NCER

National Centre for Econometric Research

NCER Working Paper Series

Limited Information Estimation and Evaluation of DSGE Models

Martin Fukac and Adrian Pagan

Working Paper #6

November 2006

Limited Information Estimation and Evaluation of DSGE Models *

Martin Fukac† and Adrian Pagan‡

November 27, 2006

Abstract

We advance the proposal that DSGE models should not just be estimated and evaluated with reference to full information methods. These make strong assumptions and therefore there is uncertainty about their impact upon results. Some limited information analysis which can be used in a complementary way seems important. Because it is sometimes difficult to implement limited information methods when there are unobservable non-stationary variables in the system we present a simple method of overcoming this that involves normalizing the non-stationary variables with their permanent components and then estimating the estimating the resulting Euler equations. We illustrate the interaction between full and limited information methods in the context of a well-known open economy model of Lubik and Schorfheide. The transformation was effective in revealing possible mis-specifications in the equations of LS's system and the limited information analysis highlighted the role of priors in having a major influence upon the estimates.

Contents

*Research of the second author was supported by ARC Grant No. DP0449659

†Centre for Economic Research and Graduate Education in Prague and Czech National Bank

‡Queensland University of Technology.

1 Introduction

DSGE models are becoming widely used in both academic and central bank research. In the case of the latter there is naturally great interest in the ability of the models to adequately represent the data. The question of how to do this engaged econometricians during the second half of the 20th century. Initially emphasis was placed upon ways of summarizing *system fit* because the recommended estimation method for the system parameters was FIML. After 2SLS emerged as the estimator that was most widely used, largely for computational reasons, more attention was paid to how the individual equations of the system fitted the data. In the late 60s and early 70s however, many model builders became dissatisfied with such an orientation, largely due to their experience that, whilst the individual equations seemed to fit the data closely, when combined into a system there were obvious deficiencies. This led them to recommend that one should study the properties of complete models as part of the model development process - an attitude summed up very well by the adage " Simulate early and simulate often".[1] An important part of this re-orientation was a focus upon seeing how well the complete model tracked the data.

Within the DSGE tradition the emphasis on model evaluation seems to have begun with a focus upon the moments of a small set of variables. More recently, there has been an increasing use of system estimation and evaluation methods that derive from FIML (and Bayesian versions of it), probably because of the improved computational facilities. It is arguable that this approach is now the norm and single equation methods have been largely ignored.[2] There are some reasons for this shift involving a possible improved estimator performance when a systems estimator is adopted, an argument which needs more detailed analysis than is possible in this paper. But the shift in emphasis has also meant that what evaluation has been done on these models is largely from the system perspective, basically offering a comparison with a VAR, and rarely involves an examination of the individual structural equations of the DSGE model i.e. of the Euler equations. This seems unfortunate. The stimulus to system-wide measures of macroeconomic model performance in the 1970s arose since the single equations of the models seemed to fit the data well, and therefore evaluation tools were needed that treated the system as a whole. These were viewed as a complement rather than a substitute to single equation methods. For this reason it seems useful to examine the Euler equations of any DSGE model in order to determine the extent to which they fit the data as a *supplement* to any systems tests. One advantage of this approach is that it is often easier to see where the specification of the DSGE model is weak, and any such information can suggest suitable re-specifications. Section 2 of the paper therefore sets out the ways in which one might want to estimate and test the Euler equations when these are taken individually, rather than as a complete system. To do this we need to estimate the parameters of the individual equations. One might use the FIML estimates of the complete system but it is more logical to look at LIML estimates that only use the information contained in the Euler equation under study, and we set out various ways in which this might

[1] We owe this quote to the late Chris Higgins.

[2] An exception is the literature on the small system known as the New Keynesian Policy Model that incorporates a Phillips curve, an IS curve and an interest rate rule. Even there however, although estimation has sometimes been by single equation methods, this has not been true of evaluation.

be done.

Performing evaluation at the level of the Euler equations may not be easy, largely due to the presence of unobserved variables, and this is particularly apparent when one has unobserved factors which are designed to account for the fact that variables in a DGSE model (and the data) are integrated. In section 3 we outline a method to get around this problem. In doing so we consider whether it is possible to utilize data in the Euler equations from which a permanent component has been removed via some filtering operation. We show that such a strategy can lead to inconsistent estimates of the parameters of the Euler equations and that one needs to allow for the nature of the filter that removed the permanent component in order to avoid such biases. In section 4, we utilize the methods to study the adequacy of a well known open-economy model used by Lubik and Schorfheide (2006). Section 5 concludes.

2 DSGE Model Structure and Estimation

DSGE models have the following stylized representation

$$B_0 z_t = B_1 z_{t-1} + D x_t + C E_t z_{t+1} + G u_t \tag{1}$$

where z_t is a vector of $n \times 1$ variables, x_t is a set of observable, and u_t a set of unobservable shocks. Generally z_t will be the logs of variables. 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 in this paper. By observable we will mean that the shocks can be recovered from a statistical model of x_t. By unobservable we will mean that the shocks are defined by the economic model. The system above consists of a set of Euler equations describing optimal choices and a set of identities. The latter may be associated with income identities and the equations describing the evolution of stocks. We ignore the latter although these will of course be crucial to whether the complete system is able to adequately match the data. The parameters in the DSGE model will be designated as θ.

The solution to this system has the form

$$z_t = P z_{t-1} + \sum_{j=0}^{\infty} \Pi_1^j (\Pi_2 E_t x_{t+j} + \Pi_3 E_t G u_{t+j})$$

where P satisfies $B_0P - B_1 - CP^2 = 0, \Pi_1 = (B_0 - CP)^{-1}C, \Pi_2 = (B_0 - CP)^{-1}D$ and $\Pi_3 = (B_0 - CP)^{-1}G$. In the case where the x_t and u_t are AR(1) processes with matrices Φ_x and Φ_u this reduces to a Vector Autoregression with Exogenous Variables (VARX) system for z_t of the form

$$z_t = Py_{t-1} + \overline{D}x_t + \overline{G}u_t, \tag{2}$$

where $\overline{D} = \sum \Pi_1^j \Pi_2 \Phi_x^j$ and $\overline{G} = \sum \Pi_1^j \Pi_3 \Phi_u^j$. Using (2) one can then find an expression for $E_t z_{t+1}$

$$E_t(z_{t+1}) = Pz_t + \overline{D}E_t(x_{t+1}) + \overline{G}E_t(u_{t+1}).$$

Hence one could solve for the conditional expectation of z_{t+1} once $P, \overline{D}$ and $\overline{G}$ are known. Since these are functions of θ, once θ is estimated we can construct the expectation in a way that is consistent with the DSGE model. Alternatively, one could estimate these parameters in an unconstrained way either by regressing z_{t+1} on z_t and x_t (if u_t was white noise) or z_t against z_{t-1}, z_{t-2}, x_t and x_{t-1} (if there was a VAR(1) in u_t).

Instead of estimating the complete system we believe it is worthwhile using only the information in the Euler equation of interest to produce estimates of its parameters. We will focus upon a representative Euler equation in the system of the form

$$z_{1t} = B_{10}z_t + B_{11}z_{t-1} + D_1x_t + C_1E_tz_{t+1} + \zeta_{1t}, \tag{3}$$

where we have normalized on one of the endogenous variables in the equation, with others appearing on the RHS in $B_{10}z_t$. The unknown parameters in B_{10}, B_{11} etc. will be termed η. Now some of the DSGE parameters will appear in this equation. These will be θ_1. It may not be possible to estimate θ_1 from η as the $\dim(\theta_1)$ may exceed $\dim(\eta)$. However, assuming for the moment that it is possible, then the question arises of whether we can identify and estimate η. Problems in estimating η will arise from the presence of RHS endogenous variables and E_tz_{t+1}. In standard simultaneous equation methods the first problem is effectively overcome by constructing a synthetic system composed of the simultaneous equation whose parameters are being estimated and auxiliary equations which determine any RHS endogenous variables. Let the auxiliary equation parameters be ϕ. If η and ϕ are estimated jointly one would be performing LIML. Pagan (1979) used this idea to explore the relation of LIML and 2SLS. The latter is of course found by applying two-step maximum likelihood to the synthetic system but with ϕ replaced by OLS estimates. The ideas above readily extend to the Euler equation case - the auxiliary VARX system representing any RHS endogenous variables is combined with the Euler equation to form estimates of $E_t z_{t+1}$ and these are then effectively used as a regressor. One advantage of constructing estimates in this way is that the same software can be utilized to generate both full information and limited information estimates. Another advantage is that one can generate limited information Bayesian estimates of θ_1, provided of course it is possible to recover these parameters using limited information.

Once the η parameters have been estimated one can utilize them to either test the consistency of the assumptions made in generating the systems estimators or to examine the specification of the Euler equation. If it is possible to estimate θ_1 one might construct a Hausman test based on comparing the limited and full information estimates of θ_1. If only a sub-set of the θ_1 are identified using limited information then we could do such a comparison with the remaining parameters in θ_1 set to the full information estimates. One might also perform specification tests on the Euler equation itself by specifying a new form and re-estimating its parameters. It seems preferable to do this with limited information methods so as to avoid the possibility that the specification error may be elsewhere in the system and it is causing biases in the estimates of the parameters of this Euler equation. By using the sub-system estimator one avoids the problems of contamination caused by specification errors in other equations. Finally, by using limited information methods one can better break find which sets of information contribute most to the parameter estimates. The precision of these estimates may be influenced by the use of a systems estimator but it can also be a consequence of the imposition of prior information if the data is not very informative about the parameter. Separating these two influences on precision seems important. By using a limited information estimator we can abstract from the the first influence and can therefore study the impact of the prior much more directly.

3 Adapting Estimation to Handle $I(1)$ Shocks

3.1 The Effects of Filtering

The discussion above has proceeded as if z_t was a stationary random variable. Where the situation becomes more complex is if the observed data is an $I(1)$ process and the factors driving it are unobservable. Recourse is often had to first filtering the data to remove a permanent component and then working with the filtered data in place of the original levels variables in the Euler equations. But, since the basic Euler equations relate to the levels of the variables, it is not clear that they will still hold once a permanent component is filtered out. Indeed, as we will demonstrate, one normally needs to adjust the Euler equations when one is using filtered variables. Failure to do so introduces an omitted variable and this can cause inconsistent estimators of the parameters of these equations.

To illustrate the argument we work with the Euler equation (3). For convenience we will assume that there are no RHS endogenous variables entering contemporaneously i.e. $B_{10} = 0$. Since the variables z_t are in logs we transform the $I(1)$ variables to a new series $\overline{z}_t = z_t - z_t^f$, which are stationary, and where the vector of permanent components z_t^f are constructed by filtering the data z_t. In doing this we assume that the permanent component is unobservable. If it derived from x_t filtering would not be needed as we would utilize the relationship to remove the permanent component of z_t.

The Euler equation is re-expressed in terms of transformed variables to produce

$$\begin{aligned} z_{1t} - z_{1t}^f &= B_{11}(z_{t-1} - z_{t-1}^f) + C_1 E_{t+1}(z_{t+1} - z_{t+1}^f) + Dx_t \\ &\quad + C_1 E_t(z_{t+1}^f) + B_{11} z_{t-1}^f - z_{1t}^f + \zeta_{1t}, \end{aligned}$$

where $\zeta_{1t} = Gu_t$. Therefore

$$\begin{aligned} \overline{z}_{1t} &= C_1 E_t(\overline{z}_{t+1}) + B_{11}\overline{z}_{t-1} + \{C_1 E_t(z_{t+1}^f) + B_{11} z_{t-1}^f - z_{1t}^f\} \\ &\quad + Dx_t + \zeta_{1t} \end{aligned} \tag{4}$$

It is clear from (4) that one cannot assume that the Euler equations will apply to the transformed variables unless it can be verified that the term in curly brackets is zero. In general this will not be true. To understand why this is so it is useful to consider the special case where z_{1t} is the only $I(1)$ variable, so that we don't need to transform any contemporaneous RHS endogenous variables, $C_1 E_t \overline{z}_{t+1} = E_t z_{1t+1}$, and $C_1 + B_{11} = 1$. The last restriction often occurs with models such as this (so called hybrid models). Putting these features together makes the term under examination

$$\begin{aligned} & \{C_1 E_t(z^f_{1,t+1} - z^f_t) + B_{11}(z^f_{t-1} - z^f_{1t}\} \\ = \quad & C_1 E_t \Delta z^f_{t+1} - B_{11} \Delta z^f_t . \end{aligned} \tag{5}$$

Now suppose the Beveridge-Nelson (BN) decomposition is used for computing z^f_t. We know that, regardless of whether the filtered estimate is constructed from either multivariate or univariate data, and also regardless of whether the variables are co-integrated or not, the BN estimate of z^f_t has the property that Δz^f_t is white noise, so that $E_t \Delta z^f_{t+1} = 0$. Consequently, with this filter only the term $B_{11} \Delta z^f_t$ is left in (5). Whether this is uncorrelated with ζ_{1t} in (3) depends upon the specification of the DSGE model and we will return to this in the next section. But we do know that $E_t \Delta z^f_{t+1}$ will generally not be zero for other filters.

To analyze what happens if z^f_t is formed using (say) the HP filter is complicated by the fact that it is a two-sided filter with time varying weights. There is however a version that has a "steady state" solution of the form

$$z^f_t = \sum_{j=-T}^{T} a_j z_{t-j}.$$

Singleton (1988) gives the weights a_j as ($\lambda = 1600$)

$$a_j = 1 - \{.894j[.056\cos(.112j) + .0558\sin(.112j)\}$$

If one looks at this expression for z^f_t it is clear that, due to the terms $\sum_{k=0}^{T} a_{-j} z_{t-j}$, $E_t(\Delta z^f_{t+1})$ will never be zero, even if Δz_t is white noise. Ignoring the fact that the term $E_t(\Delta z^f_{t+1})$ is not zero will generally bias any estimators applied to the Euler equations. Moreover it can be seen that, in the case of the HP filter, Δz^f_t has a unit root, so that the error term in the transformed Euler equation will have a unit root, even though the original shock in it may not have. This is a consequence of Harvey and Jaeger's (1993) demonstration that the underlying assumption about the DGP of z_t used in producing the HP filter is that z_t is $I(2)$. To illustrate the effect we simulated 500 observations from a DGP for z_t of the form $\Delta z_t = e_t$, where e_t is white noise, and then computed z_t^f using the HP filter ($\lambda = 1600$). The regression of the simulated Δz_t^f against Δz_{t-1}^f gives an estimated coefficient on the latter variable of 1.00. This may be a reason why one sees so many shocks having roots that are very close to unity in estimated DSGE models that have variables transformed using the HP filter.

3.2 The Single Permanent Component Case

We look at a case that is very common in DSGE models viz. when there is a single permanent component driving the system. Generally this is the log of technology, a_t. Often investigators transform the variables z_t to $\tilde{z}_t = z_t - ia_t$, where i is a unit vector, and then express the Euler equation in terms the transformed variables. But, as $\tilde{z}_t$ is unobservable, it is hard to evaluate the fit of such an equation. Consequently, we seek to replace $\tilde{z}_t$ by an observable quantity. To see how this is done we start with the "common trends representation", $z_t = J\tau_t + \nu_t = z_t^p + \nu_t$, where $\Delta\tau_t = \varepsilon_t$ are the common trends of the system, ε_t is white noise, and ν_t is some $I(0)$ component. Since there is only one $I(1)$ factor in the majority of DSGE models (and this is implicitly assumed in the discussion above), this must mean with $\tau_t = a_t^p$, with a_t^p being the permanent component of a_t and the latter will be measured by the Beveridge-Nelson decomposition. Hence

$$\begin{aligned}
\tilde{z}_t &= z_t - ia_t = Ja_t^p + \nu_t - ia_t \\
&= Ja_t^p + \nu_t - i(a_t^p + \xi_t) \\
&= (J - i)a_t^p + \nu_t - i\xi_t
\end{aligned}$$

and, because $\nu_t - i\xi_t$ is $I(0)$, it must be that $(J - i) = 0$ if $\tilde{z}_t$ is to be $I(0)$. Hence, in this one factor model, the permanent component of all series are the same. Note that z_t^p can be estimated with either a univariate or multivariate Beveridge-Nelson decomposition as in both cases ε_t would be white noise.

Now defining $\overline{z}_t = z_t - z_t^p$, it must be the case that $\overline{z}_t$ and $\tilde{z}_t$ are related in the following way

$$\begin{aligned}
\overline{z}_t &= z_t - z_t^p = z_t - ia_t + ia_t - z_t^p \\
&= \tilde{z}_t + i(a_t - a_t^p),
\end{aligned}$$

and $a_t - a_t^p$ will be the transitory component of a_t. Once a process for a_t is specified this can be calculated. If a_t is an $AR(1)$ (the most common assumption in DSGE models), with coefficient ρ_a and shock $\varepsilon_t,$ the transitory and permanent components of a_t have the properties

$$\begin{aligned}\Delta a_t^p &= \frac{\varepsilon_t}{1-\rho_a} \\ (a_t - a_t^p) &= -\frac{\rho_a}{1-\rho_a}\Delta a_t\end{aligned}$$

leading to

$$\therefore \tilde{z}_t = \overline{z}_t + i\frac{\rho_a}{1-\rho_a}\Delta a_t$$

It is also often useful to note that

$$\Delta z_t^p = i\Delta a_t^p = i\frac{\varepsilon_t}{1-\rho_a} = i\frac{(1-\rho_a L)\Delta a_t}{1-\rho_a},$$

and therefore

$$\Delta a_t = (i'i)^{-1}i\frac{1-\rho_a}{1-\rho_a L}\Delta z_t^p. \tag{6}$$

so that we can write

$$\tilde{z}_t = \overline{z}_t + i(i'i)^{-1}i\frac{\rho_a}{1-\rho_a L}\Delta z_t^p, \tag{7}$$

and the RHS variables are observable. Given the relation between $\overline{z}_t$ and $\tilde{z}_t$ we can replace $\tilde{z}_t$ in any Euler equation before estimation. We utilize this method in the next section.

4 An Example

4.1 The Model

The model we utilize is in Lubik and Schorfheide (2005) (LS). It is a small four equation model of an open economy. The IS curve describes output y_t and is specified in their paper in terms of the transformed variable $\tilde{y}_t = y_t - a_t$, where a_t is the log level of technology. Technology has an $I(1)$ structure and is the single $I(1)$ variable in the model. Hence it is a single permanent component model of the type described in the previous section. [3] The same transform is applied to (unobservable) foreign output to produce $\tilde{y}_t^*$.

$$
\begin{aligned}
\tilde{y}_t &= E_t\tilde{y}_{t+1} - [\tau+\theta](R_t - E_t\pi_{t+1}) - \alpha(\tau+\theta)E_t\Delta q_{t+1} \\
&\quad +\rho_A\Delta a_t + (\theta/\tau)E_t\Delta\tilde{y}_{t+1}^*, \qquad (8)\\
0 &< \alpha < 1, \tau^{-1} > 0
\end{aligned}
$$

In the IS equation q_t is the observable terms of trade, A_t is the log level of (unobservable) technology, α is the import share, τ is the intertemporal elasticity of substitution and $\theta = \alpha(2-\alpha)(1-\tau)$.

Their open economy Phillips curve is

$$
\begin{aligned}
\pi_t &= \beta E_t\pi_{t+1} + \alpha\beta E_t\Delta q_{t+1} - \alpha\Delta q_t + \frac{\kappa}{\tau+\theta}\tilde{y}_t \\
&\quad +\frac{\kappa\theta}{\tau[\tau+\theta]}\tilde{y}_t^*, \qquad (9)
\end{aligned}
$$

where π_t is the domestic inflation rate, β is the discount factor and κ is a "price stickiness" parameter.

The exchange rate equation is

$$\Delta e_t - \pi_t = -(1-\alpha)\Delta q_t - \pi_t^*, \qquad (10)$$

where e_t is the log of the exchange rate and π_t^* is the (unobservable) foreign inflation rate.

The policy rule for the nominal interest rate (R_t) is

$$R_t = \rho_R R_{t-1} + (1-\rho_R)\left[\psi_1\pi_t + \psi_2\tilde{y}_t + \psi_3\Delta e_t\right] + \varepsilon_t^R. \qquad (11)$$

Exogenous variables evolve as AR(1) processes where shocks ε_t are all *i.i.d.*

$$
\begin{aligned}
\Delta q_t &= \rho_q\Delta q_{t-1} + \varepsilon_t^q. && (12)\\
\Delta a_t &= \rho_a\Delta a_{t-1} + \varepsilon_t^a && (13)\\
\tilde{y}_t^* &= \rho_{y^*}\tilde{y}_{t-1}^* + \varepsilon_t^{y^*}, && (14)\\
\pi_t^* &= \rho_{\pi^*}\pi_{t-1}^* + \varepsilon_t^{\pi^*}, && (15)
\end{aligned}
$$

[3] In fact these are deviations from their steady state values but this would only introduce intercepts into the equations so we ignore that in the exposition below.

4.2 Transforming the Model

The Euler equations will now be transformed to be in terms of observables $\overline{y}_t = y_t - y_t^p$, where the "p" indicates the permanent component. Given the nature of a_t, and using $E_t(\Delta y_{t+1}^*) = (\rho_{y^*} - 1)y_{t-1}^*$ and $E_t(\Delta q_{t+1}) = \rho_q \Delta q_t$, the IS equation becomes

$$\begin{aligned} \overline{y}_t &= E_t\overline{y}_{t+1} - [\tau + \theta](R_t - E_t\pi_{t+1}) - \alpha(\tau+\theta)\rho_q\Delta q_t \\ &\quad -\frac{\theta}{\tau}(1-\rho_{y^*})\tilde{y}_t^*, \end{aligned} \tag{16}$$

Now the Euler equation only has one shock in it. This should not be surprising as this is a case where $B_{11} = 0$ in (5), and so no extra terms appear in the Euler equation if one uses the Beveridge-Nelson permanent component as the normalizing element. This would not be true if we had used HP filtered data. The presence of the term $\rho_a \Delta a_t$ is then seen to stem from the fact that $a_t \neq a_t^p$ whenever there is serial correlation in the growth rates of technology. Hence the choice of normalizing factor a_t is not the best when there is serial correlation in technology growth.

In a similar way we can re-write the Phillips curve and interest rate rules as[4]

$$\begin{aligned} \pi_t &= \beta E_t\pi_{t+1} - \alpha(1-\beta\rho_q)\Delta q_t + \frac{\kappa}{(\tau+\theta)}\overline{y}_t \\ &\quad + \frac{\kappa\rho_a}{(\tau+\theta)(1-\rho_a L)}\Delta y_t^p + \frac{\kappa\theta}{\tau[\tau+\theta]}\tilde{y}_t^* \end{aligned} \tag{17}$$

$$\begin{aligned} R_t &= \rho_R R_{t-1} + (1-\rho_R)\left[(\psi_1+\psi_3)\pi_t + \psi_2\overline{y}_t + \psi_3(\Delta e_t - \pi_t)\right] \\ &\quad + \psi_2\frac{(1-\rho_R)\rho_a\Delta y_t^p}{(1-\rho_a L)} + \varepsilon_t^R \end{aligned} \tag{18}$$

It is worth noting that the move to $\overline{y}_t$ means that the (17) and (18) need to be adjusted to reflect this change and an additional variable Δy_t^p appears in

[4] We can see some interesting features of this model. If $\rho_a = 0$ then the structural errors in the IS and Phillips curves are proportional to $\tilde{y}_t^*$ and so they are perfectly correlated. This means that there may be singularity problems in the system and this would create difficulties for MLE.

them. Hence care has to be exercised when using filtered variables and the appropriate adjustments made to ensure that these do not create estimation biases.

For later reference we will want to consider a hybrid form of the curve. Then the underlying Euler equation will be of the form

$$\begin{aligned} y_t &= cE_t y_{t+1} + (1-c)y_{t-1} - [\tau+\theta](R_t - E_t\pi_{t+1}) - \alpha(\tau+\theta)E_t\Delta q_{t+1} \\ &\quad -(\theta/\tau)E_t\Delta\tilde{y}^*_{t+1}. \end{aligned} \tag{19}$$

Consequently, if we normalize with a_t, the term $\rho_a\Delta a_t$ in (8) will become $(c-1+c\rho_a)\Delta a_t$. Thereupon, when we use the IS equation expressed in terms of $\overline{y}_t$, there will be an extra term $-(1-c)\Delta y^p_t$ present in it.

4.3 Solving the Transformed Model

As discussed in section 2 the solution to the transformed model (with a hybrid IS curve) will have the form

$$z_t = \Gamma_1 R_{t-1} + \Gamma_2\Delta q_t + \Gamma_3 v_t$$

where $z'_t = \begin{bmatrix} \overline{y}_t & \pi_t & \Delta e_t - \pi_t & R_t \end{bmatrix}$ and $v'_t = \begin{bmatrix} \Delta A_t & \tilde{y}^*_t & \pi^*_t & \varepsilon^R_t \end{bmatrix}$. After substituting

$$v_t = \Phi v_{t-1} + \varepsilon_t$$

into this expression we have

$$\begin{aligned} z_t &= \Gamma_1 R_{t-1} + \Gamma_2\Delta q_t + \Gamma_3\Phi v_{t-1} + \Gamma_3\varepsilon_t & (20) \\ &= \Gamma_1 R_{t-1} + \Gamma_2\Delta q_t + \Gamma_3\Phi\Gamma_3^{-1}(z_{t-1} - \Gamma_1 R_{t-2} - \Gamma_2\Delta q_{t-1}) + \Gamma_3\varepsilon_t & (21) \end{aligned}$$

DYNARE gives the solution (20). From that output one can determine Γ_1, Γ_2 and Γ_3 and, with known Φ, it is possible to derive the equation (21), which, apart from the white noise shocks, is expressed in terms of observables. Now consider the equation for $\overline{y}_t = Sz_t$, where $S = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$. This will be

$$\overline{y}_t = w'_t\delta + S\Gamma_2\Delta q_t + S\Gamma_3\varepsilon_t$$

where $w'_t = \begin{bmatrix} R_{t-1} & \overline{y}_{t-1} & \pi_{t-1} & \Delta e_{t-1} - \pi_{t-1} & R_{t-2} & \Delta q_{t-1} \end{bmatrix}$. Hence

$$E_t(\overline{y}_{t+1}) = w'_{t+1}\delta + S\Gamma_2\rho_q\Delta q_t,$$

showing that the $E_t(\overline{y}_{t+1})$ can be found once Γ_1, Γ_2 and Γ_3 are computed. It would also be possible to form an estimate of $E_t(\overline{y}_{t+1})$, either by a regression of $\overline{y}_{t+1}$ against w_{t+1} and Δq_t or by estimating the parameters in δ and $\Gamma_2\rho_q$ freely. The same is true of $E_t(\pi_{t+1})$.

5 Analysis of the Equations of the LS Model

5.1 Full Information Estimation

Table 1 gives the Full Information Bayesian posterior estimates found after applying the priors that LS used.[5]. We also present results that use a different prior for ρ_q. As pointed out in Fukac and Pagan (2006) the OLS estimate of ρ_q found by regressing Δq_t against Δq_{t-1} is negative, but the prior used by LS precludes negative values. Hence we have replaced that prior with a normal density. However, this change only has minor effects. Later when we perform model evaluation we use the original LS estimates. Table 1 estimation results come from applying Dynare Version 3.064. Table 2 shows what happens when we move to a system transformed to observables i.e. with the system expressed in terms of $\tilde{y}_t$ and also $\overline{y_t}$. There is no reason to think that these will be the same unless the model is correctly specified, since the $\tilde{y}_t$ equations contain a variable that depends on how a_t is generated, while using $\overline{y}_t$ makes no such assumption. There are some notable differences - the influence of the exchange rate upon interest rate decisions that was the main focus of attention in LS is markedly reduced and the estimates of α, τ and κ are much less precise. This suggests that we need to look at the estimates obtained when only limited information is used.

5.2 Limited Information Analysis

5.2.1 The Real Exchange Rate Equation

To begin the analysis of the LS system one should start with (10). Under the assumptions made in this model Δq_t and π_t^* are strongly exogenous processes, and so (10) is actually a regression equation, with $\Delta e_t - \pi_t$ as dependent variable, Δq_t as the regressor and with first order serially correlated errors.

[5] We thank Thomas Lubik for providing the UK estimates.

Table 1:

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

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

We know therefore that one can estimate α, σ_{π^*} and ρ_{π^*} from this equation without reference to the remainder of the system.

The exact MLE estimates of the parameters of (10) are found using MicroFit Version 5 and are presented in Table 3. We will call this the LIML estimator since it is a single equation estimator.

Now it is clear that these are very different to the Full Information Bayesian estimates in Table 1. A negative value for α is certainly unattractive since it is meant to be an import share, but the implication of the MLE estimates is more that one can't estimate it with any precision. In times past a "wrong sign" might well have suggested to an investigator that there are specification problems with the equation. To look at what causes the difference between the Full Information Bayesian estimates and the LIML we stand back from the system and simply estimate the exchange rate equation

Table 2:

	Original model				Model with hybrid IS curve			
	$\tilde{y}$		$\bar{y}$		$\tilde{y}$		$\bar{y}$	
c*	-	-	-	-	0.92	[0.84, 1]	0.53	[0.46, 0.58]
ψ_1	1.16	[0.97, 1.33]	1.35	[1.09, 1.59]	1.19	[0.95, 1.40]	1.20	[0.91, 1.45]
ψ_2	0.4	[0.18, 0.61]	0.34	[0.15, 0.52]	0.51	[0.16, 0.76]	0.12	[0.05, 0.19]
ψ_3	0.12	[0.07, 0.17]	0.06	[0.01, 0.10]	0.13	[0.08, 0.19]	0.11	[0.05, 0.16]
ρ_R	0.69	[0.63, 0.77]	0.73	[0.66, 0.80]	0.70	[0.60, 0.78]	0.77	[0.71, 0.84]
α	0.12	[0.07, 0.18]	0.05	[0.02, 0.07]	0.13	[0.07, 0.17]	0.07	[0.04, 0.10]
κ	1.97	[1.18, 2.67]	0.72	[0.42, 0.99]	1.86	[1.05, 259]	0.32	[0.22, 0.44]
τ	0.44	[0.6, 0.64]	0.12	[0.07, 0.15]	0.39	[0.25, 0.52]	0.12	[0.07, 0.16]
ρ_q	-0.17	[-0.29, -0.05]	-0.18	[-0.28, -0.06]	-0.19	[-0.31, -0.07]	-0.19	[-0.31, -0.06]
ρ_A	0.18	[0.11, 0.24]	0.24	[0.14, 0.32]	0.21	[0.12, 0.28]	0.17	[0.10, 0.23]
ρ_{y^*}	0.97	[0.96, 0.99]	0.97	[0.95, 0.98]	0.98	[0.95, 0.99]	0.84	[0.77, 0.89]
ρ_{π^*}	0.39	[0.29, 0.50]	0.29	[0.16, 0.43]	0.37	[0.25, 0.47]	0.30	[0.20, 0.40]
σ_R	0.33	[0.26, 0.39]	0.32	[0.26, 0.37]	0.34	[0.26, 0.42]	0.25	[0.21, 0.28]
σ_g	1.33	[1.17, 1.50]	1.33	[1.15, 1.49]	1.34	[1.16, 1.50]	1.35	[1.19, 1.56]
σ_A	0.57	[0.48, 0.65]	1.12	[0.33, 2.07]	0.58	[0.49, 0.67]	0.63	[0.26, 0.93]
σ_{y^*}	0.83	[0.32, 1.49]	0.74	[0.33, 1.27]	0.58	[0.35, 0.83]	0.59	[0.35, 0.81]
σ_{π^*}	3.31	[2.90, 3.73]	3.22	[2.80, 3.57]	3.25	[2.84, 3.61]	3.24	[2.82, 3.65]
$\sigma_{\epsilon^{bn}}$	-	-	0.67	[0.57, 0.75]	-	-	0.88	[0.71, 1.05]

Note: * We take the prior for the habit persistence parameter c ∼beta(0.2,0.3).

using Limited Information Bayesian methods. Figure 1 then shows that there is a large difference between the posteriors of the Full Information (system) (FIBE) and Limited Information (single equation) (SIBE) Bayesian estimators. The imprecision that is indicated by the MLE is present in the SIBE results. Moreover the fact that the mean of the posterior for SIBE is virtually the same as the prior reveals that there is very little information in the sample about α. This point is made very clear when one looks at Figure 2, which shows how the likelihood changes with α, and what the criterion being used to get the mode of the posterior is (the criterion being the sum of the log likelihood plus the log of the prior). It's evident that this criterion

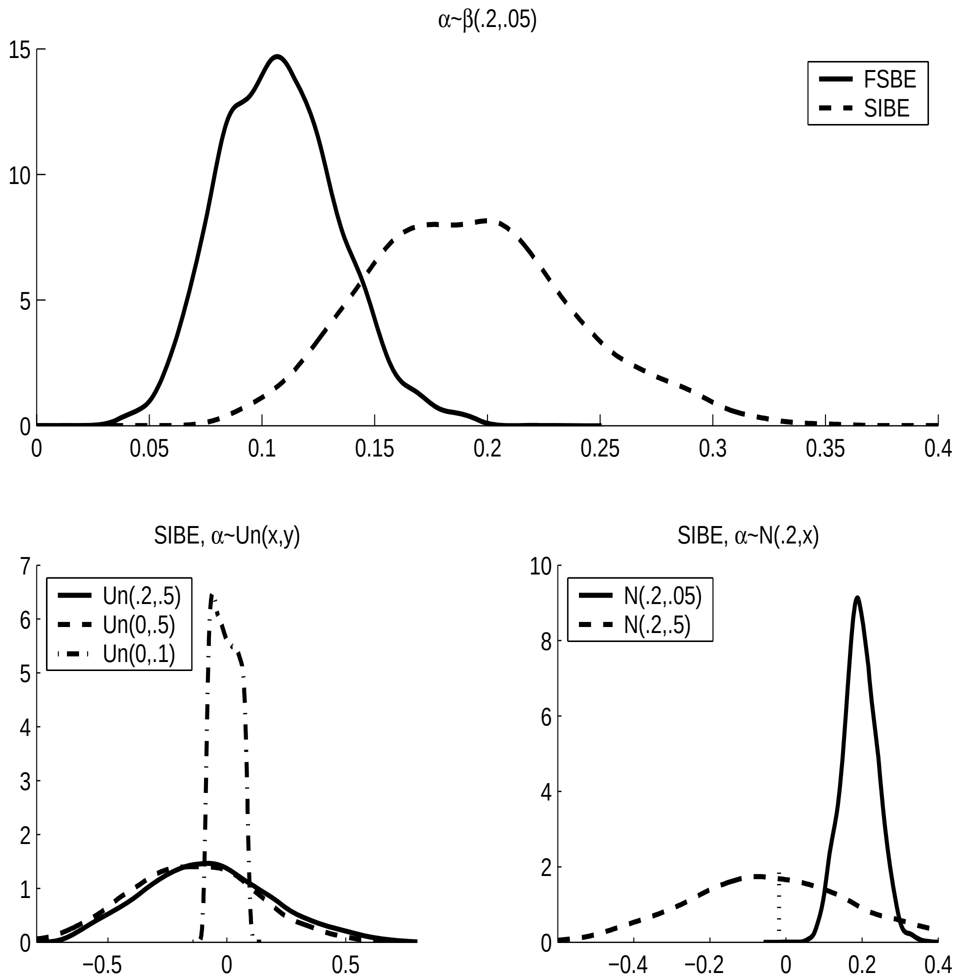
Table 3: LIML Estimates of the Parameters of (10)

	est	std dev	t	
α	-.113	.26	.43	
ρ_{π^*}	.073	.11	.65	
σ_{π^*}	3.195	.36		

is dominated by the prior component. Given this the Limited Information modal estimates reflect the location of the prior mean. If it is centered on zero one will get results very similar to the MLE. To show this we impose a normal prior with a zero mean and the same standard deviation as for the beta density. The results are in the bottom part of Figure 1.

Although there is nothing surprising in these findings, the point is that the Bayesian *systems* estimates *suggest the opposite* i.e. that there is a good deal of information in the sample. The implied standard deviation of the posterior for α is .034 which would produce a t ratio of around 4, as contrasted with the .43 of the single equation MLE. Now, given the exogeneity assumptions in place, the extra information is clearly not from the data on the real exchange rate but must arise from either the imposition of cross-equation restrictions due to the presence in the system of forward expectations or the assumption that the shocks in the LS model are uncorrelated. If these assumptions are incorrect then there will be a bias introduced into the posterior of α. It seems a big assumption to believe that these will be accurate and we will see later that there is strong evidence against them. One wonders at the wisdom of using the complete system to estimate parameters that can be estimated without reference to it.

The same situation applies to ρ_{π^*}. In this case however the divergence does not come from the use of a Full Information estimator. Figure 3 shows that the posterior we get is virtually the same for the Full and Limited Information estimators. What accounts for the difference now is the nature of the prior and its location. If we choose a uniform density that has the same mean and variance as the beta density we basically get the MLE. Moreover, if we choose a normal prior with the same standard deviation as the beta prior, and also one that is five times higher, than we would again get the MLE evidence. Thus the parameterization of a given prior is now the principal determinant



α~β(.2,.05)
FSBE
SIBE
0
5
10
15
0
0.05
0.1
0.15
0.2
0.25
0.3
0.35
0.4
SIBE, α~Un(x,y)
Un(.2,.5)
Un(0,.5)
Un(0,.1)
−0.5
0
0.5
SIBE, α~N(.2,x)
N(.2,.05)
N(.2,.5)
−0.4
−0.2
0
0.2
0.4

Figure 1:

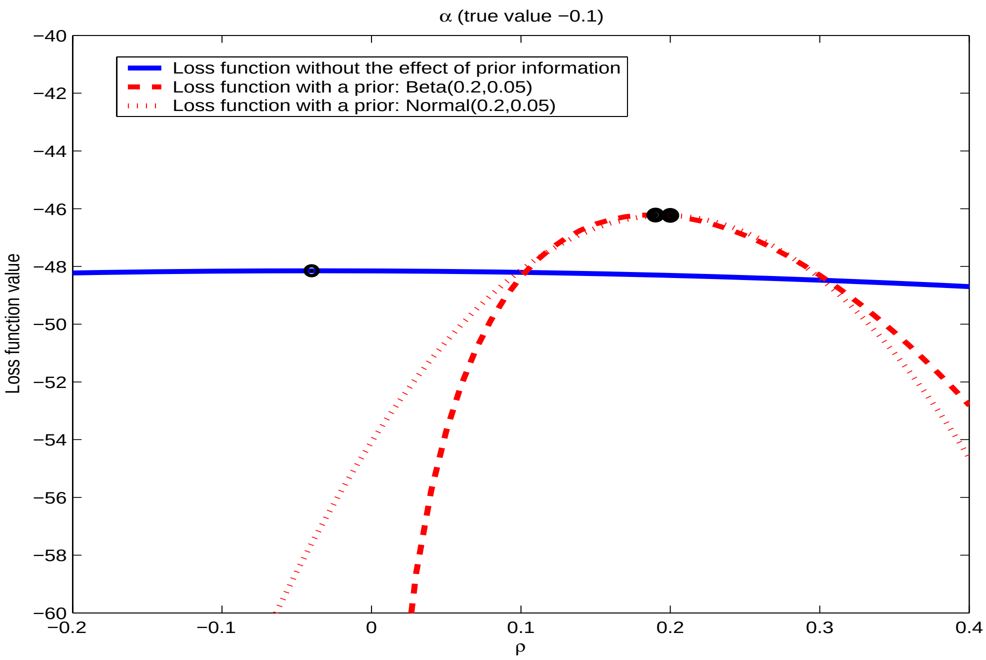

α (true value −0.1)
Loss function without the effect of prior information
Loss function with a prior: Beta(0.2,0.05)
Loss function with a prior: Normal(0.2,0.05)
Loss function value
ρ
−40
−42
−44
−46
−48
−50
−52
−54
−56
−58
−60
−0.2
−0.1
0
0.1
0.2
0.3
0.4

Figure 2:

of the modal estimates. The uncertainty, as judged by the posterior standard deviation of .011, is the same as for the LIML estimator, so that it is the location of the posterior that has been affected, and this is important if one wished to decide if ρ_{π^*} was zero. The data only has a limited influence on the posterior location due to the strength of the prior. Exactly where one gets the prior knowledge that the autoregressive parameter has a very limited range of values is unknown. In this simple case one can discover these difficulties but, in more complex applications, it seems unlikely. Thus one has to be very wary of where the information supplied with systems Bayesian estimators come from, and studying the limited information estimators as a complement seems imperative.

5.2.2 The IS Curve

The IS equation has the form

$$\begin{aligned}\overline{y}_t &= E_t\overline{y}_{t+1} + \eta_1(R_t - E_t\pi_{t+1}) + \eta_2\Delta q_t + \tilde{y}_t^* \\ \tilde{y}_t^* &= \eta_3\tilde{y}_{t-1}^* + \eta_4 e_t. \end{aligned} \tag{22}$$

Now we need to decide which of the model assumptions we want to retain. We have already done this implicitly by making the coefficient on $E_t(\overline{y}_{t+1})$ equal to one. Later we will review that choice. This means that there are 6 DSGE parameters that appear in (22), $\tau, \alpha, \rho_{\tilde{y}^*}$, $\sigma_{\tilde{y}^*}, \rho_q$ and σ_q, so that it is impossible to estimate all six from this equation with limited information methods. However, if an equation for the terms of trade, (12), is adjoined then we will have a small system that incorporates the six parameters. To estimate the parameters we need to complete the system with equations for π_t and R_t. We do that by adjoining unrestricted versions of VARX equations in (21).

A second decision that is required is how we are to estimate this new synthetic system. In order to retain comparability with LS's values we decided to use Bayesian methods i.e. we used the same priors as they did for τ,α, $\rho_{\tilde{y}^*}$ and $\sigma_{\tilde{y}^*}$, while the priors for the Δq_t equation were those given in Table 1 earlier. The VARX parameters were set to the OLS values from fitting (21).[6] Since Δe_{t-1} appears in the VARX equations for π_t and R_t an equation to generate it was needed. We decided to use (10) but with $(1-\alpha)$ replaced

[6] The alternative was to jointly estimate these with the IS equation. Mostly this gives similar results but the computational demands are much higher.

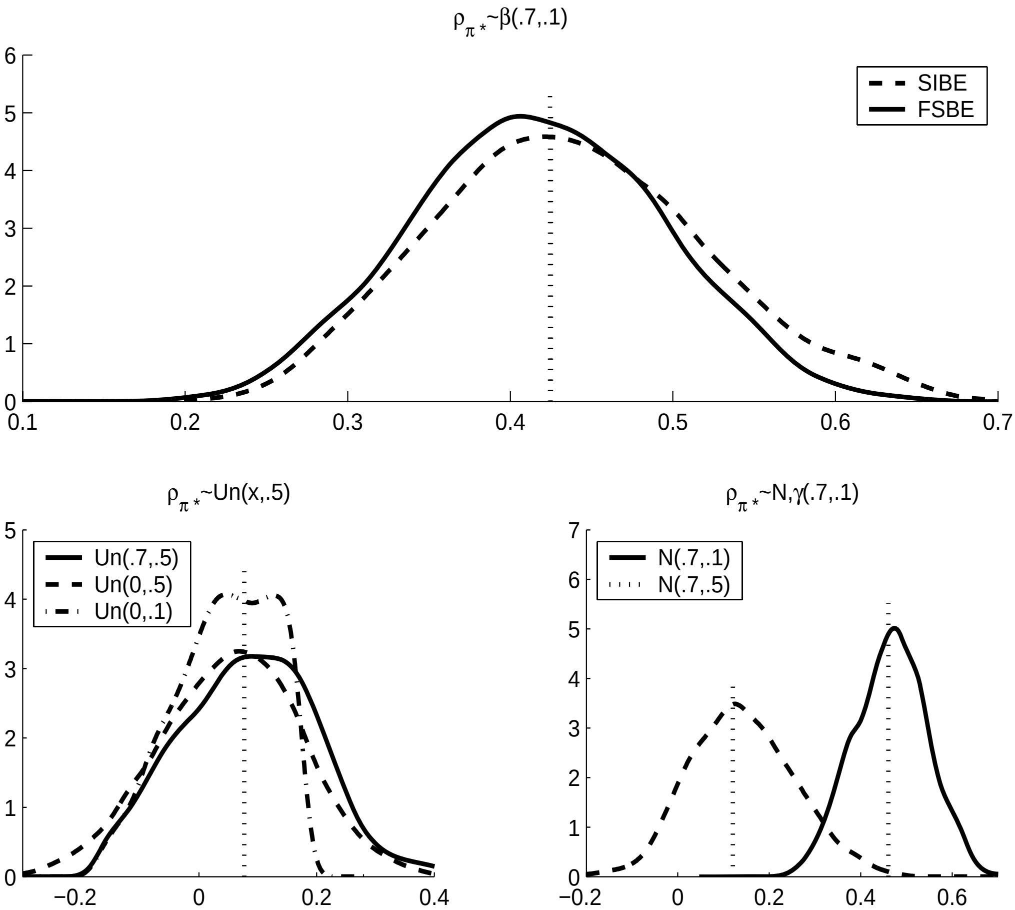


Figure 3:

by c_q, so that the information in the exchange rate equation is not used for estimating α except insofar as it affects expectations. An alternative would be to use a VARX equation for Δe_t but, in fact, the main determinants were ΔR_t and Δq_t. DYNARE was then used to estimate this system. Thus the solutions for $E_t(y_{t+1})$ and $E_t(\pi_{t+1})$ are consistent with the synthetic system. In order to generate $\overline{y}_t$ for use in the system we need to compute y_t^p, and this was done by applying the Beveridge-Nelson decomposition to data on y_t. An assumption was made that Δy_t was an AR(2), although the answers are not sensitive to making it an AR(1) or higher order.

Finally, if we accept that there are likely to be specification problems with the equation, what alternative forms suggest themselves? The most likely mis-specification is that the equation is in fact a hybrid one i.e. it has the form (19). The limited information (2SLS type) estimates of such an equation were found and Table 4 presents these. For comparative purposes Table 3 contains full information estimates of the hybrid curve using both the $\tilde{y}_t$ and $\overline{y}_t$ forms. It is interesting that there is strong evidence of a hybrid form once we work with $\overline{y}_t$. Moreover, the coefficient of $E_t\overline{y}_{t+1}$ is well away from zero in the latter form, unlike the case with $\tilde{y}_t$, which may be why LS did not use such a form.

Examining the limited information results it is clear that the estimates of α and τ become very imprecise when the mis-specification is corrected, and $\tilde{y}_t^*$ now seems to have unit root behaviour. This may be indicative of the fact that the equation is still mis-specified. Returning to the non-hybrid model it is interesting to observe that α and τ are estimated about as precisely with limited information as with full information. This could be because the IS equation is the dominant source of information about them. An alternative however is that there is little information in the data about these parameters and so the results we are getting just reflect the prior. To assess this we computed the modes of these parameters for the non-hybrid model under two scenarions - when the priors are those used by LS and when they are normal with the same mean but with a large standard deviation. The LS-modes were (.07,.08) and the normal-priors gave modes of (.003,.007). Hence the prior seems to the dominant influence on the estimation of the coefficients and the data embodies in the IS curve has little role.

Table 4: LIBE (2SLS) of the IS Curve

	Non-hybrid IS curve		Hybrid IS curve	
α	0.11	[0.05 0.21]	0.02	[0.01 0.02]
τ	0.12	[0.06 0.30]	0.01	[0.00 0.02]
ρ_{y^*}	0.91	[0.86 0.93]	0.99	[0.98 0.99]
ρ_{π^*}	0.48	[0.32 0.63]	0.36	[0.26 0.48]
c	-	-	0.68	[0.66 0.69]
σ_{y^*}	0.79	[0.41 1.11]	0.79	[0.44 1.17]
σ_{π^*}	3.39	[2.95 3.48]	3.51	[3.07 4.07]

5.2.3 The Phillips Curve

Limited Information estimates of the Phillips curve in (17) were found by using the same strategy as for the IS curve. The DSGE model parameters present in the equation are α,τ, κ, ρ_{y^*}, σ_{y^*}, ρ_g, σ_q, ρ_A. Therefore one parameter must be fixed and ρ_A was set to .2. Then equations for $\Delta q_t, \Delta e_t$ were adjoined along with VARX equations for $\overline{y}_t$ and R_t. *Prima facie* it looks as if there is useful information in the equation for α, τ although there is clearly very little about κ (and this is the only equation κ appears in). Indeed, combining the results from Table 2 about κ with these, it seems as if there is only weak information in the sample and that the strength of the evidence in Table 1 relies a great deal upon the normalization with respect to an unobservable. As this parameter determines the effect of variations in demand upon inflation, it is a crucial parameter in macroeconomic models, so that one would need to be very cautious in adopting it.[7] Regarding α, τ we again allow the priors on α, τ to have the same mean as the LS versions but with a substantial standard deviation. The modes are (.16,.22) and (.12,.14) respectively, so it does appear as if the Phillips curve contains some evidence on these parameters, and this is the source of the systems results. A check on the

[7] It is noticeable that there are differences between the estimates of κ found using the $\tilde{y}_t$ and $\overline{y}_t$ systems. One reason for this is that we are using the assumption in the DSGE model that Δa_t is uncorrelated with the other shocks. If true this means that Δy_t^p is uncorrelated with them. Later we will see that this may well be incorrect and that would produce biases in the coefficient estimates of both formulations of the system, possibly in different directions.

specification once again involved estimating a hybrid model. This was done by writing $\beta E_t \pi_{t+1}$ as $\beta c E_t \pi_{t+1} + (1-c)\beta\pi_{t-1}$ and estimating c. As Table 5 shows there is clear evidence of this. Interestingly, unlike the IS equation, the α and τ parameters are more precisely estimated than with full information estimation which suggests that some of the mis-specifications that can happen in the complete system have had an impact upon the estimation of α and τ.

Table 5: LIBE (2SLS) of the Phillips Curve

	Non-hybrid PC curve		Hybrid PC curve	
α	0.17	[0.11 0.23]	0.17	[0.12 0.22]
τ	0.22	[0.03 0.39]	0.24	[0.05 0.48]
κ	0.11	[0.03 0.19]	0.11	[0.04 0.19]
ρ_{y^*}	0.37	[0.19 0.57]	0.08	[0.00 0.19]
ρ_{π^*}	0.23	[0.07 0.38]	0.24	[0.07 0.61]
c	-	-	0.70	[0.59 0.79]
σ_{y^*}	2.92	[0.69 5.09]	3.72	[0.74 7.96]
σ_{π^*}	3.26	[2.88 3.72]	3.26	[3.82 3.74]

5.3 Some Residual and Tracking Diagnostics

A different perspective is to be had by using LS's parameter estimates to compute $\overline{y}_t$, $E_t(\overline{y}_{t+1})$ and $E_t(\pi_{t+1})$. Once these quantities have been formed $\tilde{y}_t^*$ can be found from (16) and (17) and the interest rate shock ε_t^R from (18). Table 6 fits an AR(2) to these series. According to the LS full information estimates the serial correlation in $\tilde{y}_t^*$ should be .97 and the interest rate shock should have none. Clearly the estimated model is inconsistent with these predictions. Moreover, whilst the correlation between the two estimates of $\tilde{y}_t^*$ should be unity, it is actually .68. Worse is the fact that the correlation between the interest rate shock and $\tilde{y}_t^*$ is either -.35 (using the IS estimate of $\tilde{y}_t^*$) or -.7 (using the Phillips curve estimate) rather than zero, which was the basis of the Bayesian estimates. It is also the case that Δy_t^p from the model is correlated with $\tilde{y}_t^*$ with a correlation of .3, while Δy_t^p using the Beveridge Nelson decomposition is correlated with its estimate of $\tilde{y}_t^*$ from the Phillips curve of .70 (non-hybrid IS model). This almost certainly accounts for the differences in parameter estimates in Table 2 found with these different measures of the permanent component. Neither should really be trusted as all estimates are based on the assumption that Δa_t is uncorrelated with $\tilde{y}_t^*$. The latter seems to be an odd assumption to make as it implies that the transitory component of world technology developments are uncorrelated with domestic ones.

Table 6: AR(2) Models Fitted to Euler Equation Residuals

Eq		AR(1)	t		AR(2)	t
y_t		.32	3.0		.37	3.5
π_t		.17	1.6		.50	5.0
R_t		-.33	-2.8		.17	1.5

A final diagnostic that is useful is a comparison of the expectations of inflation provided by the estimated LS model with the actual inflation rate. Fig 4 presents this. It is clear that, although expectations track the downward movements to inflation that occurred over the 1990s, it overdoes this, with deflation being a feature, particularly towards the end of the sample.

6 Conclusion

We have advanced the proposal that DSGE models should not just be estimated and evaluated with reference to full information methods. These make strong assumptions and therefore there is uncertainty about their impact upon results. Some limited information analysis which can be used in a complementary way seems important. Because it is sometimes difficult to implement limited information methods when there are unobservable non-stationary variables in the system we present a simple method of overcoming this that involves normalizing the non-stationary variables with their permanent components and then estimating the estimating the resulting Euler equations. We illustrate the interaction between full and limited information methods in the context of a well-known open economy model of Lubik and Schorfheide. The transformation to observable variables that we employ

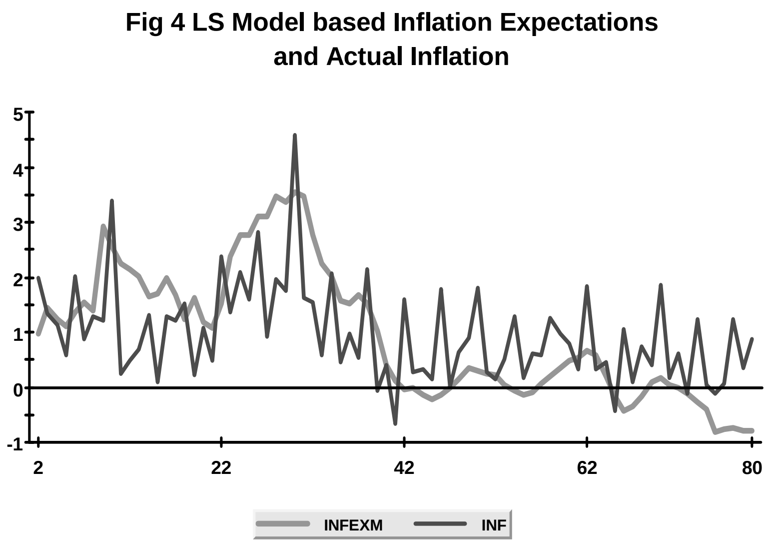


Figure 4:

was effective in revealing possible mis-specifications in the equations of LS's system and the limited information analysis highlighted the role of priors in having a major influence upon the estimates. The major influence of priors seems to be one that needs attention, particularly if policy actions are being recommended that depend upon parameter estimates that are largely produced by the type of prior density used rather than the data.

7 References

Binder, M and M.H.. Pesaran (1995), "Multivariate Rational Expectations Models and Macroeconomic Modelling: A Review and Some New Results" in M.H. Pesaran and M Wickens (eds) *Handbook of Applied Econometrics: Macroeconomics,* Basil Blackwell, Oxford.

Fukac, M.and A.R.Pagan (2006) "Issues in Adopting DSGE Models for Use in the Policy Process", *CAMA Working Paper 10/2006*, Australian National University

Harvey, A.C. and A. Jaeger, (1993), "De-Trending, Stylized Facts and the Business Ccyle", *Journal of Applied Econometrics*, 8, 231-247

Lubik, T.A. and F. Schorfheide (2005) " Do Central Banks respond to exchange Rate Movements? A Structural Investigation", *Journal of Monetary Economics,* (forthcoming)

Pagan, A. R. “Some Consequences of Viewing LIML as an Iterated Aitken Estimator”, *Economics Letters* (1979), 369–372.

Singleton, K.J, (1988) "Econometric Issues in the Analysis of Equilibrium Business Cycle Models," *Journal of Monetary Economics* 21, 361-386.