0 x_t + G_0^{+}\Phi(y_{t-1} - D_0 x_{t-1}), \tag{1.10}$$

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, McPhail, Pioro and Rennison (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, Najarian and Pagan (2003) for this use.

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.[9] 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 (1.10), 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. However, 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

[9]One problem with such moment comparisons occurs 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.

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 $\left[\frac{var(y_t^*)}{var(y_t)}\right]^{1/2}$.

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 only to 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 (1.10). In many ways, the problem is the same as a dynamic factor extraction when one wishes only to 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 a spectra.

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-meaned, 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-meaned, 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 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, Pagan and Scott (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, Lassen, Linde and Villani (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 Negro, Schorfheide, Smets and Wouters (2004). They 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.[10] 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

[10] In fact they eventually use a VAR rather than a VECM because the model implied co-integrating restrictions, which 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.

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 a 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 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 $G_0 var(u_t) G_0'$. Because $var(u_t)$ 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.

1.4.4 An Example: A Simple, Open Economy Model

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 (forthcoming) (LS), which is a simplified version of Gali and Monachelli (2002).

The model has four equations. An open economy IS curve

$$\begin{aligned}
\tilde{y}_t &= E_t \tilde{y}_{t+1} - [\tau + \alpha(2-\alpha)(1-\tau)](R_t - E_t \pi_{t+1}) - \alpha[\tau + \alpha(2-\alpha)(1-\tau)]\rho_q \Delta q_t \\
&\quad + \rho_A dA_t - \alpha(2-\alpha)\frac{1-\tau}{\tau}(1-\rho_{y^*})\tilde{y}_t^*, \qquad (1.11) \\
0 &< \alpha < 1, \tau^{-1} > 0;
\end{aligned}$$

an open economy Phillips curve

$$\begin{aligned}\pi_t &= \beta E_t \pi_{t+1} - \alpha(1-\beta\rho_q)\Delta q_t + \frac{\kappa}{\tau + \alpha(2-\alpha)(1-\tau)}\tilde{y}_t \\ &+ \frac{\kappa\alpha(2-\alpha)(1-\tau)}{\tau[\tau+\alpha(2-\alpha)(1-\tau)]}\tilde{y}_t^*; \end{aligned} \tag{1.12}$$

an exchange rate equation

$$\Delta e_t = \pi_t - (1-\alpha)\Delta q_t - \pi_t^*; \tag{1.13}$$

and the policy rule

$$R_t = \rho_R R_{t-1} + (1-\rho_R)\left[\psi_1 \pi_t + \psi_2 \tilde{y}_t + \psi_3 \Delta e_t\right] + \varepsilon_t^R. \tag{1.14}$$

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

$$\Delta q_t = \rho_q \Delta q_{t-1} + \varepsilon_t^q. \tag{1.15}$$

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:

$$\begin{aligned} dA_t &= \rho_a dA_{t1} + \varepsilon_t^a & (1.16)\\ \tilde{y}_t^* &= \rho_{y^*}\tilde{y}_{t-1}^* + \varepsilon_t^{y^*}, & (1.17)\\ \pi_t^* &= \rho_{\pi^*}\pi_{t-1}^* + \varepsilon_t^{\pi^*}, & (1.18)\end{aligned}$$

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 $lnA_t = \ln A_{t-1} + dA_t$ 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. $\beta = .99$

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 $R4_t$, quarterly exchange rate change, Δe, and terms of trade growth, Δq. The series were de-meaned, and the data are related to the model variables as follows:

$$[\Delta y_t, \pi 4_t, R4_t, \Delta e, \Delta q] = [\Delta\tilde{y}_t + dA_t, 4\times\pi_t, 4\times R_t, \Delta e, \Delta q].$$

LS estimate the model by Bayesian methods. Table 1.2 contains estimates of the parameters.[11]

Table 1.2: Prior distributions and parameter estimation results

	Prior distribution			Estimates		LS's estimates	
	Density	Mean	std	Mean	90% Interval	Mean	90% Interval
ψ_1	Gamma	1.50	0.20	1.31	[1.04, 1.57]	1.17	[0.92,1.41]
ψ_2	Gamma	0.25	0.13	0.49	[0.26, 0.72]	0.40	[0.15,0.63]
ψ_3	Gamma	0.25	0.13	0.15	[0.07, 0.21]	0.12	[0.07,0.18]
ρ_R	Beta	0.50	0.20	0.79	[0.74, 0.84]	0.68	[0.60, 0.77]
α	Beta	0.20	0.05	0.09	[0.05, 0.13]		
r	-	2.50	calibrated				
κ	Gamma	0.50	0.10	0.74	[0.55, 0.93]	0.65	
τ	Gamma	0.50	0.20	0.25	[0.16, 0.34]		
ρ_q	Beta	-0.20	0.20	-0.17	[-0.30, -0.05]		
ρ_A	Beta	0.20	0.05	0.20	[0.12, 0.28]		
ρ_{y^*}	Beta	0.90	0.05	0.97	[0.95, 0.99]		
ρ_{π^*}	Beta	0.80	0.10	0.45	[0.33, 0.58]		
σ_R	InvGamma	0.50	4	0.28	[0.23, 0.33]		
σ_g	InvGamma	1.50	4	1.32	[1.16, 1.47]		
σ_A	InvGamma	1.50	4	0.56	[0.47, 0.65]		
σ_{y^*}	InvGamma	1.50	4	0.68	[0.34, 1.06]		
σ_{π^*}	InvGamma	0.55	4	3.34	[2.89, 3.74]		

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 1.2.

Replicating LS's results, we performed a robustness check on their conclusion about whether the policy rule (1.14) 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 the mode between 0.10 and .20, and different from zero.

Now, if one just fitted (1.15) 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 centered on zero as the prior. However, the answers were much the same. Nevertheless, the example raises the question of whether one should try to get more efficient estimates

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

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 with 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 1.3 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" that 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 supressing 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 centered 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 "absurb 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".

Table 1.3: Maximum likelihood estimation

Parameter	Initial condition	Range	Estimate	std
ψ_1	1.50	$[0\ \infty]$	1.73	0.38
ψ_2	0.25	$[-\infty\ \infty]$	-0.06	0.05
ψ_3	0.25	$[0\ \infty]$	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	$[0\ \infty]$	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	$[0\ \infty]$	0.58	0.21
σ_g	1.50	$[0\ \infty]$	1.32	0.10
σ_A	1.50	$[0\ \infty]$	0.03	0.02
σ_{y^*}	1.50	$[0\ \infty]$	17.0	32.1
σ_{π^*}	0.55	$[0\ \infty]$	3.16	0.54

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-meaned, 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 U.K., 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 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 (1.13):

$$e_t = p_t - (1-\alpha)q_t - p_t^*. \tag{1.19}$$

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.[12] 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 (1.19), and therefore its volatility can be adjusted so that the model produces an exchange rate volatility that matches the data. It is hard to see this as an "explanation". To dig deeper one might ask whether a standard deviation of quarterly world inflation of 3.7 is realistic. Over this period, most estimates would suggest something closer to 1.6.

Table 1.4: Descriptive statistics of actual data

	mean	std	Δy	$\pi 4$	$R4$	Δe	Δq
Δy	0.68	0.56	1				
$\pi 4$	3.85	3.47	-0.15	1			
$R4$	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

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 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

[12] Note the peak in autocorrelation for the fourth lag of inflation. Despite what Lubik and Schorfheide (forthcoming) state it seems as if the inflation rate for the U.K. may be seasonally unadjusted.

Table 1.5: Implied Descriptive statistics for DSGE Model Variables

	mean	std	Δy	$\pi 4$	$R4$	Δe	Δq	ΔA	y^*	π^*
Δy	0	0.72	1							
$\pi 4$	0	6.49	0.01	1						
$R4$	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

Table 1.6: Autocorrelation of actual time series

	Lag				
	1	2	3	4	5
Δy	0.37	0.26	0.33	0.26	0.23
$\pi 4$	0.07	0.33	-0.05	0.65	-0.13
$R4$	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 1.7: Implied Autocorrelation of DSGE Model Variables

	Lag				
	1	2	3	4	5
Δy	0.03	0.02	-0.01	0	0
$\pi 4$	0.78	0.69	0.65	0.63	0.61
$R4$	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

ratio of the implied to observed standard deviations on the diagonal. Since it was assumed in 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.

Table 1.8: VARX(1) Model Parameter estimates from data

	Δy_t	$\pi 4_t$	$R4_t$	Δe_t
Δy_{t-1}	0.3214	1.2365	0.0931	-1.5965
	(*3.0045*)	(*2.0705*)	(*0.5367*)	(*-2.8895*)
$\pi 4_{t-1}$	-0.0173	-0.3367	0.0630	-0.0656
	(*-0.8235*)	(*-2.8638*)	(*1.8462*)	(*-0.6036*)
$R4_{t-1}$	-0.0278	0.8366	0.9416	0.0327
	(*-1.1763*)	(*6.3329*)	(*24.5427*)	(*0.2682*)
Δe_{t-1}	-0.0035	-0.0965	-0.0195	0.1148
	(*-0.1905*) (	*-0.9315*)	(*-0.6486*)	(*1.1972*)
Δq_t	-0.0313	0.2329	-0.0574	-1.0673
	(*-0.7131*)	(*0.9490*)	(*-0.8061*)	(*-4.7003*)
constant	-0.0091	-0.0285	-0.0786	-0.1070
	(*-0.1574*)	(*-0.0884*)	(*-0.837*)	(*-0.3579*)

Note: t-statistics in parentheses

Table 1.9: Implied VARX(1) Parameter estimates from DSGE Model

	Δy_t	$\pi 4_t$	$R4_t$	Δe_t
Δy_{t-1}	0.0400	-0.5280	-0.1283	-0.1201
	(*3.97*)	(*-9.75*)	(*-9.59*)	(*-2.58*)
$\pi 4_{t-1}$	-0.0423	0.4583	0.1104	0.0047
	(*-20.87*)	(*43.13*)	(*42.03*)	(*0.52*)
$R4_{t-1}$	0.0510	0.5230	0.8503	0.1500
	(*18.38*)	(*35.95*)	(*236.33*)	(*12.01*)
Δe_{t-1}	0.0026	-0.0004	0.0339	0.4055
	(*1.28*) (	*-0.03*)	(*12.71*)	(*43.80*)
Δq_t	0.0363	-0.0846	-0.1153	-0.9887
	(*6.56*)	(*-2.91*)	(*-16.07*)	(*-39.68*)
constant	-0.0083	-0.0661	-0.0023	0.0281
	(*-1.10*)	(*-1.67*)	(*-0.23*)	(*0.83*)

Note: t-statistics in parentheses

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.[13] The

[13]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 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 1.10: Covariance Matrix of Residuals of VARX(1) Model from Data

Equation	Δy_t	$\pi 4_t$	$R4_t$	Δe_t
Δy_t	0.2478			
$\pi 4_t$	-0.1443	7.7209		
$R4_t$	0.0826	0.7506	0.6512	
Δe_t	0.1427	-1.6908	0.2320	6.6090

Table 1.11: Implied Covariance Matrix of VARX(1) residuals from DSGE Model

Equation	Δy_t	$\pi 4_t$	$R4_t$	Δe_t
Δy_t	0.5054			
$\pi 4_t$	0.1832	13.884		
$R4_t$	-0.1424	1.2314	0.8493	
Δe_t	-0.3453	-0.9621	0.3441	10.2325

dotted lines show what happens when shocks are taken to have zero autocorrelations (so that 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, 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 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 strong subsequent appreciation of sterling is also missed by the model, but this was also a feature that was badly forecasted by most commentators.

1.5 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

Table 1.12: 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

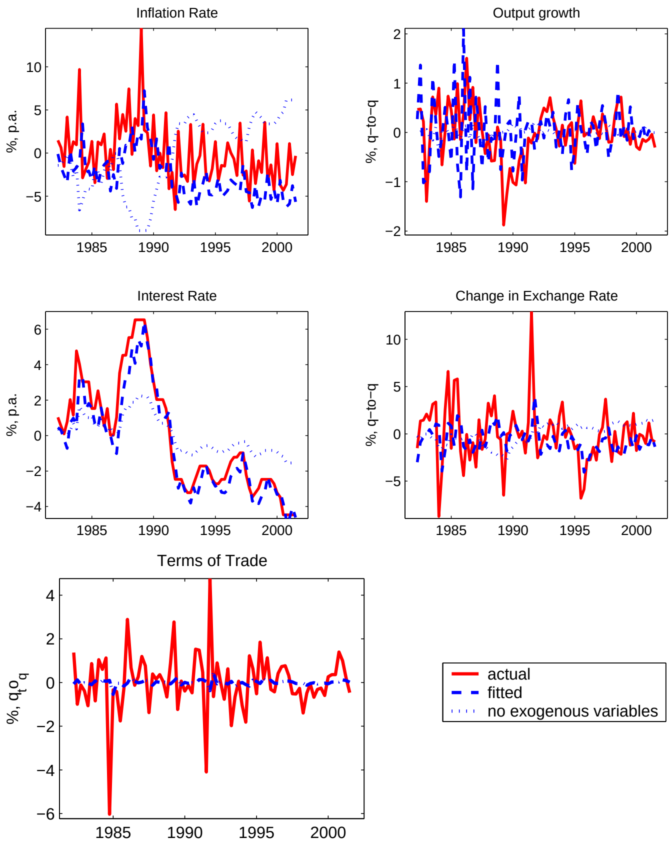


Figure 1.1: Estimation results and the model predictions

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.

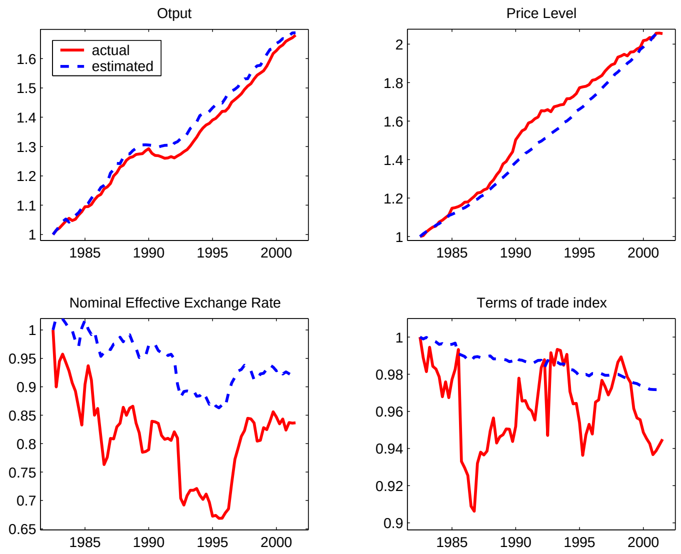


Figure 1.2: Estimation results in levels

1.5.1 Difficult and Poor data

DGSE 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.

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 does not 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 does not 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.

1.5.2 What Do We Do If the Base Model Does Not Fit the Data?

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 co-integration between variables, we might be prepared to change the model parameters to ensure that this holds even if the fit to dynamics become 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 (1.4) be re-written as

$$\Delta y_t^* = (P - I)y_{t-1}^* + D_0 x_{t-1} + D_0 \Delta x_t + G_0 u_t, \tag{1.20}$$

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 v_t are taken to be white noise. Then

$$\begin{aligned} \Delta\zeta_t &= \{(F_1 + F_2 - I)y_{t-1} + N x_{t-1}\} - \{(P - I)y_{t-1}^* + D_0 x_{t-1}\} \\ &\quad + F_2 \Delta y_{t-1} + N \Delta x_t - D_0 \Delta x_t + \nu_t - G_0 u_t, \end{aligned} \tag{1.21}$$

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

$$\begin{aligned} \Delta\zeta_t &= \alpha_D \psi_{t-1}^D - \alpha_M \psi_{t-1}^M + F_2 \Delta y_{t-1} + (N - D_0)\Delta x_t \\ &\quad + \nu_t - G_0 u_t, \end{aligned}$$

where "D" refers to the data and "M" to the model. In this expression, we have distinguished the co-integrating errors ψ_t^D and ψ_t^M. Provided the long-run implications of the model are the same as the data, both 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, $F_2 = 0, N = D_0$, and $\psi_{t-1}^D = \psi_{t-1}^M$. Thereupon, adding the extra regressors $\psi_{t-1}^D, \Delta y_{t-1}$, and Δx_t into the ECM generated by the model, provides 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 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, Harrison, Piscatelli and Scott (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 (1.4) 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.

1.5.3 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 heart 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 to be 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 as in Forni, Lippi and Reichlin (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.

1.6 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 the 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, Wheatley and Wilkinson (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 it 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 Methropolis-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 the 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.

Chapter 2

Heterogeneous Expectations, Adaptive Learning, and Forward-Looking Monetary Policy

2.1 Introduction

Empirical literature – eg., Mishkin and Schmidt-Hebbel (2006) – provides evidence that, in inflation targeting regimes, long-term inflation expectations are anchored to the target. On the other hand, short-term inflation expectations (1 to 2 year-horizon expectations, which is the time horizon of effective monetary policy) are typically more volatile, and there is a high degree of expectations heterogeneity – eg., Mankiw and Wolfers (2003). Central bankers face a problem of anchoring these short-term expectations. This raises the question of what are the effects of private agents having different expectations than a central bank.

The contribution of this chapter is in its focus on short-term transitional dynamics, from imperfect and heterogeneous expectations to a perfect homogeneous (rational) expectations environment. There are two groups of agents that establish a simple heterogeneous expectations environment. The private agents (households and firms) and the central bank are assumed to have imperfect information, and they form different expectations. I am interested in how forward looking monetary policy can influence the speed and volatility of the convergence process in a 1 to 10 year time horizon in such an environment. Orphanides and Williams (2003) and Ferrero (2007) show that a central bank, operating in an imperfect knowledge, but a homogenous expectations environment, has the potential to improve the process. I extend this problem to the heterogeneous expectations case. I build on Honkapohja and Mitra (2005) who show the conditions under which a heterogeneous expectations economy can converge to a stationary, rational expectations equilibrium (REE). I use a numerical analysis to study the convergence process under these conditions.

I find that if private agents (households and firms), and the monetary authority disagree about the expected inflation rate, then in an inflation targeting regime, the central bank should not respond aggressively to deviations from an inflation target that it itself anticipates. Weaker responses improve economic stability in the short run. They lead to falls in inflation volatility, in output volatility, and in the central bank's expected loss. This is in contrast to the findings for the imperfect, homogeneous expectations environment (e.g., Orphanides and Williams, 2003; and Ferrero, 2007).

Heterogeneous expectations cause a mismatch in subjective real interest rates. The mismatch leads to higher volatility in both inflation and output than would occur when expectations are homogeneous across the economy. One of the worst scenarios occurs when private agents predict less inflation than the central bank. That leads the central bank to raise the policy interest rate. For private agents, who expect lower inflation, the *ex ante* real interest rate is higher. Higher real rates cause private agents to substitute from current consumption so that aggregate demand drops. But consumption drops more than it would have if the private agents expected the same inflation rate as the central bank because it is the subjective real rate that matters. So the effect of monetary policy is stronger than the central bank itself intends. A similar situation, but with opposite implications, arises when the central bank expects low future inflation, and private agents expect high inflation. Implied, subjective real interest rates are low for private agents, which results in the economy growing at the cost of unnecessarily high inflation.

The role of monetary policy is complex in a heterogeneous expectations environment. A central bank's aversion to price inflation implies strong policy responses to deviations of expected inflation from the desired target. But if the central bank is too responsive, it multiplies the effect of the mismatch in the real rates even more. In the short run, the mismatch matters most for monetary policy. In the mid to long run, this phenomenon naturally disappears, and optimal monetary policy is standard as in a homogeneous expectations environment.

The following text describes a simple numerical analysis of a New Keynesian model to assess its dynamics under imperfect and heterogeneous knowledge on the part of economic agents. A particular focus is on the implications that heterogeneous expectations have for the optimal behaviour of the central bank. The next section sets up the experiment laboratory: a workhorse model, adaptive learning mechanism, and the source of expectations heterogeneity. In the third section, dynamics of the model environment are studied and basic observations are summarized. The fourth section provides the economic intuition for the results. The last part concludes with a general discussion about the results and what lesson can be taken for monetary policy.

2.2 The model

The New Keynesian business cycle model is used as an approximation of the economy. As an extension to the standard model, the assumption that monetary policy is perfectly credible is relaxed. As a result, private firms and households – as one economic group - form different expectations from the central bank. All agents use an adaptive (econometric) learning mechanism to learn about the actual structure of the economy, and they are allowed to disagree in their views. The only source of expectations heterogeneity in my set-up is that the private agents and the central bank give different weights to past forecasting errors. They differ in the opinion about how much of the innovation is due to a fundamental error in their forecasting model – in other words, about how much they should update the model structure – and how much of it is due to an unanticipated shock that hits the economy.

The basic model is standard; e.g., see Walsh (2003), Ch.5.4, or Honkapohja and Mitra (2005). The aggregate dynamics are given by the IS curve (2.1), which is the households' Euler equation, linearized around a flexible price equilibrium; and the Phillips curve (2.2), which is the linearization of the firms' pricing rule. In a perfect-knowledge environment, the model is

$$x_t = E_t x_{t+1} - \sigma (i_t - E_t \pi_{t+1}) + g_t, \tag{2.1}$$
$$\pi_t = \beta E_t \pi_{t+1} + \lambda x_t + u_t. \tag{2.2}$$

x_t is the output gap, defined as the deviation of actual output from the output arising in a flexible price environment; π_t is the inflation rate, and i_t is the interest rate set by the central bank. g_t and u_t are demand and cost-push shocks, respectively, assumed to follow AR(1) processes. β, σ, and λ, are the households' time preference parameter, risk aversion parameter, and inflation to output gap elasticity parameter, respectively.

The nominal side of the economy is anchored – at a zero inflation target rate – by a discretionary, expectations-based policy rule:

$$i_t = \theta_0 + \theta_\pi E_t \pi_{t+1} + \theta_x E_t x_{t+1}. \tag{2.3}$$

i_t is the nominal interest rate set by the central bank. θ_0 collects constant terms like the equilibrium real interest rate and inflation target. θ_π and θ_x are the policy weights on inflation and the output gap, respectively. Here, the central bank *cannot* observe the shocks $\{g_t, u_t\}$ when making policy decisions.[1]

All expectations operators $E_t(.) = E_t(.|\Omega_t)$ stand for perfect knowledge, rational expectations. Ω_t is the perfect-knowledge information set, $\Omega_t = \{\beta, \lambda, \sigma, \theta_0, \theta_\pi, \theta_x, g_t, u_t, g_{t-1}, u_{t-1}, ...\}$.

Definition 1. Economic agents have *perfect knowledge* if an information set Ω_t is available at time t, where

$$\Omega_t = \{\beta, \lambda, \sigma, \theta_0, \theta_\pi, \theta_x, u_t, g_t, u_{t-1}, g_{t-1}, ...\}.$$

The information set contains the true values of the structural parameters and the current and past exogenous shocks u and g.

Definition 2. Economic agents have *imperfect, homogeneous knowledge* if all agents share the same, imperfect information set $\hat{\Omega}_t$ at time t, where

$$\hat{\Omega}_t = \{\hat{\Theta}_t, \kappa_t, u_t, g_t, u_{t-1}, g_{t-1}, ...\}.$$

$\hat{\Theta}_t$ is the imperfect, time varying belief about the true structural parameters $\{\beta, \lambda, \sigma, \theta_0, \theta_\pi, \theta_x\}$, and κ_t represents the information gain (willingness to learn, or the sensitivity to new information).

[1] Evans and Honkapohja (2003b) derive an expectations-based rule where a central bank observes shocks in the current period.

Definition 3. There are two groups of agents: (P) private agents, and (CB) the central bank. The private agents and central bank have *imperfect, heterogeneous knowledge* if the information they have differs, and it is not perfect, $\Omega_t^P \neq \Omega_t^{CB}$.

The workhorse model The key assumption of this paper is that the perfect information set Ω_t is not available to agents. Agents have imperfect and heterogeneous knowledge, which leads to a heterogeneous expectations formation. The workhorse model takes the form

$$x_t = \hat{E}_t^P x_{t+1} - \sigma \left(i_t - \hat{E}_t^P \pi_{t+1} \right) + g_t, \tag{2.4}$$

$$\pi_t = \beta \hat{E}_t^P \pi_{t+1} + \lambda x_t + u_t, \tag{2.5}$$

$$i_t = \theta_0 + \theta_\pi \hat{E}_t^{CB} \pi_{t+1} + \theta_x \hat{E}_t^{CB} x_{t+1}, \tag{2.6}$$

where $\hat{E}_t^P(.) = E_t(.|\Omega_t^P)$ are the subjective, imperfect-knowledge expectations of private agents, and $\hat{E}_t^{CB}(.) = E_t(.|\Omega_t^{CB})$ are the subjective, imperfect knowledge expectations of the central bank. The individual imperfect information sets Ω_t^P and Ω_t^{CB} are subsets of the perfect knowledge set, $\Omega_t^P, \Omega_t^{CB} \subset \Omega_t$.

I am interested in biasing this economy to an absurd extreme. To deviate from the homogeneous expectations case (the pooling-information assumption), and at the same time to avoid the problems of infinite-order expectations due to "forecasting others' forecasts", as raised by Towsend (1983), I suppose that agents believe that everyone shares their own expectations. And they do not learn from experience other than their own. Agents in this set-up are very naive. Even though it might not seem to be a very realistic assumption, it is useful.[2] It sets bounds for the results that one might expect for the convex combinations of two extreme assumptions - this one, and the imperfect homogeneous knowledge assumption.

Adaptive learning mechanism Expectations heterogeneity turns on an adaptive learning technology. The learning mechanism described below, resembles the assumption about the agents' knowledge. Honkapohja and Mitra (2005) show that the move from the perfect knowledge model to the imperfect and heterogeneous knowledge model is possible under Euler-equation learning. If all agents are learning (using recursive least squares, and the E-stability conditions hold), the originally heterogeneous knowledge Ω_t^P and Ω_t^{CB} is enriched over time so that it converges to the perfect knowledge set Ω_t. It happens even despite the very restrictive assumption that agents believe that everyone shares their expectations, and they only trust their own experience.

The adaptive learning methodology relies on agents' learning about a reduced form model. Solving for rational expectations, the minimum-state representation to the structural model (2.4)-(2.6) is

$$Y_t = a + b s_t.$$

Y_t is the vector of endogenous variables, $(x_t, \pi_t)'$, s_t is the vector of exogenous shocks $(g_t, u_t)'$, and a and b are the matrices collecting structural parameters. Their derivation is in Appendix A.

[2] It is at least as much realistic/unrealistic as the perfect, homogeneous, knowledge assumption.

The (P) private agents' and (CB) central bank's *perceived law of motion* (PLM) for the economy (2.4)-(2.6) is assumed to be

$$\hat{Y}_t = \hat{a}_t^i + \hat{b}_t^i s_t,$$

where $\{\hat{a}_t^i, \hat{b}_t^i\} \in \Omega_t^i$, and $i = \{P, CB\}$ are the time-varying matrices of the model primitives, representing beliefs about the true structure $\{a, b\}$. Implicitly in this framework, agents have perfect knowledge about the structure of the economy, but they have imperfect knowledge about the true values of some of the structural parameters. Consequently, private agents and the central bank both learn about the structural matrices $\{a, b\}$ over time. The learning behaviour takes the form of econometric learning (recursive least squares). In the adaptive learning literature, it is believed that such a mechanism resembles the actual behaviour of agents very closely – for all see Evans and Honkapohja (2001). The algorithm is

$$\begin{aligned} \xi_t^i &= \xi_{t-1}^i + \kappa_t^i (R_t^i)^{-1} X_t (Y_t - X_t' \xi_{t-1}^i), && (2.7) \\ R_t^i &= R_{t-1}^i + \kappa_t^i (X_t X_t' - R_{t-1}^i). && (2.8) \end{aligned}$$

$i = \{P, CB\}$, $\xi_t^i = [\hat{a}_{11}^i, \hat{a}_{21}^i, \hat{b}_{11}^i, \hat{b}_{12}^i, \hat{b}_{21}^i, \hat{b}_{22}^i]'$ is the vector of the PLM parameters. X_t is the matrix of appropriately stacked exogenous shocks s_t, and κ_t^i is the information gain. I also call this gain as the willingness to learn, or the sensitivity to new information. The information gain is the only source of heterogeneity. R_t^i is the information matrix available at time t to a group i.

2.3 Model dynamics

This section analyzes the dynamics of the workhorse model. The goal is to assess the implications that expectations heterogeneity has in a forward-looking monetary policy regime for short-run economic fluctuations. I focus on two questions. The first question is, what is the contribution of expectations heterogeneity to inflation and output volatility? The benchmark is the standard, rational expectations model with optimized monetary policy. The second question is, how can a central bank's behaviour minimize the fluctuations in heterogeneous expectations environments?

To address both questions, the plan is to perform an intervention analysis. I expose the model economy to a one-period, unitary, cost-push shock, to a one-period, unitary, demand shock, and to a combination of the two. The REE serves as benchmark dynamics. There are no monetary policy shocks. To summarize the results, the central bank's expected loss is the prism. I report (i) the half life of the shock to central bank's expected loss (denoted as HL; it is the time it takes from the initial effect of the shock to the value of one half of the response amplitude), and (ii) the amplitude of the response deviation from the rational-expectations dynamics (denoted as max). It is the maximum deviation of imperfect knowledge dynamics from REE dynamics. If it is positive, the adaptive learning (AL) economy is more responsive to the shock than under the rational expectations (RE); if negative, less.

Model calibration The paper Clarida, Gali and Gertler (2000) calibrates the workhorse model. The calibrated values are: $\sigma = 1$, $\beta = 0.99$, and $\lambda = 0.3$. Optimal weights are derived for the policy rule (2.3).[3] Assuming that a central bank puts 1/3 weight on output stabilization, and 2/3 weight on inflation stabilization $(\theta_\pi^*, \theta_x^*) = (1.5, 1)$; see Appendix F. For comparison purposes, I also use two sets of non-optimal policy weights: $(\theta_\pi, \theta_x) = (1.3, 1), (2.5, 1)$. In all the simulations, I assume an econometric learning algorithm, which means that whenever a new piece of information (observation) arrives, the agents re-estimate their forecasting models. The recursive econometric learning is represented by (2.7) and (2.8) with $\kappa_t^i = c_i(t-15)^{-1}$, where t denotes time, and $i = \{CB, P\}$; c_i is a positive constant and stands for a bias in the information gain. If $c_i = 1$, κ_t^i represents the recursive least squares technique. If $c_i > 1$, there is a greater willingness to update than under standard econometric learning. $\kappa_t^i \to 0$ as $t \to \infty$, thus the effect of $c_i \neq 1$ matters only initially. Next, I calibrate the autocorrelation in demand and cost-push shocks to be 0.2. The reason for such a small number is that high persistence in the output gap and inflation is delivered by adaptive learning – see for instance Milani (2007). The value is set to replicate the empirical volatility of inflation and output. All the simulations are initialized from steady state values: $\xi_0^i = [a_{11}^i, a_{21}^i, b_{11}^i, b_{12}^i, b_{21}^i, b_{22}^i]'$, for $i = \{P, CB\}$. R_0^i is an identity matrix.[4]

Experiment description First, the reference results start with the case where knowledge is imperfect but homogeneous. Both the private agents and the central bank have the same sensitivity to new information/innovations, $\kappa_t^P = \kappa_t^{CB} = \kappa_t$. Then, the same technology is used to study the heterogenous expectations case. To answer what the contribution of expectations heterogeneity to inflation and output volatility is, the two groups of agents are assumed to have different sensitivity to new information, $\kappa_t^P \neq \kappa_t^{CB}$. The focus is on the instances in which (i) the private agents are more sensitive than the central bank, and (ii) private agents are less sensitive than the central bank. To answer how can a central bank's behaviour minimizes the fluctuations under setup (i) and (ii), I compare the effects of monetary policy which is (a) optimal in the rational expectations (RE) environment; (b) more; and (c) less responsive to inflation than under the optimal (RE) setting.

2.3.1 The central bank's expected loss in the homogenous case

Table 2.1 and 2.2 give a representative set of results.[5] The diagonals of the three panels show results for the homogeneous knowledge case. The results for heterogenous knowledge lie off the diagonals and are discussed in the next subsection. Table 2.3 summarizes the impulse responses for the perfect knowledge case (RE dynamics).

Adaptive learning contribution to volatility Looking at the on-diagonal results, we can clearly see that the adaptive learning increases overall economic volatility. In all

[3] Their derivation is in Appendix C.

[4] The Matlab codes to replicate the simulation results can be obtained from the author upon request.

[5] A full grid search was performed for all possible combinations of policy parameters ($\theta_\pi \in (1, 5)$ and $\theta_x = 1$ and the information gain parameters $\{c_p, c_{CB}\} \in (0.8, 1.2)$. Due to their complexity, I decided to present only a representative set. The full set of results can be obtained upon request.

Table 2.1: The amplitude of the central bank's loss function response to demand and cost-push shocks

Amplitude			θ_π								
			1.3			1.5			2.5		
c_P	Shock		c_{CB}								
	g_0	u_0	0.8	1	1.2	0.8	1	1.2	0.8	1	1.2
0.8	1	0	0.055	0.050	0.047	0.043	0.039	0.037	0.031	0.028	0.018
	0	1	0.192	0.188	0.184	0.162	0.159	0.157	0.136	0.133	0.130
	1	1	0.232	0.221	0.214	0.194	0.188	0.183	0.187	0.182	0.179
1	1	0	0.077	0.066	0.062	0.056	0.050	0.046	0.040	0.035	0.032
	0	1	0.242	0.235	0.230	0.200	0.196	0.193	0.166	0.162	0.159
	1	1	0.307	0.281	0.270	0.244	0.235	0.229	0.230	0.224	0.219
1.2	1	0	0.109	0.084	0.077	0.073	0.062	0.057	0.049	0.042	0.037
	0	1	0.295	0.284	0.277	0.239	0.234	0.229	0.197	0.192	0.187
	1	1	0.407	0.344	0.328	0.339	0.311	0.297	0.272	0.265	0.258

Note: For shocks, "0" means no shock, "1" is a unitary shock.

Table 2.2: The half life of the central bank's loss function response to demand and cost-push shocks

			θ_π								
			1.3			1.5			2.5		
c_P	Shock		c_{CB}								
	g_0	u_0	0.8	1	1.2	0.8	1	1.2	0.8	1	1.2
0.8	1	0	1000+	990	981	993	971	970	488	742	720
	0	1	1000+	994	993	1000+	1000+	1000+	222	225	233
	1	1	997	996	995	1000+	1000+	1000+	153	149	151
1	1	0	1000+	995	993	996	993	977	227	285	389
	0	1	1000+	995	994	859	816	812	140	137	139
	1	1	1000+	997	993	1000+	821	867	102	96	95
1.2	1	0	607	995	1000+	430	644	925	153	151	192
	0	1	1000+	1000+	1000+	504	462	454	102	97	97
	1	1	1000+	1000+	1000+	993	996	1000+	78	71	69

Note: For shocks, "0" means no shock, "1" is a unitary shock.

Table 2.3: The summary of impulse responses for the rational expectations

Shock		Amplitude	θ_π		
g_0	u_0		1.3	1.5	2.5
1	0	7	0.98	1.02	1.29
0	1	7	0.72	0.52	0.43
1	1	7	1.49	1.52	1.58

considered cases, $c_P = c_{CB} = \{0.8, 1, 1.2\}$, the amplitude (Table 2.1), and the half life (Table 2.2) is a positive number, which means that the impulse response of the expected CB's loss is bigger than under the RE dynamics for all t. For example, if a demand shock hits the economy, $u_0 = 1$, and $c_P = c_{CB} = 0.8$ and $\theta_\pi = 1.3$, the expected central bank's loss is higher by 0.192 basis points. The total loss amplitude is then 1.072; the sum of the RE response, 0.98; and the contribution of adaptive learning, 0.192. The HL of the shock exceeds the RE case by more than 1000 periods (1000+). If $\theta_\pi = 2.5$, then the contribution to the response amplitude is 0.136, and the total amplitude is 1.116. The HL is only 488 periods longer than in the RE case (Table 2.3), where it takes the economy only 7 periods to converge to its steady state.

We can also see that with higher sensitivity to new information, $c_i > 0.8$, the HL shortens. Examining the right panel of Table 2.2 for the demand shock again, we see that the half life drops from 222 periods, through 137, to only 97 periods, as $c_P = c_{CB}$ increases from 0.8 to 1.2. The same results hold for a cost-push shock, and for cost-push and demand shocks occurring jointly, $g_0 = u_0 = 1$.

Remarkably, the policy that is efficient under rational expectations does not perform very well in an imperfect knowledge case. It is a similar finding as in Orphanides and Williams (2003). From Tables 2.1 and 2.2, it follows that the optimal (RE) policy is outperformed by a more inflation responsive policy.[6]

Monetary policy effect The key result for the homogeneous case is that a monetary policy can effectively influence both economic variability and the speed of learning (the convergence to the REE). The numbers in Table 2.1 and 2.2 demonstrate that increasing the inflation responsiveness from $\theta_\pi = 1.3$ to 2.5, the deviation from RE dynamics lowers. The shock response amplitude drops in all three cases. Also the speed of learning significantly improves. Its relation to the policy reactiveness is highly non-linear. All these results confirm the findings made by Orphanides and Williams (2002), and Ferrero (2004). "Policy should respond more aggressively to inflation under imperfect knowledge than under perfect knowledge [...] in order to anchor inflation expectations and foster macroeconomic stability" (Orphanides and Williams, 2003, p.26).

The results for the imperfect, homogeneous knowledge case can be summarized in two points:

- overall volatility increases with higher sensitivity to new information, but the increase in volatility is offset by faster learning;
- if the policy reacts aggressively to inflation, the central bank's expected loss decreases; the speed of learning increases.

2.3.2 The central bank's expected loss in the heterogeneous case

The results under heterogenous expectations differ dramatically from the benchmark case. The summary of impulse response characteristics is again in Table 2.1 and 2.2. They lie

[6] "...policies that would be efficient under rational expectations can perform poorly when knowledge is imperfect" (Orphanides and Williams, 2003, p.26).

off the diagonal of the three panels. There are two dimensions to read the results in: the effect of different sensitivity to new information ($\{c_P, c_{CB}\} = \{0.8, 1, 1.2\}$) and the policy inflation reactiveness ($\theta_\pi = \{1.3, 2.5\}$). We can read the following story from Table 2.1 and 2.2:

The effect of expectations heterogeneity

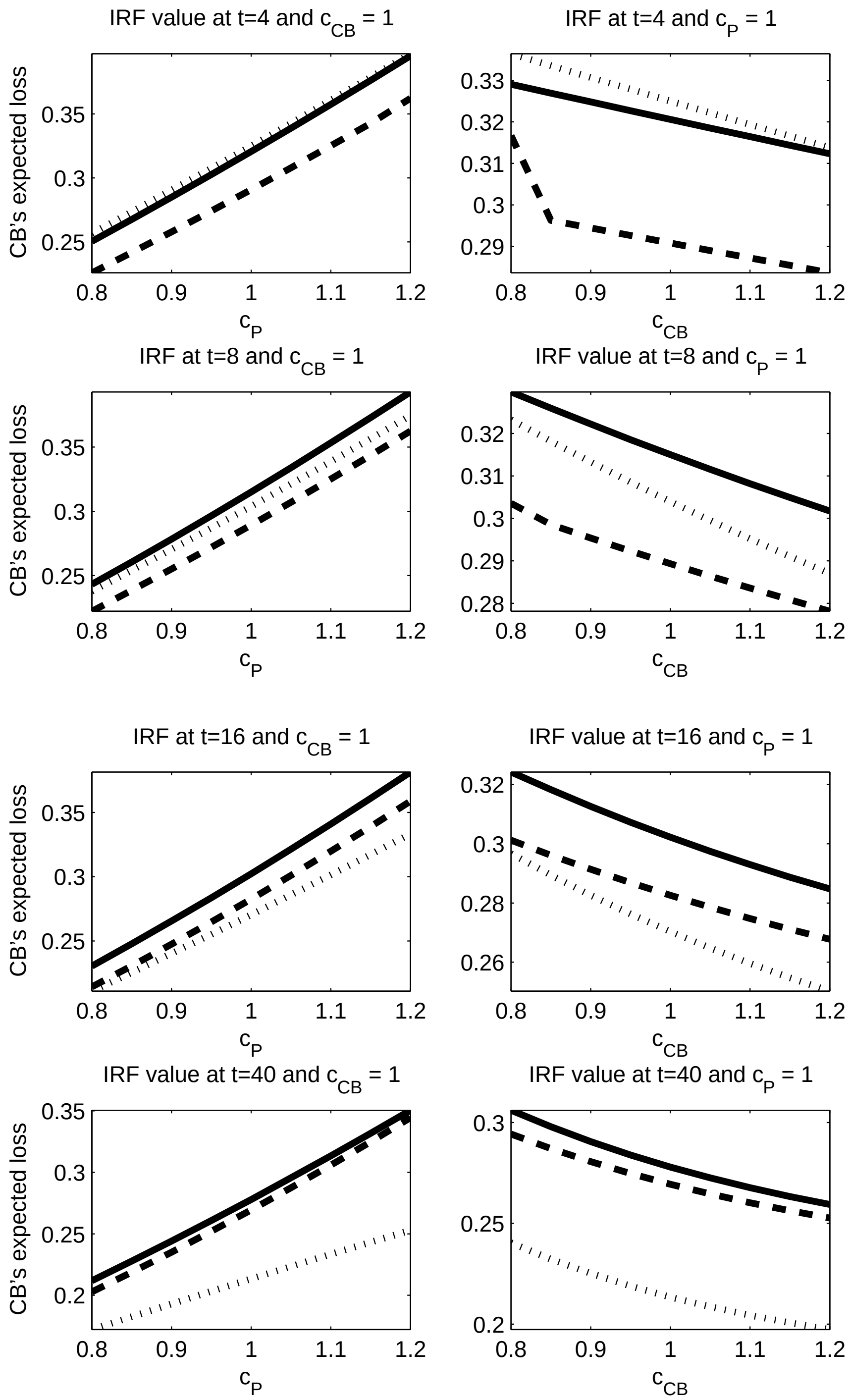
- As the private sector becomes more sensitive to new information, economic variability increases. Fixing c_{CB}, the impulse amplitude of the CB's expected loss increases with c_P.
- As the central bank becomes more sensitive to new information, economic variability decreases. Fixing c_p and shifting c_{CB}, amplitude falls.
- As both the central bank and private sector become more new-information sensitive (c_i increases), and at the same time monetary policy is less inflation responsive ($\theta_\pi = 1.3$), volatility falls.
- As the private sector becomes more sensitive to new information, the speed of convergence increases.
- As the central bank becomes more sensitive to new information, and at the same time the policy is too responsive, the speed of convergence decreases. If the policy is not sufficiently responsive, the results are inconclusive.

The effect of monetary policy

- As policy becomes more inflation responsive, economic volatility decreases.
- A relative reduction in volatility depends on the degree of expectations heterogeneity (c_P-c_{CB} ratio).
- As policy becomes more inflation responsive, speed of convergence significantly increases.

The aggregated results indicates, that in the short run, monetary policy has a limited operational space, and non-trivial implications. In Figure 2.1, I plot the impulse response function of the central bank's expected loss for a different policy reactiveness to inflation. The impulse responses are presented in a static form. They are cut for horizons $t = \{4, 8, 16, 40\}$, and plotted for a different information gain bias, $\{c_P, c_{CB}\}$. Periods 4 to 8 are here to represent the short run. The medium run, and long run is represented by 16, and 40 periods, respectively.

Figure 2.1: Impulse responses of the central banks' expected loss



Note: The *dash line* is the response when $\theta_\pi = 1.3$; a *continuous line* when $\theta_\pi(\alpha = 1/3)^* = 1.5$; and a *dotted line* when $\theta_\pi = 2.5$. The economy is subject to a unitary cost-push and demand shock ($u_0 = g_0 = 1$), and the impulse response function is evaluated in 3 dimensions: time, and information gain bias c_p, and c_{CB}. The graphs above show the impulse response value at time $t = \{4, 8, 16, 40\}$.

Figure 2.1 reveals that in the short run, monetary policy, which is less responsive to inflation ($\theta_\pi < 1.5$), delivers the lowest expected loss for almost all combinations of $\{c_P, c_{CB}\}$ – the dashed line lies at the bottom at $t = 8$. On the other hand, this is not a long-lasting phenomenon. There is a rising stabilizing effect of an inflation responsive policy ($\theta_\pi > 1.5$) in the mid-run. At $t = 16$, it depends on the degree of information gain bias, but the dotted line shows that the responsive policy performs the best in most cases. In the long run, for $t \geq 40$, the inflation responsive policy is a dominant one – the dotted line is always connected with the lowest expected losses.

Similarly as found in the homogeneous expectations case, optimal (RE) monetary policy does not perform very well under heterogeneous expectations. In Figure 2.1, the continuous line, where $\theta_\pi^* = 1.5$, is not connected with the lowest expected loss. Even though it can be the second best option in some cases, it is always dominated by a more or less inflation responsive policy.

Robustness The robustness of results is checked by changing the relative sizes of the shocks. The basic results remain mostly unchanged. As the variance of the demand shock g_t gets bigger in relative terms, there is a polarization of the policy effect at the short and long horizon. At $t = 20$, an inflation-reactive policy clearly dominates. On the other hand, as the variance of cost-push shock u_t gets bigger, in relative terms – approximately twice as big – the picture slightly changes. At the short horizon, there is already a region of expectations heterogeneity in which an inflation responsive policy is preferred, and over the time it dominates.

2.4 Some intuition behind the results

One of the characteristics of adaptive learning is that, eventually, the boundedly rational equilibrium path converges to the REE, Evans and Honkapohja (2003a). Even though the two groups of agents are assumed not to communicate with one another, they can reach homogeneous and perfect knowledge. Over time, both groups, unintentionally (purely based on their own errors), end up with the same forecasting model and the same expectations. This is why we observe in Figure 2.1 that an inflation-responsive policy starts to dominate after 16 quarters. It is because expectations become homogeneous, and the economic environment evolves towards the RE. An important difference is that this does not hold early on, and a too responsive policy can actually considerably destabilize the economy. This observation leads to the conclusion that when expectations are heterogeneous, monetary policy should not be too active in order to improve stability in the short-term.

To understand the results, it might help us, if the model (2.4)-(2.6) is rewritten in a more suitable form so that we can see the effects of expectations heterogeneity more easily. For simplicity, I also assume that $\hat{E}_t^P x_{t+1} - \hat{E}_t^{CB} x_{t+1} = 0$. Then the workhorse model can be written as

$$\begin{aligned} x_t &= -\sigma\theta_0 + \sigma\theta_\pi(\theta_\pi^{-1}\hat{E}_t^P \pi_{t+1} - \hat{E}_t^{CB}\pi_{t+1}) + g_t, \\ \pi_t &= \lambda\sigma\theta_0 + \lambda\sigma\theta_\pi \left[\frac{\lambda\sigma + \beta}{\lambda\beta\theta_\pi}\hat{E}_t^P \pi_{t+1} - \hat{E}_t^{CB}\pi_{t+1}\right] + u_t + \lambda g_t. \end{aligned}$$

Demand shock A demand shock first hits the output gap, and then affects inflation for a while. Beginning in REE, agents were expecting equilibrium values of inflation and the output gap. In the RE and persistence-less environment, the shock would have just a one period impact. Under adaptive learning it influences expectations for subsequent periods. Being surprised, agents update their forecasting models. The policy rate is set so that it neutralizes the shock. A positive demand shock will cause an upward correction in the PLM's parameters, which will yield higher predictions of inflation and the output gap for future periods. The policy rate reacts to those values. Because increasing c_{CB} causes higher expected values for inflation and output gap, monetary policy is suddenly more restrictive - policy rate increases. This is why the expected CB's loss declines as c_{CB} increases.

Using the same logic, we can interpret the effect of increasing private sector sensitivity to new information. A demand shock transmits to the future via expectations. Private agents update their model similarly to the central bank. Their expectations, however, influence the economic dynamics directly. A positive shock motivates model updates, yielding higher inflation and output gap forecasts in the future. Higher output gap expectations imply a higher current output gap, and consequently higher inflation. Higher inflation expectations have a direct effect on inflation, which increases, and an indirect effect on the output gap via a decrease in the real interest rate, which influences the output gap positively.

Cost-push shock Assuming no persistence, a cost-push shock has an immediate impact on contemporaneous inflation and expectations, through which it transmits further. In the next period, since no other shock occurs, inflation should return to the REE. But because the private agents and the central bank update their model – biasing upward their expectations, the inflation rate and the output gap increase above the RE values. The mechanism of monetary policy is the same as in the previous shock case. An inflation averse policy pushes inflation down to the RE dynamics, and since the policy is now more aggressive than under the RE, the output gap decreases more, and becomes more responsive.

The central bank's new-information sensitivity decreases the inflation rate responsiveness to a cost-push shock, but increases the responsiveness of the output gap. Again, monetary policy becomes more restrictive than under the RE since the central bank predicts higher inflation due to the model updates, it tightens the interest rate, which closes the output gap, and the inflation rate returns to the RE dynamics. Thus by changing c_{CB}, we can explain the decrease in the responsiveness of inflation, accompanied by the increase in the responsiveness of the output gap.

The private sector's sensitivity to new information helps the cost shock to propagate to inflation. As private agents become more innovation sensitive, they anticipate higher inflation than under full knowledge, and thus increase the actual inflation rate. With increasing c_P, agents update their models more, and produce higher forecasts of inflation. This immediately increases inflation due to higher expected inflation in the future. Agents also update their forecasts of the output gap. They will anticipate the reaction of the central bank, which they assume has the same expectations as themselves, which will lead to a policy rate adjustment. Since c_P will bias a policy reaction upwards, private agents will assume a lower output gap than under the RE. This explains why the output gap becomes more reactive if the private sector is more information sensitive. This phenomenon is observable particularly if the central bank is too responsive to inflation.

2.5 Concluding discussion

The world is simpler if knowledge and beliefs are homogeneous. If knowledge is homogeneous, a central bank's aversion to price inflation helps to decrease inflation variability and speeds up learning. The speed of learning affects the persistence of inflation and its variability. If a central bank wishes to minimize its expected loss, it is desirable that agents learn the economy's actual law of motion as fast as possible. If knowledge and beliefs are heterogeneous, the central bank should not be too anti-inflationary, because if the bank is less inflation responsive, short-run economic stability improves. Thus, in a heterogeneous expectations world, the first goal of the central bank should aim to make expectations homogeneous across the economy, in order to minimize inflation target and output volatility. Once expectations have become homogeneous, the standard policy recommendations apply.

How can a central bank make expectations homogeneous in the short run? The expectations homogeneity is closely related to enhancing policy effectiveness. In this simple model, there are two ways this may work. First, either the central bank learns and adopts private agents' expectations. Or second, private agents get to know and acquire the central bank's expectations. (And, of course, the two processes could be combined, with both sets of expectations converging on each other). In practice, neither is simple. The first will require reliable measures of private sector expectations. Central banks usually have surveys of private sector expectations on future economic developments. But the information that such surveys yield might be unreliable. The data collected may not truly represent market expectations, which drive agents' market behaviour - they could be subject to systematic measurement errors (due to inaccurate or collusive, game-playing responses, perhaps). In fact, the central bank can never be sure if the data being collected are useful for immediate policy decisions. Those considerations suggest it might be better for private agents to borrow central bank expectations, than the reverse. How can this be done though? And can it be relied upon? Central bank communications, through publications, speeches, and press conferences, clearly provide a crucial educational function. But when credibility is absent, the logic of this paper is that the central bank really must furnish evidence of its commitment and capability to turn its expectations into reality.

A major challenge for future research is an analytical solution to the problem addressed in this paper. Even having a simple model, heterogeneous expectations and adaptive learning lead to an analytically uneasily tractable problem. Model transition functions are highly non-linear, which complicates and limits a comparative statics analysis. An analytical evaluation of the speed of learning, as in Ferrero (2007), also seems complex. At present, a numerical analysis seems to be the most viable approach.

Chapter 3

Inflation Expectations in the Czech Interbank Market

3.1 Introduction

In this chapter, I am interested in testing the similarity of the surveyed inflation expectations and the inflation expectations of Czech interbank market participants. The expectations of market participants – the *market expectations*, are unobservable and one can make inferences about them only indirectly, through market prices or other market indicators. I develop a simple econometric test here that evaluates the statistical difference between the surveyed, and the (latent) market inflation expectations using information included in the 1-year Prague Interbank Offer Rate (1Y Pribor).

The test treats inflation expectations as independent variables, and the Fisher rule is assumed to be the true asset pricing rule. I use the rule to model the 1Y Pribor as a function of inflation expectations, and the real interest rate. The market expectations are allowed to violate the rational expectations hypothesis (REH), and the relationship itself is further allowed not to hold perfectly in the short run. These assumptions make the Fisher rule, which is otherwise a theoretical concept, more flexible.

Conditional on the assumed asset pricing rule, I find that the difference between the surveyed and market expectations is not statistically significant. The surveyed expectations seem to be a good proxy of the market expectations.

The text is structured as follows. The methodology is formulated in the next section 2. The data set and estimation results with their interpretation follow in sections 3 and 4, respectively. In section 5, I discuss the results, and section 6 concludes.

3.2 The Methodology

In this section, I develop a test to assess the null hypothesis: "*the surveyed expectations are the same as market expectations.*" It is important to emphasize that in the methodology developed below, the expectations' rationality, as assumed by the REH, is not questioned. Instead, the assumption is weakened and the expectations may be biased and inefficient.

This assumption is a difference from the literature like Fama and Gibbons (1982), Mishkin (1990b), and Mishkin (1990a), which has to deal with unobservable market expectations.

The Fisher rule Testing H_0, the Fisher rule – an equilibrium relationship between inflation expectations, *ex ante* real interest rate, and the nominal interest rate – is assumed as the true pricing rule for the money market, i.e.,

$$i_t^m = \Pi_{t,m}^e + r_t^m + \rho_t, \tag{3.1}$$

where i_t^m is the nominal interest rate valid from the beginning of period t to $t+m$, $\Pi_{t,m}^e$ are the true market expectations formed at the beginning of period t for the $t+m$ time horizon, and r_t^m is the corresponding *ex ante* real interest rate. To capture short-run deviations from the long-run relationship, I augment the Fisher rule with the risk premium term, ρ_t, which is assumed to follow a stationary process, orthogonal to expected inflation, and the real interest rate.

In a standard way, let us assume that the inflation expectations are based on information Ω available at time t. Furthermore, following assumptions common to the economic literature, let us assume that the *market inflation expectations* at time t for the $t+m$ horizon are equal to the actual inflation rate π_{t+m}, but are subject to disturbances ε_{t+m}. Formally written

$$\Pi_{t,m}^e = E(\pi_{t+m}|\Omega_t) = \pi_{t+m} + \varepsilon_{t+m}. \tag{3.2}$$

Under the REH, it holds that ε_t has zero mean conditional upon the information set Ω_t - they are martingale differences with respect to past information. Note that the non-zero auto-covariance structure of errors causes only inefficiency of expectations but does not affect their consistency.For the testing procedure developed below, no particular form of the market expectations is prescribed, i.e. ε_{t+m} can be iid or have an AR of GARCH structure.[1] As the objective is to find the relation of the surveyed expectations to the market ones, it is convenient to keep ε_{t+m} in a nearly unspecified form.

Now substituting (3.2) in (3.1) yields

$$i_t^m = \pi_{t+m} + \varepsilon_{t+m} + r_t^m + \rho_t. \tag{3.3}$$

Depending on the correlation of expected inflation, $\pi_{t+m} + \varepsilon_{t+m}$ and *ex ante* real interest rate, r_t^m, as argued in Mishkin (1990b), the relation between the nominal interest rate and expected inflation does not need to be one-to-one. Hence, we augment (3.3) into a more general form

$$i_t^m = \phi_1(\pi_{t+m} + \varepsilon_{t+m}) + r_t^m + \rho_t,$$

The Fisher rule where $\phi_1 = 1$ counts only for a special case. Now adding and subtracting the mean values of the real interest rate, unanticipated inflation and risk premium $(\phi_1 E(\varepsilon_{t+m}), E(r_t), E(\rho_t))$, and the above expression becomes

$$i_t^m = \phi_0 + \phi_1 \pi_{t+m} + v_t. \tag{3.4}$$

where $\phi_0 = E(\phi_1 \varepsilon_{t+m}) + E(r_t) + E(\rho_t)$, and $v_t = \phi_1 \varepsilon_{t+m} + r_t + \rho_t - \phi_0 = \phi_1(\Pi_{t,m}^e - \pi_{t+m}) +$

[1] Note that a GARCH structure is perfectly compatible with the REH as ARCH errors are martingale differences with respect to the past information set.

$r_t + \rho_t - \phi_0$, collecting all latent/unobservable variables. The constant term ϕ_0 will be in general different from zero. No specific assumptions are imposed on ε_{t+m}, and thus, in general, $E(\varepsilon_{t+m})$ may be non-zero. Similar can be said about the mean values of the *ex ante* real interest rate, $E(r_t)$ and the risk premium, $E(\rho_t)$, which both can be expected to be of positive magnitudes.

Dealing with market expectations unobservability For v_t in formulation (3.4), one is not able to separate the real interest rate, risk premium, and the expectation error. It is impossible to distinguish what part of the variation of v_t accounts for the unanticipated inflation and what part accounts for the *ex ante* real interest rate and risk premium.[2] The testing procedure here avoids this limitation. The empirical test is built solely on the 'comparison' of expectations processes, where only one of them has to be directly observable. For the null-hypothesis, I assume a correspondence between the *surveyed expectations* (observable process) and the *market expectations* (unobservable process).

The error term v_t in expression (3.4) can be written for the *market expectations* and *survey expectations* separately:

$$v_t = \phi_1(\Pi^e_{t,m} - \pi_{t+m}) + r_t + \rho_t - \phi_0, \tag{3.5}$$

and

$$vv_t = \psi_1(\pi^e_{t,m} - \pi_{t+m}) + rr_t + \rho\rho_t - \psi_0, \tag{3.6}$$

where ψ_0 and ψ_1 have the same interpretation as ϕ_0 and ϕ_1, respectively, but as the surveyed expectations $\pi^e_{t,m}$ are introduced a different denotation is chosen. From here the idea of the test is straightforward. The null-hypothesis is

$$H_0 : \qquad \pi^e_{t,12} = \Pi^e_{t,12},$$

where again $\pi^e_{t,12}$ denotes the *surveyed inflation expectations* formed at the beginning of period t for the time horizon till $t+12$, and $\Pi^e_{t,12}$ are the *true market inflation expectations* for the same period of time. We can see that under H_0, the error terms (3.5) and (3.6) must coincide, i.e., $\phi_0 = \psi_0$ and $\phi_1 = \psi_1$.

Testing formula Substituting vv_t instead of v_t in (1.4) and rearranging it, one obtains

$$i_t = (\phi_0 - \psi_0) + (\phi_1 - \psi_1)\pi_{t+12} + \psi_1\pi^e_{t,12} + rr_t + \rho\rho_t. \tag{3.7}$$

For testing purposes, we set $\phi_0 - \psi_0 = a_r$,$\phi_1 - \psi_1 = a$, $\psi_1 = b$ to obtain the final form of the Fisher rule to be tested

$$i_t = a_r + a\pi_{t+12} + b\pi^e_{t,12} + u_t. \tag{3.8}$$

Because the *ex ante* real interest rate and risk premium are unobservable, they are put as the error term u_t in (3.8). If the above assumptions and following requirements hold,

[2] On this fact, Campbell, Lo and MacKinlay (1997), in section 1.5.2., build their argument that the rational expectations hypothesis is not testable. It can be only assumed.

then one may conclude that $\pi^e_{t,12} = \Pi^e_{t,12}$. (i) if $\psi_0 = \phi_0$ and $\psi_1 = \phi_1$, then it must be that $a_r = a = 0$ and (ii) b must be statistically significant and positive. From (i), we see that testing H_0 on (3.8), we perform a joint hypothesis test. First, we test for the equality of surveyed and market expectations. Secondly, we test for the equality of *ex ante* real interest rates and risk premiums.

Three major problems to pay attention to when performing the test The first concerns the *character of disturbances* in (3.8). The disturbances in (3.8) contain the variable part of *ex ante* real interest rate. If the real interest rate is a function of productivity, it may be the case that the disturbance term follows a non-stationary process. In general, when residuals contain a unit root, one ends up with a spurious regression. Given this possibility, an essential part of the methodology is to test the disturbances for stationarity. The augmented Dickey-Fuller (ADF) test and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) test are employed. Indeed, if a unit root is identified, the test cannot be used.[3] If the disturbances in (3.8) do not follow a unit-root process they are likely to have an autocorrelation structure. In this case, the test can still be used, one just has to correct for the autocorrelation effect on the parameters' standard error estimates. The parameters' unbiasedness and consistency property remain unaffected otherwise.

In equation (3.3a), it is stated that the nominal interest rate and expected inflation do not need to move in a one-to-one relation. Following an analysis by Mishkin (1990b), this may arise if there is a correlation between expected inflation and the *ex ante* real interest rate. Since, the real interest rate is in the error term, OLS estimates are inconsistent and the 'instrumental variable approach' must be employed instead. This is the method used here.

The second problem one should be aware of is the *problem of multi-collinearity.* The occurrence of multi-collinearity is very likely because of the nature of the formation of expectations. The closer the expectations are to the REH, the more severe the problem is because both π_{t+12} and $\pi^e_{t,12}$ are regressors.

The third problem is the *problem of joint hypothesis.* The methodology crucially relies on the assumption that the Fisher rule holds. Thus if it happens the null-hypothesis is rejected, one cannot be certain whether it is because the market expectations are truly different from the surveyed ones or because the Fisher rule is an invalid assumption. Indeed, this is common to all tests of this type. However, the problem is lessened here. As mentioned in the introductory section, the focus on the problem of inflation expectations is from a macroeconomic perspective, and there the Fisher rule is a standard way to capture the interest rate behavior. In addition, the results in the next section strongly suggest that the Fisher rule holds.

The test in a nutshell To test H_0, we estimate (3.8) and test whether the parameters a_r, a, and b are statistically significant. If the parameters a_r and a are insignificant

[3]The Bohm-Bawereck hypothesis may offer a solution. As the latent variable for the *ex ante* real interest rate, the real GDP growth can be used. Following the same idea as with expected inflation, the real GDP growth becomes one of the explanatory variables in (1.8), and the unanticipated part of the GDP growth enters the error term. In a standard economic environment, the unanticipated real growth ought to be stationary and so is the whole error term in (1.8).

and parameter b is significant, then H_0 cannot be rejected. If H_0 cannot be rejected, the surveyed inflation expectations are likely to coincide with the market expectations. Accounting for the test weaknesses, an essential part of the test has to be (i) the test of residuals stationarity and the consequent adjustment of critical values, and (ii) a check of the multi-collinearity magnitude.

3.3 The Data

The data set consists of the monthly data of the Czech Interbank money market. The source of the data is the Czech National Bank and the Czech Statistical Office. For testing purposes, three series of monthly data are employed: (i) the monthly average of nominal one-year Interbank interest rate (Prague InterBank Offer Rate, PRIBOR 1Y), (ii) year-to-year CPI inflation, and (iii) year-to-year expected change in the CPI. The expectations have been collected at the beginning of each month.[4] Figure 3.1 compares surveyed inflation expectations and the 1Y Pribor. Figure 3.2 plots the expectations against actual inflation.

Before moving further, let us discuss the data timing. A period t denotes a month. PRIBOR 1Y is an average of daily rates within the period t. CPI is a measure of the price level within a given period t. It is a within measure because the data are collected in the second decade of each month. We can roughly think of it as of a monthly price level average. When computing the actual inflation rate, one must be careful about the data timing. For our purposes, the year-to-year inflation rate at time $t+12$ is the relative change in the CPI between period $t-1$ and $t+11$. Because the information set available at time t to the agents contains only the CPI of $t-1$ as the latest information about the actual price level, i.e., $\Omega_t = \{CPI_{t-1}, CPI_{t-2}, ...\}$, then when comparing $\pi^e_{t,12}$ with actual inflation over this period, π_{t+12} as denoted in equation (8), the actual inflation rate should be computed as $\pi_{t+12} = 1 - \frac{CPI_{t+11}}{CPI_{t-1}}$.

[4] The details of the survey can be found at *http://www.cnb.cz*. The data range is from 1999:05 to 2005:05. The whole data set can be obtained upon request from the author.

Figure 3.1: Comparison of Inflation Expectations and Pribor 1Y

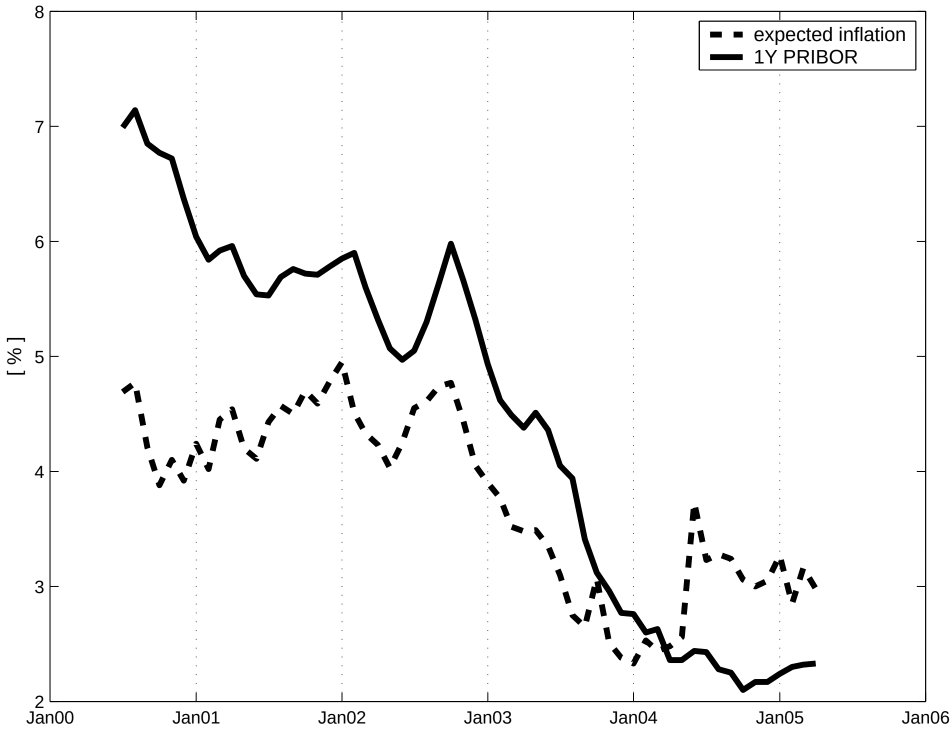
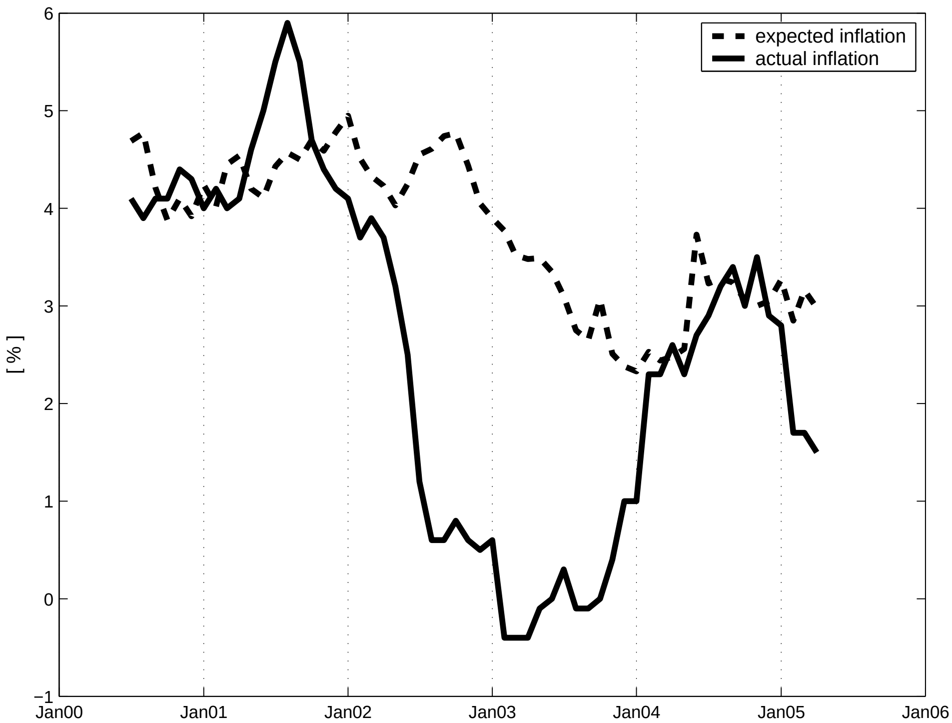


Figure 3.2: Comparison of Inflation Expectations and Actual Inflation



3.4 The Results

The estimation results of equation (3.8) are reported in Table 3.1 and Table 3.2. The test is performed for two data samples. The first sample ranges from May 1999 to March 2003.[5] The second one ranges from May 1999 to May 2004. Two different samples are considered because of the structural break which occurred in March 2003. The jump correction in expectations, accounting for a 1.17% increase in expected inflation, is explained by the expected growth of oil prices due to the beginning of the second war in Iraq, and an expected tax reform. Having the two samples, I control for the structural change effect on the results.

For the sample excluding the structural break (Table 3.1), the null-hypothesis that the market expectations are equal to the surveyed expectations cannot be rejected at the standard level of significance. The parameters a_r and a are not significantly different from zero, while parameter b is found to be significantly different from zero. The goodness of fit is about 80% (without accounting for autocorrelation in residuals) which suggests the joint hypothesis problem is not binding.

Similar results are obtained for the complete data sample. Here, equation (3.8) is expanded for a dummy variable to control for the structural break. In Table 3.2, we see again parameters a_r and a are not significantly different from zero, while parameter b is significant and positive. The dummy variable parameter $a_{r,dummy}$ is found significant. Its value of -1.7 can be accounted for the change in the mean value of unanticipated inflation $E(\varepsilon_{t+m})$, which constitutes the constant term in (1.8) and which has shifted in March 2003 by about 1.2%. The negative value of $a_{r,dummy}$ indicates that the impact of the tax reform and other exogenous shocks were not anticipated in the market expectations and thus not reflected in the market price. Even though the requirements to accept the null-hypothesis are not met here, because the constant terms of unanticipated inflation differ, we can conclude at least the variable parts of (surveyed and market) expectations are very similar. Parameters a and b have similar values and character for both data samples considered. Hence, we might conclude that the surveyed expectations are very close to the market expectations also on the full length sample. The variable components of expectations are still close to each other, and the variable components are the ones most interesting for any analysis.

Table 3.1: 2SLS results: $i_t = a_r + a\pi_{t+12} + b\pi^e_{t,12} + \epsilon_t$, 1999:05-2003:03

Parameter	Estimate	t-stat.	Critical value 5%	DW	$\bar{R}^2$	No. of obs.
a_r	0.3434	0.4868	4.5052			
a	0.1532	2.5789	4.3978	0.31	0.79	43
b	1.1016	5.4525	4.6014			

Note: instrumental variables $\{\pi_{t+11}, \pi^e_{t-1,12}, \Delta GDP_t\}$.

[5] The dating is from the expectations surveys perspective. For instance, May 1999 stands for the month when expectations for the period May 1999-May 2000 were formed.

Table 3.2: 2SLS results: $i_t = a_r + a\pi_{t+12} + b\pi^e_{t,12} + \epsilon_t$, 1999:05-2004:05

Parameter	Estimate	t-stat.	Critical value 5%	DW	$\bar{R}^2$	No. of obs.
a_r	-0.5230	-1.1005	4.5052			
$a_{r,dummy}$	-1.6769	-6.8732	4.5052			
a	0.0853	1.5881	4.3978	0.62	0.84	58
b	1.3599	10.1255	4.6014			

Note: instrumental variables $\{\pi_{t+11}, \pi^e_{t-1,12}, \Delta GDP_t\}$.

Properties of residuals and critical values of the test A very low value of Durbin-Watson statistics is found which indicates either the presence of a unit root in the residuals or their strong positive autocorrelation. To test for the former, the augmented Dickey-Fuller and Kwiatkowski-Phillips-Schmidt-Shin tests are used. Since the results, summarized in Appendix D at the end of the dissertation, do not suggest statistically significant evidence for the presence of the unit root, we argue in favor of a strong positive autocorrelation in the residuals. Because of the autocorrelation, the parameters' standard errors are biased downwards, and the pivotal statistics of t-tests are biased upwards. By simulating new critical values, the effect of autocorrelation on the test results is eliminated. The critical values reported in the tables are not the standard ones but those adjusted for the autocorrelation effect. Details on the simulation are presented in Appendix E. Another reason for simulating the t-test's critical values, instead of using the Student's distribution, is also the presence of multi-collinearity. Appendix F investigates this issue more closely. Because the variance inflation factor (VIF), the measure of the magnitude of collinearity, is 1.52, and a critical value is 10, it is concluded that multi-collinearity is not a significant problem. Despite this however, the parameters' standard errors are affected and simulating critical values partially account for it.

3.5 Discussion

Why the expectations of financial analysts should affect market dealers The surveyed expectations are formed by financial analysts who do not have any direct connection to the market makers in order to influence their actions. In practice, it is more than likely that the dealers do not pay any attention to the analysts' inflation expectations, or rather the dealers do not have any inflation expectations at all.[6] Their particular objective, as professionals, is to maximize their profit and their actions mostly have speculative motives. Despite this, the financial analysts' expectations seem to have a predictive power for money market price.

Possible explanation Money-market dealers do not necessarily need to know the inflation expectations the market price includes or even how the credit price is formed. They are only "endowed" with an excess or deficit of money which they trade. The

[6] I have interviewed a few market dealers.

deficits or excesses of loanable funds are determined by the credit market, which may be taken as exogenous to the Interbank market.

The credit market is the main channel used to transmit capital to the economy. Banks (lenders) are the market price setters, and borrowers are price-takers. The price on the credit market is derived from the price on the Interbank money market which, in contrast to the credit market, may be considered a competitive one. For banks as the major lenders of capital, it is profitable to bias their inflation forecasts (expectations) upwards. If we believe that bankers take into account a nominal depreciation of money when forming the credit price, overshooting these expectations increases their *ex post* real revenues. Given that the borrowers are price-takers, overshooting is accepted. Let us assume that a bank on the Czech credit market sets the one-year nominal interest rate on credit so that it is composed of an individual PRIBOR 1Y estimate plus a risk premium and a profit margin. The individual bank's PRIBOR 1Y estimate is composed of a required minimum *ex ante* real return plus an expected nominal depreciation (expected inflation). Having a price on the credit market, the lenders face a possible lack or excess of loanable funds (deposits from clients). To utilize them, they are motivated to enter the Interbank market and trade them. Under the assumption that the pricing rule is the same for all banks in the credit market, there will be only a moderate correction in the market PRIBOR in order to make the market clear. The new market PRIBOR is recursively reflected in the price on the credit market.

3.6 Conclusion

I find that surveyed inflation expectations have a significant predictive power for the 1Y Pribor monthly movements, modeled by the Fisher rule. Based on the econometric test, I developed, surveyed expectations are not statistically different from the market expectations.

The research has produced more questions than answers. The findings call for a study on why surveyed expectations proxy the Czech Interbank money market expectations. Although I make some headway towards explaining how the inflation forecasts are transmitted into the market price, a serious attempt to find the final answer should still be made.

In principal, a similar logic as in this paper, can be used to test the similarity of the surveyed expectations and the expectations adopted by the central bank. One of the central questions in the future, is to investigate the CNB's monetary policy pass-through to the interbank interest rates, and than further down the chain to the rates effective for firms' and consumers' decision making.

Bibliography

Adolfson, M., S. Lassen, J. Linde, and M. Villani, "Forecasting Performance of an Open Economy Dynamic Stochastic General Equilibrium Model," *Sveriges Riksbank Working Paper*, 2005, *190.*

Altug, S., "Time to Build and Aggregate Fluctuations: Some New Evidence," *International Economic Review*, 1989, *30*, 889–920.

Alvarez-Lois, P., R. Harrison, L. Piscatelli, and A. Scott, "Taking DSGE Models to the Policy Environment," in "IRFMP/IMP Conference on DSGE Modelling at Policymaking Institutions: Progress and Prospects" Washington, DC December 2005.

Amano, R., K. McPhail, H. Pioro, and A. Rennison, "Evaluating the Quarterly Projection Model: A Preliminary Investigation," *Bank of Canada Working Paper*, 2002, *2002-20.*

An, S. and F. Schorfheide, "Bayesian Analysis of DSGE Models," *mimeo*, 2005, p. University of Pennsylvania.

Binder, M. and M.H. Pesaran, *Handbook of Applied Econometrics: Macroeconomics*, Basil Blackwell, 1995.

Binette, A., S. Murchison, P. Perrier, and A. Rennison, "An Introduction to TOTEM," *Bank of Canada, mimeo*, 2005.

Black, R., D. Laxton, D. Rose, and R. Tetlow, "The Bank of Canada's New Quarterly Projection Model Part 1: The Steady-State Model: SSQPM," *Bank of Canada Technical Report*, 1994, *72.*

Breunig, R., S. Najarian, and A.R. Pagan, "Specification Testing of Markov Switching Models," *Oxford Bulletin of Economics and Statistics*, 2003, *65*, 703–725.

Brubakk, L., T.A. Husebo, S. McCaw, D. Muir, B. Naug, K. Olsen, O. Roisland, T. Sveen, and B.R. Wilhelmsen, "An Overview of the Norwegian Economy Model (NEMO)," *mimeo*, 2005, p. Bank of Norway.

Brundy, J.M. and D.W. Jorgenson, "Efficient Estimation of Simultaneous Equations by Instrumental Variables," *Review of Economics and Statistics*, 1971, *53*, 207–234.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay, *The Econometrics of Financial Markets*, Princeton University Press, 1997.

Canova, F., "Statistical Inference in Calibrated Models," *Journal of Applied Econometrics*, 1994, *9*, S123–S144.

___ **and L. Sala**, "Learning about the Parameters and the Dynamics of DSGE Models: Identification and Estimation," in "Model Evaluation Conference" Oslo May 2005.

Clarida, R., J. Gali, and M. Gertler, "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 2000, *115*, 147–180.

Clements, M. and D.F. Hendry, *Forecasting Non-stationary Economic Time Series*, MIT Press, 1999.

Cogley, T. and J.M. Nason, "Impulse Dynamics and Propagation Mechanisms in a Real Business Cycle Model," *Economics Letters*, 1993, *43*, 77–81.

Colleti, D., B. Hunt, D. Rose, and R. Tetlow, "The Bank of Canada's New Quarterly Projection Model, Part 3: The dynamic model: QPM," *Bank of Canada Technical Report*, 1996, *75.*

Davidson, J.E.H., D.F. Hendry, F. Srba, and S. Yeo, "Econometric Modelling of the Aggregate Time-Series Relationship Between Consumers' Expenditure and Income in the United States," *Economic Journal*, 1978, *88*, 661–692.

Enders, W., *Applied econometric time series*, first ed., New York: John Wiley and Sons., Inc., 1995.

Evans, G.W. and S. Honkapohja, *Learning and Expectations in Macroeconomics*, 2 ed., New Jersey: Princeton University Press, 2001.

___ **and** ___ , "Adaptive learning and Monetary Policy Design," *Journal of Money, Credit and Banking*, 2003, *35*, 1045–1072.

___ **and** ___ , "Expectations and the stability problem for optimal monetary policies," *Review of Economic Studies*, 2003, *70*, 807–824.

Fama, E. F. and M. R. Gibbons, "Inflation, real returns and capital investment," *Journal of Monetary Economics*, 1982, *9*, 297–323.

Forni, M., M. Lippi, and L. Reichlin, "Opening the Black Box: Structural Factor Models versus Structural VARs," *mimeo*, 2003, (ECARES, University Libre de Bruxelles).

Fuhrer, J.C. and G.P. Olivei, "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 March 2004.

Gali, J. and T. Monachelli, "Monetary Policy and Exchange Rate Volatility in a Small Open Economy," *NBER Working Paper*, 2002, *8850.*

Gramlich, E.M., "Remarks," in "Conference on Models and Monetary Policy" Federal Reserve Bank Board of Governors March 2004.

Granger, C.W.J. and P. Newbold, "Spurious regressions in econometrics," *Journal of Econometrics*, 1974, *2*, 111–120.

Gregory, A.W., A.R. Pagan, and G.W. Smith, *Models, Methods and Applications of Econometrics*, Blackwell, 1993.

Harrison, R., K. Nikolov, M. Quinn, G. Ramsay, A. Scott, and R. Thomas, "The Bank of England Quarterly Model," *Bank of England*, 2005.

Helliwell, J.F., G.R. Sparks, F.W. Gorbet, H.T. Shapiro, I.A. Stewart, and D.R. Stephenson, "The Structure of RDX2," *Bank of Canada Staff Study*, 1971, *7*.

Honkapohja, S. and K. Mitra, "Performance of Monetary Policy with Internal Central Bank Forecasting," *Journal of Economic Dynamics and Control*, 2005, *29* (forthcoming), 627–658.

Ireland, P., "A Method for Taking Models to the Data," *Journal of Economic Dynamics and Control*, 2004, *28*, 1205–1226.

Johansen, S., "What is the Price of Maximum Likelihood," in "Model Evaluation Conference" Oslo May 2005.

Kapetanios, G., A.R. Pagan, and A. Scott, "Making a Match: Combining Theory and Evidence in Policy-oriented Macroeconomic Modelling," *Journal of Econometrics*, 2005, *forthcoming.*

Kodde, D.A., F.C. Palm, and G.A. Pfann, "Asymptotic Least Squares Estimation: Efficiency Considerations and Applications," *Journal of Applied Econometrics*, 1990, *5*, 229–243.

Kwiatkowski, D., P. Phillips, P. Schmidt, and Y. Shin, "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Are Non Stationary?," *Journal of Econometrics*, 1992, *54*, 159–178.

Laxton, D. and P. Pesenti, "Monetary Policy Rules for Small, Open, Emerging Economies," *Journal of Monetary Economics*, 2003, *50*, 1109–1146.

Lubik, T.A., "A Simple Structural and Empirical Model of the Antipodean Transmission Mechanism," in "DSGE Workshop, reserve Bank of New Zealand" 2005.

____ **and F. Schorfheide**, *NBER Macroeconomics Annual*, NBER, 2005.

____ **and** ____ , "Do Central Banks Respond to Exchange Rate Movements? A Structural Investigation," *Journal of Monetary Economics*, forthcoming.

Mankiw, N.G. and J. Wolfers, "Disagreement about Inflation Expectations," *NBER Macroeconomics Annual*, 2003, *18.*

Mavroeidis, S., "Weak Identification of Forward-Looking Models in Monetary Economics," *Oxford Bulletin of Economics and Statistics*, 2004, *66*, 609–635.

Medina, J.P. and C. Soto, "Model for Analysis and Simulations (MAS)," in "Conference on Dynamic Stochastic General Equilibrium (DSGE) Models" Reserve Bank of New Zealand July 2005.

Milani, F., "Expectations, learning and macroeconomic persistence," *Journal of Monetary Economics*, 2007, *54*, 2065–2082.

Mishkin, F., "The information in the longer maturity term structure about future inflation," *The Quarterly Journal of Economics*, 1990, *105*, 815–828.

____, "What does the term structure tell us about future inflation?," *Journal of Monetary Economics*, 1990, *25*, 77–76.

Mishkin, F.S. and K. Schmidt-Hebbel, "Does inflation targeting make a difference?," *CNB Working Paper*, 2006, (13).

Murphy, C.W., *Macroeconomic Modelling in Australia*, Australian Economic Papers,

Nason, J.M. and G.W. Smith, "Identifying the New Keynesian Phillips Curve," *Federal Reserve Bank of Atlanta Working Paper*, 2005, *2005-1.*

Negro, M. Del, F. Schorfheide, F. Smets, and R. Wouters, "On the Fit and Forecasting Performance of New Keynesian Models," *Federal Reserve Bank of Atlanta Working Paper*, 2004, *2004-37.*

Nickell, S., "Error Correction, Partial Adjustment and All That," *Oxford Bulletin of Economics and Statistics*, 1985, *47*, 119–130.

Orphanides, A. and J. C. Williams, "Imperfect Knowledge, Iflation Expectations, and Monetary Policy," *Working Paper of the Federal Reserve Bank of San Francisco*, May 2002.

Phillips, P.C.B., "Understanding Spurious Regressions in Econometrics," *Journal of Econometrics*, 1986, *33*, 311–40.

Powel, A.A. and C.W. Murphy, *Inside a Modern Macroeconometric Model*, Vol. 428 of Lecture Notes on Economics and Mathematical Systems, Springer, 1995.

Silverman, B.W., *Density Estimation for Statistics and Data Analysis*, New York: Chapman and Hall, 1986.

Smets, F. and R. Wouters, "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, 2003, *1*, 1123–1175.

Stone, A., T. Wheatley, and L. Wilkinson, "A Small Model of the Australian Macroeconomy: An Update," *Reserve Bank of Australia Discussion Paper*, 2005, *2005-11.*

Towsend, R.M., "Forecasting the forecasts of others," *Journal of Political Economy*, 1983, *93*, 546–588.

Walsh, C.E., *Monetary Theory and Policy*, second ed., MIT Press, 2003.

Watson, M.W., "Measures of Fit for Calibrated Models," *Journal of Political Economy*, 1993, *101*, 1011–1041.

Appendix A

MSV representation

Using the method of undetermined coefficients, we derive the exact form of the minimum state variable (MSV) representation for the model considered in the text. Starting with the reduced form and assuming rational expectations, i.e., $\hat{E}_t^P(.) = \hat{E}_t^{CB}(.) = E_t(.)$, we get

$$Y_t = M_0 + (M_1 + M_2)E_t Y_{t+1} + P\epsilon_t, \tag{A2.1}$$

where

$$\epsilon_t = F\epsilon_{t-1} + \varepsilon_t.$$

Now assume the MSV form takes the form

$$Y_t = a + b\epsilon_t. \tag{A2.2}$$

Taking the appropriate expectations needed in (A2.1), one obtains

$$E_t Y_{t+1} \quad = \quad a + bF\epsilon_t.$$

Plugging these expectations back into (A2.1) yields

$$Y_t = M_0 + (M_1 + M_2)a + [(M_1 + M_2)bF + P]\epsilon_t. \tag{A2.3}$$

Using the method of undetermined coefficients, it follows that the MSV solution must satisfy

$$\begin{aligned} M_0 + (M_1 + M_2)a &= a, \\ (M_1 + M_2)bF + P &= b. \end{aligned}$$

Solving for the matrices a and b we get

$$\begin{aligned} a &= (I - M_1 - M_2)^{-1}M_0, \\ vec(b) &= [\mathbf{I} - F' \otimes (M_1 + M_2)]^{-1} vec(P). \end{aligned} \tag{A2.4}$$

Appendix B

Determinacy and E-stability

To analyze the conditions under which the incomplete knowledge model (2.4)-(2.7) converges to the true model, REE form, the methodology developed by Evans and Honkapohja (2001) is employed. In principle the methodology consists of two parts. First, the rational expectation equilibrium of the model is examined. I look for conditions under which the REE is *determined*. In the adaptive-learning terminology, the REE is said to be determined if it is found to be unique. Second, I check for the learnability of the REE. The question is, if economic agents have incomplete knowledge, can they learn the REE? The conditions that guarantee the REE are attainable under the adaptive learning mechanism and are called the *E-stability conditions*.[1]

REE Determinacy

To examine the rational expectation equilibrium of the model (2.4)-(2.6), we begin by rewriting the model in a matrix form *reduced form*

$$\hat{Y}_t = M_0 + M_1 \hat{E}_t^P \hat{Y}_{t+1} + M_2 \hat{E}_t^{CB} \hat{Y}_{t+1} + P s_t, \tag{A2.1}$$

where $\hat{Y}_t = [x_t, \pi_t]$, $s_t = [g_t, u_t]$, M_0 is an intercept vector. Because of a zero inflation target, all intercepts are zero. For that reason, I omit M_0 in further derivations.

$$M_1 = \begin{bmatrix} 1 & \sigma \\ \lambda & \beta + \lambda\sigma \end{bmatrix}, M_2 = \begin{bmatrix} -\sigma\theta_x & -\phi\theta_\pi \\ -\lambda\sigma\theta_x & -\lambda\sigma\theta_\pi \end{bmatrix}, P = \begin{bmatrix} 1 & 0 \\ 1 & \lambda \end{bmatrix}.$$

To analyze the REE determinacy, we shall assume for now a complete knowledge environment, $\hat{E}_t^P(.) = \hat{E}_t^{CB}(.) = E_t(.)$. Then rearranging the reduced form one obtains

$$\tilde{Y}_t = M E_t \tilde{Y}_{t+1} + P s_t, \tag{A2.2}$$

where $M = M_1 + M_2$.

Proposition 1. *The model (2.4)-(2.6) has a unique and stable rational expectations equilibrium if the eigenvalues of matrix M in (2.10) have real parts lest than one.*

[1]For details on the methodology, I refer to Evans and Honkapohja (2001) and Evans and Honkapohja (2003a), where adaptive learning in a homogeneous environment is explained, and to Honkapohja and Mitra (2005) for an extension to heterogeneous learning.

Proof standard outcome of the difference equation theory.

E-Stability

The second issue is to analyze the conditions under which the REE is learnable. We already know when the REE exists and is unique. We are now interested in whether, having incomplete knowledge, we can learn such an REE eventually. If the REE is determined, the model has the *minimum state variable* (MSV) representation

$$Y_t = a + bs_t. \tag{A2.3}$$

a, and b are the (3x1) and (3x3) matrices of the model primitives. Their exact form is derived in Appendix B.

We recall that the perceived law of motion (PLM) is

$$\hat{Y}_t = \hat{a}_t^i + \hat{b}_t^i s_t. \tag{A2.4}$$

$i = \{P, CB\}$. The subscript t on the matrices indicates the time dependence of the matrices as the agents learn using (2.7) and (2.8). s_t follows an AR(1) process, $s_t = Fs_{t-1} + e_t$, where e_t is white noise. The private agents and central bank use their PLMs to form expectations

$$\hat{E}_t^i \hat{Y}_{t+1} \;=\; \hat{a}_t^i + \hat{b}_t^i F s_t. \tag{A2.5}$$

Substituting (2.14) back into the reduced form (2.9), one obtains the economy's *actual law of motion* (ALM)

$$Y_t \;=\; \left(M_1 \hat{a}_t^P + M_2 \hat{a}_t^{CB}\right) + \left(P + M_1 \hat{b}_t^P F + M_2 \hat{b}_t^{CB} F\right) s_t. \tag{A2.6}$$

The mapping from PLM to ALM is formalized to

$$T[a, b] \;=\; [M_1 \hat{a}_t^P + M_2 \hat{a}_t^{CB}, P + M_1 \hat{b}_t^P F + M_2 \hat{b}_t^{CB} F], \tag{A2.7}$$

where $T : R^2 \rightarrow R$ is a map between perceived parameters and their true (equilibrium) values.

We are interested in the fixed point of T. Honkapohja and Evans (2002) show that E-stability is achieved if the steady state in the following differential equation is locally stable

$$\frac{d}{d\tau}(a, b) = T[a, b] - (a, b). \tag{A2.8}$$

Furthermore, Honkapohja and Mitra (2005) and Evans and Honkapohja (2003a) show that the map under heterogeneous and homogeneous expectations is equivalent. Using their result, I rewrite (A2.7) by equating $\hat{j}_t^P = \hat{j}_t^{CB} = \hat{j}_t$ for $j = \{a, b\}$. Hence, (A2.7) becomes

$$T[a, b] = \left[(M_1 + M_2)\hat{a}_t, P + (M_1 + M_2)\hat{b}_t F\right] \tag{A2.9}$$

and can be easily assessed.

Proposition 2. *The REE of the model (2.4)-(2.7) is E-stable under heterogeneous expectations if and only if the corresponding model with homogeneous expectations is E-stable. Hence, the real parts of the eigenvalues of*

$$\begin{aligned} DT_a(a) &= I \otimes (M_1 + M_2) \\ DT_b(b) &= F' \otimes (M_1 + M_2) \end{aligned}$$

must be less than one. $\otimes$ *is the Kronecker product.*[2]

Proof see Evans and Honkapohja (2003a) for the proof.

[2]Having the map from the PLMs to ALM

$$T[a, b] = \left[M_0 + (M_1 + M_2)\hat{a}_t, P + (M_1 + M_2)\hat{b}_t F)\right].$$

we take derivatives with respect to $\hat{a}_t$ and $\hat{b}_t$. Using the rules for the derivatives of matrices, we get

$$\begin{aligned} DT_a(a) &= \frac{d}{d\hat{a}_t}\left[M_0 + (M_1 + M_2)\hat{a}_t\right] = I \otimes (M_1 + M_2), \\ DT_b(b) &= \frac{d}{d\hat{b}_t}\left[P + (M_1 + M_2)\hat{b}_t\right] = F' \otimes (M_1 + M_2). \end{aligned}$$

Appendix C

Optimal Expectations-Based Policy Rule

The central bank minimizes a quadratic loss function

$$\min_{\{x_t,\pi_t\}} \qquad V = \frac{1}{2}E_t\left\{\sum_{i=0}^{\infty}\beta^i\left[\alpha x_{t+i}^2 + (1-\alpha)(\pi_{t+i}-\pi^*)^2\right]\right\}$$

subject to

$$\begin{aligned} x_t &= \hat{E}_t^{CB}x_{t+1} - \sigma\left(i_t - \hat{E}_t^{CB}\pi_{t+1}\right) \\ \pi_t &= \lambda x_t + \beta\hat{E}_t^{CB}\pi_{t+1}. \end{aligned}$$

Note that the central bank assumes that the private sector trusts the bank's expectations and adopts them for their own decisions. The central bank a priori assumes that monetary policy is credible. Further, we assume the bank does not observe current period exogenous shocks u_t and v_t.

The first order condition to the problem is

$$\alpha x_t + \alpha(1-\alpha)(\pi_t - \pi^*) = 0.$$

Using the FOC, the Phillips curve, and IS curve to solve for i_t, we obtain the optimal policy rule under discretion. When we assume that the inflation target π^* is zero, then the expectations-based policy rule takes the form

$$i_t = \theta_0 + \theta_\pi\hat{E}_t^{CB}\pi_{t+1} + \theta_x\hat{E}_t^{CB}x_{t+1},$$

where $\theta_\pi = 1 + \frac{(1-\alpha)\lambda\beta}{\lambda^2(1-\alpha)+\alpha}$, and $\theta_x = \frac{1}{\sigma}$, $\theta_0 = 0$.

Appendix D

Test critical values

In this appendix, the methodology for obtaining the critical values reported in the text is outlined. The methodology relies on Monte Carlo experiments. Mishkin (1990b) was the motivation for this approach.

The methodology can be summarized in the few following steps:

1. Analyze the time series of model variables, i.e., PRIBOR 1Y, year-to-year inflation, and inflation expectations, on the unit root.
2. Apply the Box-Jenkins methodology on the data. The outcome ought to be an ARIMA(p,d,q) model with the best fit as possible.
3. Simulate the estimated models from the previous step.
4. Using the simulated time series, estimate equation (3.8) and save the results on the t-tests.
5. Repeat steps 3.) and 4.) 10,000 times.
6. From step 5.), construct a new distribution for critical values.

The simulated critical values are reported in Table D.1.

Table D.1: The empirical critical values

Parameter	25%	10%	5%	1%
a_r	1.7339	3.6738	4.5052	7.2765
a	1.7232	3.4252	4.3978	6.8293
b	1.8041	3.4825	4.6014	6.8393

Appendix E

Tests for residuals stationarity

The regression residuals from (3.8) are tested for a unit root here. Testing for a unit root is crucial for the regression results. Because residuals are partially estimates of the real interest rate, there might be an economic reason to believe that a unit root is present. If it is so, then the results are spurious and non-usable.[1] To test for the unit root, the augmented Dickey-Fuller (ADF) and the Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests are employed. Each of them tests for a unit root but from a different view. The null-hypothesis of the ADF test is that a *time series is not stationary*, while the null of KPSS test is that *a series is stationary*. Applying both tests provides a more complex picture than using only one of them.

The results for the ADF test are summarized in Table E.1. Following Enders (1995), and the ADF test is based on testing $H_0 : \gamma = 0$ in

$$\Delta\hat{\epsilon}_t = \gamma\hat{\epsilon}_{t-1} + \Sigma_{i=2}^{p}\beta_i\Delta\hat{\epsilon}_{t-i+1} + \upsilon_t.$$

Table E.1: Augmented Dickey-Fuller Test Results

Parameter	Estimate	t-test	t-crit.	F-test (p-value)	Q-test (crit.val.)
γ	-0.2192	-7.5765	-1.95		
β_1	-0.0115	-0.1414	-1.69	5.65	17.28
β_2	-0.1728	-2.3414	-1.69	(0.02)	(25.70)

The critical values are for the 5% level of significance. As seen, the null-hypotheses of residuals being non-stationary is rejected ($\gamma \neq 0$).

The KPSS test results are summarized in Table E.2. The test statistics are computed for the lag truncation parameter, l, from 0 to 8. As argued by Kwiatkowski, Phillips, Schmidt and Shin (1992), for $l = 8$, the test has the largest power. Including $l = 0$, the test also accounts for autocorrelation. More details can be found in (Kwiatkowski et al. 1992). The critical value for the test at the 5% significance level is 0.463.

[1]For details on spurious regression, we refer to Granger and Newbold (1974) who experimentally demonstrated consequences of unit root on regression results and to Phillips (1986) for the formalization of the former.

Table E.2: Kwiatkowski-Phillips-Schmidt-Shin Test Results

l	0	1	2	3	4	5	6	7	8
$\hat{\epsilon}_t$	**1.10**	**0.62**	**0.47**	0.38	0.32	0.29	0.26	0.25	0.24

The results from the KPSS test suggest that the regression residuals from estimating (3.8) are stationary. For $l = 2$, the test result is on the margin of statistical significance. Putting the result together with the ADF test, it may be concluded that the residuals do not contain a unit root and follow a stationary process. Consequently, the parameter estimates reported in the text are unbiased and consistent, although, inefficient.

Appendix F

Collinearity test

In this appendix, the collinearity issue is addressed. First, a formal test for the presence of collinearity is performed. At the same time, the test is also a test of the rational expectations hypothesis. Second, the effect of collinearity on parameters' estimates and their standard errors is quantitatively analyzed.

The test for collinearity

The test is standard in considering the econometric literature. It is based on a *variance inflation factor* estimation. Because it tests for the linear relationship between π_{t+12} and $\pi^e_{t,12}$, the test is at the same time a test of the REH.

The REH is usually tested on the following inflation-prediction equation:

$$\pi_{t+12} = c_1 + c_2\pi^e_{t,12} + v_{t+12}.$$

If the REH holds, parameter c_1 is zero, c_2 is equal to 1, and v_{t+12} is an i.i.d. process with zero mean and finite variance. In this case, the two variables are evidently collinear.

When the inflation-prediction equation is estimated, the $\bar{R}^2$ is used to evaluate collinearity. To this purpose a variance inflation factor (VIF) is computed:

$$VIF = \frac{1}{1 - \bar{R}^2}.$$

Usually, we face a problem of collinearity if $VIF > 10$.

The estimation results of the inflation-prediction equation are summarized in Table F.1.

Table F.1: The REH test: $\pi_{t+12} = c_1 + c_2\pi^e_{t,12} + v_{t+12}$

Parameter	Estimate	t-stat.	Critical value 5%	DW	$\bar{R}^2$	No. of obs.
c_1	-4.3353	-2.9835				
c_2	1.7051	4.8067		0.16	0.34	44

Note: The critical values are simulated similarly as in Appendix A.

First, we can see that $c_1 \neq 0$, $c_2 \neq 1$, and the residuals are positively autocorrelated. As a consequence, the REH cannot be accepted. Second, a more important result, the VIF is 1.5, which is far from the values where collinearity causes estimation problems.

The Quantitative Assessment of Collinearity Effect

Let us derive the estimates $\{\hat{a}_r, \hat{a}, \hat{b}\}$ of equation (1.8):

$$i_t = a_r + a\pi_{t+12} + b\pi^e_{t,12} + \epsilon_t.$$

To find the estimates to the parameters, the ordinary least-square criterion is used:

$$\sum_{t=1}^{N} \left(i_t - a_r - a\pi_{t+12} - b\pi^e_{t,12}\right)^2 \rightarrow min.$$

Minimizing the criterion gives rise to a set of normal equations

$$\begin{aligned}
0 &= \sum 2(i_t - a_r - a\pi_{t+12} - b\pi^e_{t,12})(-1), \\
0 &= \sum 2(i_t - a_r - a\pi_{t+12} - b\pi^e_{t,12})(-\pi_{t+12}), \\
0 &= \sum 2(i_t - a_r - a\pi_{t+12} - b\pi^e_{t,12})(-\pi^e_{t,12}),
\end{aligned}$$

which can be conveniently rewritten as

$$\begin{aligned}
0 &= \bar{i} - \hat{a}_r - \hat{a}\bar{\pi} - \hat{b}\bar{\pi}^e, \\
0 &= cov(i_t, \pi_{t+12}) - \hat{a}\ var(\pi_{t+12}) - \hat{b}\ cov(\pi_{t+12}, \pi^e_{t,12}), \\
0 &= cov(i_t, \pi^e_{t,12}) - \hat{a}\ cov(\pi_{t+12}, \pi^e_{t,12}) - \hat{b}\ var(\pi^e_{t,12}),
\end{aligned}$$

where $\bar{x} = \frac{1}{N}\sum_{t=1}^{N} x_t$. Or in a matrix form

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} \begin{bmatrix} var(\pi_{t+12}) & cov(\pi_{t+12}, \pi^e_{t,12}) \\ cov(\pi^e_{t,12}, \pi_{t+12}) & var(\pi^e_{t,12}) \end{bmatrix} = \begin{bmatrix} cov(i_t, \pi_{t+12}) \\ cov(i_t, \pi^e_{t,12}) \end{bmatrix},$$

where $\begin{bmatrix} var(\pi_{t+12}) & cov(\pi_{t+12}, \pi^e_{t,12}) \\ cov(\pi^e_{t,12}, \pi_{t+12}) & var(\pi^e_{t,12}) \end{bmatrix}$ is the information matrix $X'X$.

Solving for the parameter estimates gives

$$\begin{aligned}
\hat{b} &= \frac{1}{1-\rho^2}\frac{cov(i_t, \pi^e_{t,12})}{var(\pi^e_{t,12})} - \frac{\rho}{1-\rho^2}\frac{cov(i_t, \pi_{t+12})}{std(\pi_{t+12})\ std(\pi^e_{t,12})}, \\
\hat{a} &= \frac{cov(i_t, \pi_{t+12})}{var(\pi_{t+12})} - \hat{b}\rho\frac{std(\pi^e_{t,12}}{\pi_{t+12}}, \\
\hat{a_r} &= \bar{i} - \hat{a}\bar{\pi} - \hat{b}\bar{\pi}^e,
\end{aligned}$$

where $\rho = \frac{cov(\pi_{t+12},\pi^e_{t,12})}{std(\pi_{t+12})\, std(\pi^e_{t,12})}$.[1]

Next, to evaluate the impact of collinearity on the t-test, we have to analyze its effect on the parameters' standard errors. For simplicity, let us assume that regression residuals are homoscedastic and uncorrelated. Then the parameters' variance can be expressed as

$$\begin{aligned}
var\begin{bmatrix}\hat{a}\\ \hat{b}\end{bmatrix} &= \sigma^2(X'X)^{-1} = \sigma^2\begin{bmatrix} var(\pi_{t+12}) & cov(\pi_{t+12},\pi^e_{t,12}) \\ cov(\pi^e_{t,12},\pi_{t+12}) & var(\pi^e_{t,12})\end{bmatrix}^{-1} = \\
&= \sigma^2\begin{bmatrix} var(\pi_{t+12}) & \rho\, std(\pi_{t+12})\, std(\pi^e_{t,12}) \\ \rho\, std(\pi^e_{t,12})\, std(\pi_{t+12}) & var(\pi^e_{t,12})\end{bmatrix}^{-1} \\
&= \frac{\sigma^2}{var(\pi_{t+12})var(\pi^e_{t,12})(1-\rho^2)}\begin{bmatrix} var(\pi_{t+12}) & \rho\, std(\pi_{t+12})\, std(\pi^e_{t,12}) \\ \rho\, std(\pi^e_{t,12})\, std(\pi_{t+12}) & var(\pi^e_{t,12})\end{bmatrix} \\
var(\hat{a}_r) &= \bar{\pi}^2 var(\hat{a}) + \bar{\pi^e}^2 var(\hat{b}).
\end{aligned}$$

The analysis is limited for the above case only. Because the simulated critical values account for collinearity, this limitation is acceptable for a further quantitative analysis.

To quantitatively analyze the effect of collinearity on the parameters' estimates and t-tests, their values are simulated for different magnitudes of correlation, ρ, between π_{t+12} and $\pi^e_{t,12}$. The analysis is conducted on the following set of descriptive statistics:

Statistics	Estimate
ρ	0.6499
$cov(i_t, \pi_{t+12})$	1.7089
$cov(i_t, \pi^e_{t,12})$	0.7157
$std(\pi_{t+12})$	2.0678
$std(\pi^e_{t,12})$	0.7115
$var(\pi_{t+12})$	4.2760
$var(\pi^e_{t,12})$	0.2548
$\bar{i}$	5.1581
$\bar{\pi}$	2.5791
$\bar{\pi^e}$	4.0116

Table F.2: The Calibrated Values - Actual Data Statistics

From the graphical results (Figures F.1-F.3), it follows that only parameter a is sensitive to the value of ρ. Its estimate differs considerably for $\rho = 0$ and $\rho = .6$ and so

[1]Note that when $\rho = 0$, the solution has the standard form

$$\begin{aligned}
\hat{b} &= \frac{cov(i_t, \pi_{t+12})}{var(\pi^e_{t,12})}, \\
\hat{a} &= \frac{cov(i_t, \pi_{t+12})}{var(\pi_{t+12})}, \\
\hat{a_r} &= \bar{i} - \hat{a}\bar{\pi} - \hat{b}\bar{\pi^e}.
\end{aligned}$$

do the t-statistics. What is, however, important for our purposes, the parameter is not statistically significant for $\rho = (.5, .9)$. By the nature of expectations, the correlation between $\pi^e_{t,12}$ and π_{t+12} should not be very low, even though the REH does not hold.

The parameter a_r is always statistically insignificant, and b, is statistically significant under the simulation setup. Hence, it might be concluded that the results presented in the paper are robust for $\rho = (0.5, 0.9)$.

Figure F.1: Dependence of $\hat{a_r}$ and its t-statistics on ρ.

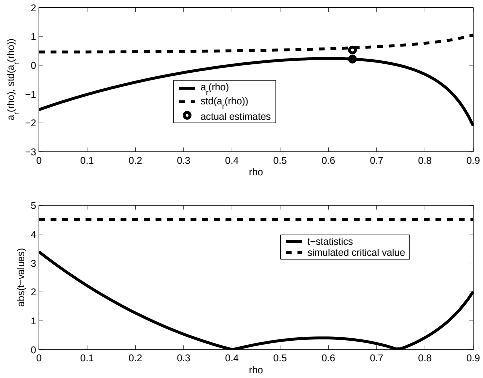


Figure F.2: Dependence of $\hat{a}$ and its t-statistics on ρ.

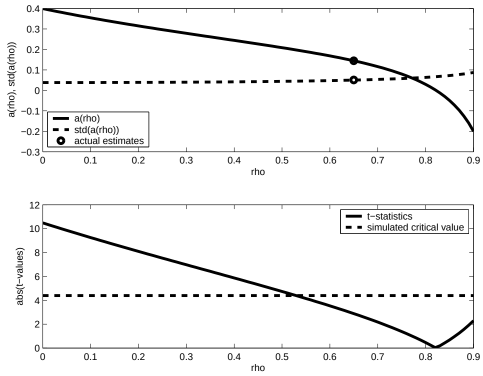


Figure F.3: Dependence of $\hat{b}$ and its t-statistics on ρ.