Limited Information Estimation and Evaluation of DSGE Models *

Martin Fukac[†] and Adrian Pagan[‡]

January 29, 2008

Abstract

We advance the proposition that DSGE models should not just be estimated and evaluated with full information methods. These require that the complete system of equations are specified properly. Some limited information analysis which focuses upon specific equations is therefore a useful complement to full system analysis. Two major problems occur when implementing limited information methods. These are the presence in the system of forward-looking expectations and unobservable non-stationary variables. We present methods for dealing with both of these difficulties and illustrate the interaction between full and limited information methods in the context of the Phillips curve in a model set out in Lubik and Schforheide(2007).

Contents

*Research of the second author was supported by ARC Grant No. DP0449659
[†]Reserve Bank of New Zealand
[‡]Queensland University of Technology and University of New South Wales

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] Reasons often cited for this shift involve a possible improvement in estimator performance, an argument which however needs more detailed analysis than we can provide here. 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 it rarely involves an examination of the individual structural equations of the DSGE model i.e. of the Euler and other structural 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, it was only when one looked at their performance in a system that deficiencies became apparent. Therefore evaluation tools were needed that treated the system as a whole. These were viewed as a complement rather than as'a substitute to single equation methods. For this reason it seems useful to examine the structural equations of any DSGE model in order to determine the extent to which they fit the data as a *supplement* to any systems-based 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 might suggest suitable re-specifications.

In the standard simultaneous equations context limited information methods were implemented by augmenting the structural equation to be estimated with the reduced form equations corresponding to the "right hand" endogenous variables in that equation. 2SLS estimated the reduced form equations by OLS while LIML jointly estimated them along with the structural equa-

[1] We owe this quote to the late Chris Higgins, Secretary of the Australian Treasury.

[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, where many studies exist of these individual equations. Even there, however, although estimation of a complete system has sometimes been performed by single equation methods, this has not been true of evaluation.

tion parameters. It is natural then to look to apply limited information methods to the structural equations of DSGE models in a similar way. Two difficulties arise in performing such an extension. One is the presence of forward-looking expectations in the Euler equation. One might replace these expectations with future values of the variable and then proceed to apply 2SLS. This was the initial solution to handle such expectations, McCallum (1976), but the properties of the estimator have been disappointing, due to the problem of weak instruments. Subsequently, suggestions have been made that one should solve for the expectations from the system composed of the Euler equation and the equations completing the system e.g. Fuhrer and Olivei (2004). Kurmann (2007) has shown that there are certain difficulties with such a solution and it is even possible that one might even get inconsistent estimators of the parameters.

The second problem is that many DSGE models feature variables that are integrated processes and such variables have to be transformed to stationarity before the standard full information methods of estimation are applied. A common way to do this involves working with transformed variables that are no longer observable and so it is necessary to perform estimation when there is a difference between the total number of endogenous variables and the number that are observable. This difficulty has been resolved in the systems literature by first expressing the DSGE model in a state space form and then constructing the system likelihood by utilizing the Kalman filter, so one needs to adapt this solution to a limited information context.

Section 2 of the paper briefly states the connection between complete structural (DSGE) models and the VAR formed from the variables appearing in it. We will be extensively using both of these constructs in what follows. In section 3 of the paper it is assumed that the variables in the Euler equation and the augmenting system are "gap" variables i.e. are deviations from equilibrium values, and that these gap variables are measurable. We describe a LIML estimator that can be implemented and which avoids the difficulties identified by Kurmann. It utilizes a well-known computer program - DYNARE- to perform the estimation. Section 4 then turns to the situation when the structural equations are expressed in terms of gap variables but where these are constructed as deviations of observed variables from $I(1)$ latent variables. The solution presented in section 3 is then generalized to handle this case. In this section we also discuss what happens if the extraction of the permanent components is done in a way that is inconsistent with that implied by the system composed of the Euler equation and the augmenting system.

Section 5 looks at the estimation of the Phillips curves in Lubik and Schorfheide's (2004) New Keynesian model and suggests some methods for determining the goodness of fit of the model that are not just system methods i.e. which do not just ask if the model can match a VAR. Section 6 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 deviations from some equilibrium values, either in levels or logs. 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 generally consists of a set of Euler equations describing optimal choices, other structural equatioins and possibly a set of identities. The parameters in the DSGE model will be designated as θ.

The solution to this system has the form- see Binder and Pesaran (1995)

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

where P satisfies $B_0 P - B_1 - CP^2 = 0, \Pi_1 = (B_0 - CP)^{-1} C, \Pi_2 = (B_0 - CP)^{-1} D$ 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

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

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 $E_t z_{t+1}$

$$E_t(z_{t+1}) = P z_t + \overline{D} \Phi_x x_t + \overline{G} \Phi_u u_t. \tag{3}$$

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 the expectation can be constructed 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). The FIML estimator maximizes the likelihood of (2) after substituting (3).

3 Limited Information Estimation and Evaluation of Gaps Models

3.1 Estimation

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

$$z_{1t} = B_{10}z_t + B_{11}z_{t-1} + D_1 x_t + C_1 E_t z_{t+1} + \zeta_{1t}, \tag{4}$$

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, we will assume here 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 $E_t z_{t+1}$.

Now for simplicity we will consider a restricted version of (4) in (5)

$$z_{1t} = B_{10}z_t + B_{11}z_{t-1} + C_1 E_t z_{t+1} + \zeta_{1t}, \tag{5}$$

where the exogenous variable appearing in (4) is ignored i.e. $D_1 = 0$. In this case the solution to the complete DSGE model will be a VAR of the form (6).

$$z_t = P z_{t-1} + \overline{G} u_t, \tag{6}$$

As well as these equations we can also express the sub-set of (6) that deletes the equation corresponding to z_{1t} as

$$z_t^- = P^- z_{t-1} + \overline{G}^- u_t. \tag{7}$$

where $z_t = \begin{bmatrix} z_{1t} \\ z_t^- \end{bmatrix}$. In standard simultaneous equation methods the presence of right- hand endogenous variables in (5) was dealt with by constructing a synthetic system composed of the simultaneous equation whose parameters are being estimated and auxiliary equations which determine the right-hand side endogenous variables i.e. the equivalent of (5) and (7). 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. One advantage of constructing estimates in this way was that the same software could 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 (subject of course to their being identified from the single Euler equation).

Now there are some difficulties that can arise when seeking to perform limited information estimation of (5) due to the presence of the forward looking expectation. Initially, forward looking expectations were handled by McCallum's (1976) proposal that $E_t z_{t+1}$ be replaced by z_{t+1}, after which instrumental variables were applied to (5). Instruments were chosen from z_{t-1} (or a sub-set of it) . Most frequently one uses linear combinations of z_{t-1}. One possibility is to regress z_{t+1} against z_{t-1} and use the predictions from this as an instrument for z_{t+1}. Effectively, one is just treating z_{t+1} as another endogenous variable, and this leads one to look at how instruments were generated in the standard simultaneous equations literature, since a conditional expectation was also being used there as an instrument. There have been a number of suggestions, apart from the 2SLS solution. One proposal by Brundy and Jorgenson (1971) was to estimate the conditional expectation from the derived model reduced form information. Fuhrer and Olivei (2004) make the same suggestion in order to produce an instrument for z_{t+1}. They implement it by solving the system composed of (5) and (7), but with P^- replaced by $\hat{P}^-$, the OLS estimate of the regression of z_t^- upon z_{t-1}.

$\hat{P}^-$ can be used to compute $E_t z_{t+1}$ since this is a function of P^-, B_{10}, B_{11} and C_1. The procedure is then iterated as the estimates of B_{10}, B_{11}, C_1 are changed ($\hat{P}^-$ is held constant however). Asymptotically there can be no difference between the estimator that just uses $\hat{P}$ (the OLS estimator of the regression of z_t on z_{t-1}) and the one proposed by Fuhrer and Olivei, but their simulations suggest that there are some small sample advantages.

A different approach used by Sbordone (2004) is to recognize that P is a function of $-$ $g(B_{10}, B_{11}, C_1)-$ and to find estimates of these parameters $(\tilde{B}_{10}, \tilde{B}_{11}, \tilde{C}_1)$ by minimizing the difference between $\hat{P}$ (the OLS estimator) and $\tilde{P} = g(\tilde{B}_{10}, \tilde{B}_{11}, \tilde{C}_1)$. Basically this yields an indirect estimator of B_{10}, B_{11}, C_1, where the model equations are taken to be (5) and (7) while the auxiliary model is (6). The theory of such estimates is set out in Gourieroux et al. (1993) while Smith (1993) independently developed the method for estimating the parameters of a complete DSGE model by using the VAR as an auxiliary model.

Kurmann (2007) pointed out that there was a possible problem with these approaches when there are forward-looking expectations and only a sub-set of the structural equations is estimated. The problem arises in that, even when the complete system (1) has a unique solution, if one constructs the pseudo-system from only one of the Euler equations of the system and an augmenting VAR for z_t^- i.e. (5) and (7), this may not have a unique solution.[3] The difference derives from the fact that $E_t(z_{t+1})$ coming from the complete system is not the same as that from the augmented system. The example he provides involves a two equation system with endogenous variables z_{1t} and z_{2t}, where the endogenous variable that is the dependent variable in the structural equation being estimated, z_{1t}, is involved contemporaneously in the second structural equation. Recognition of this dependence turns out to be crucial for uniqueness. It may be lost if one simply uses the VAR equations for z_t^- i.e. z_{2t}, to construct an estimate of $E(z_{1t})$. Forcing uniqueness on the solution for z_{1t} may even rule out the correct parameter values in the Euler equation.

Now this difficulty can have an impact upon maximum likelihood estimation of (5) if we just use the augmented system. To see why we observe that the correct expected value of z_{t+1} is $E_t(z_{t+1}) = Pz_t$, while the expectation from the augmented system would be $G^* z_t$. Hence if one replaced $E_t(z_{t+1})$

[3] We will refer to these two systems in what follows as the complete and augmented models respectively.

by $G^* z_t$ the simultaneous equation to be estimated would become

$$z_{1t} = B_{10} z_t + B_{11} z_{t-1} + C_1 G^* z_t + \zeta_{1t} + C_1 (P - G^*) z_t. \tag{8}$$

The augmented-model LIML estimator would then estimate (8) along with (7). Now LIML can be shown to be an IV estimator which uses $G^* z_{t-1}$ as an instrument for z_t. This instrument is uncorrelated with ζ_{1t}, but, since the error in (8) has an extra term $C_1(P - G^*)z_t$, the expectation of this with $G^* z_{t-1}$ will not be zero. It is generally necessary that $P = G^*$ for the term to disappear (technically since G^* will be an estimator we need that $G^* \xrightarrow{p} P$). Consequently, it is important that one solve for the correct expectation of z_{t+1}. Notice that the situation is different if an IV estimator is used that simply replaces $E_t z_{t+1}$ with z_{t+1} and then instruments this with $G^* z_{t-1}$. This will produce consistent estimators of the parameters of (5) since $G^* z_{t-1}$ is uncorrelated with $\zeta_{1t} + C_1(E_t(z_{t+1}) - z_{t+1})$.

Kurmann provides a "reverse engineering" solution to the problem just identified but there is a simpler one that is easier to apply with existing computer code. One simply needs to estimate $E_t(z_{t+1})$ from the complete rather than augmented model. But this is just a matter (at least in large samples) of using the (unrestricted) VAR equations in (6) to get this expectation. Basically this VAR incorporates (in large samples) the nature of the remaining structural equations which are being ignored in solving (5) and (7). Using the expectation computed in this way in Kurmann's example produces exactly the same solution as in the complete system. Thus what the correct LIML estimator does is to estimate

$$z_{1t} = B_{10} z_t + B_{11} z_{t-1} + C_1 P z_t + \zeta_{1t} \tag{9}$$

and (7) by maximum likelihood. As before P could be estimated jointly with the structural parameters or one could replace P with $\hat{P}$ (the OLS estimate). The first of these is the analogue of LIML while the second is a 2SLS-like estimator. Notice that the errors of (9) and (7) must be allowed to be correlated as the shocks in the latter are functions of the shocks in the former.

3.2 Evaluation

To evaluate the DSGE model we propose examining the individual structural equations rather than the complete system. Such tests generally involve one of three approaches.

1. Extending the maintained equation in some direction and seeing if the extended equation dominates the maintained one.

2. Testing internal consistency of the model i.e. whether the assumptions made in constructing the model are compatible with the data.

3. Parametric encompassing tests.

Each of these can also be done with the FIML estimates of a complete system and it makes sense to perform the tests using both limited and full information. We discuss each of these testing procedures in turn.

3.2.1 Model Extension

Extending the structural equation is context dependent. Often however DSGE models only have forward-looking expectations in them and so one is interested in checking if this assumption is too strong. The class of hybrid models in which the expectations and dynamics coefficients sum to some predetermined value is a natural extension to be checked and we utilize this in the example below.

3.2.2 Internal Consistency

Internal consistency in DSGE models largely pertains to the assumptions made about the shocks. These are generally that the shocks are autoregressive processes of a particular order and that they are uncorrelated. Mostly normality is assumed as well. One therefore might be interested in checking the validity of these assumptions. Generally, DSGE models are nominally over-identified, and so one can perform such checks.

3.2.3 Parametric Encompassing

Within the literature which currently emphasizes complete model evaluation it has been parametric encompassing that has been the standard method of evaluation. Basically the implied estimated VAR parameters in (2),$\tilde{P}$, are compared to the OLS estimates $\hat{P}$, and formal statistical tests are applied.[4] This approach goes back a long way. It was used in Canova et al (1993)

[4] Again we are ignoring the fact that the solution may be a VARX rather than VAR system i.e. we set $\overline{D} = 0$.

for example. Modern versions of it try to map the differences between $\tilde{P}$ and $\hat{P}$ into a scalar λ that is meant to measure how large this deviation is. There is clearly a variant of it that would use the estimate of P that comes from solving the system composed of (9) and (7) and comparing this to $\hat{P}$. This would be a limited information encompassing test as it only imposes the restrictions from the structural equation upon the VAR.

4 Permanent Components in the Series

If the gap data is supplied independently of the model then we can proceed as above. However, if the gaps need to be consistent with the model, and this involves extracting a permanent component for the observed variables, then it is necessary to ask how this affects the estimation methods discussed above.

4.1 The Effects of Filtering

We begin by looking at the issue of estimating a single structural equation which is expressed in terms of the levels of an $I(1)$ variable z_{1t}. This could for example be an output equation and will have the form of (4). For simplicity we set $B_{10} = 0$. If this equation is transformed so as to only incorporate $I(0)$ "gap" variables which are the deviations between z_t and their permanent components (z_t^p), we will have

$$\begin{aligned} \tilde{z}_{1t} &= B_{11}\tilde{z}_{t-1} + C_1 E_{t+1}\tilde{z}_{t+1} + Dx_t + C_1 E_t(\Delta z_{1t+1}^p) \\ &\quad - B_{11}\Delta z_t^p + (C_1 + B_{11}) z_t^p - z_{1t}^p + \zeta_{1t}, \end{aligned} \tag{10}$$

where $\tilde{z}_{1t} = z_{1t} - z_{1t}^p$. There are some issues that arise in relation to this equation. Firstly, unless $(C_1 + B_{11})z_t^p - z_{1t}^p$ is $I(0)$, or x_t is co-integrated with z_t^p in a certain way, it is not possible for $\tilde{z}_t$ to be $I(0)$. In many DSGE models z_{1t} is the only $I(1)$ variable and x_t are $I(0)$ variables, such as the inflation rate and the interest rate. So in this case one needs the restriction that the model has a hybrid form i.e. one in which the coefficients of the forward-looking expectations variable and the lagged value sum to a pre-determined value. Secondly it is clear that one needs to account for the fact that the errors in the transformed equations are not the original shocks, and they will depend upon how the permanent component is measured.

If the Beveridge-Nelson (BN) decomposition is used for computing z_t^p then, regardless of whether the 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_t^p has the property that Δz_t^p is white noise, so that $E_t \Delta z_{t+1}^p = 0$. The difference between the multivariate and univariate cases is the conditioning set of information. Since we want to allow for the possibility that agents use a broader information set than that contained in the structural equation we would want to use a multivariate measure. With the BN filter there is a term $B_{11}\Delta z_t^p$ left that will need to be explicitly accounted for in the structural equation that is to be estimated i.e. one cannot ignore it and assume that the errors in the structural equations are the original shocks. Often this is not done.

For other filters however we cannot even be sure that $E_t \Delta z_{t+1}^p = 0$ or that it is uncorrelated with the regressors. This is true of the Hodrick-Prescott (HP) filter. There is a complication in showing this when the HP filter is used to form z_t^p, owing to the fact that the latter is a two-sided filter with time varying weights. There is however a version that has a "steady state" solution of the form

$$z_t^p = \sum_{j=-T}^{T} a_j z_{t-j}.$$

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

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

If one looks at the resulting expression for z_t^p it is clear that, due to the terms $\sum_{k=0}^{T} a_{-j} z_{t-j}$, $E_t(\Delta z_{t+1}^p)$ will never be zero, even if Δz_t is white noise. Ignoring the fact that the term $E_t(\Delta z_{t+1}^p)$ is not zero will generally bias any estimators applied to the structural equations[5]. Simulations show that, in the case of the HP filter, Δz_t^p is a very persistent process that behaves much like one with a unit root, so that the error term in the transformed structural equation will have a great deal of persistence, even though the original shock in it may not have. This is a consequence of Harvey and Jaeger's (1993) and Kaiser and Maravell's (2002) 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^p using the HP

[5] Of course if $B_{11} \neq 0$ one also needs to account for the term -$B_{11}\Delta z_t^p$.

filter ($\lambda = 1600$). The regression of the simulated Δz_t^p against Δz_{t-1}^p gives an estimated coefficient on the latter variable of 1.00.[6] 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.

4.2 Model Consistent Formation of Permanent Components

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. We first need to discuss how this changes the systems we work with. For simplicity we assume that all variables are $I(1)$.

Now let the z_t variables be $I(1)$ and z_t^p be their permanent component (in the case of a stationary variable these would of course be constants). As described above the structural equations will be transformed to a new form in which Δz_t^p is added on to the system.[7] Thus the system of equations will consist of

$$B_0\tilde{z}_t = B_1\tilde{z}_{t-1} + Dx_t + CE_t\tilde{z}_{t+1} + F\Delta z_t^p + Gu_t.$$

We will add to this system some equations describing Δz_t^p of the form $\Delta z_t^p = J\xi_t$, where ξ_t are the white noise shocks driving the common trends of the system. Then, defining $\zeta_t = \begin{bmatrix} \tilde{z}_t \\ \Delta z_t^p \end{bmatrix}$, $\eta_t = \begin{bmatrix} u_t \\ \xi_t \end{bmatrix}$, the system of equations can be solved for ζ_t as in section 2 to give

$$\zeta_t = P\zeta_{t-1} + H\eta_t.$$

Now, as the number of shocks in this system is less than or equal to the number of endogenous variables, $\eta_t = H^{+}(\zeta_t - P\zeta_{t-1})$ and, if η_t is a VAR(1)

[6] Regressing Δz_t^p against Δz_{t-1}^p,Δz_{t-2}^p, Δz_{t-3}^p and Δz_{t-4}^p gives coefficient values of 3.48,-4.57,2.67 and -.58 which do sum to 1. However, this regression shows that the process for Δz_t^p is not a pure random walk.

[7] We would need to impose conditions upon C, B_1 and B_0 in order to ensure that the model involves only $I(0)$ variables. In DSGE models these occur naturally.

of the form $\eta_t = \Phi\eta_{t-1} + \varepsilon_t$, we get

$$\begin{aligned}\zeta_t &= P\zeta_{t-1} + H\Phi H^+(\zeta_{t-1} - P\zeta_{t-2}) + H\varepsilon_t \\ &= (P + H\Phi H^+)\zeta_{t-1} - H\Phi H^+ P\zeta_{t-2} + H\varepsilon_t \\ &= A_1\zeta_{t-1} + A_2\zeta_{t-2} + H\varepsilon_t\end{aligned}$$

and ζ_t therefore follows a VAR(2).[8] Consequently, $E_t(\tilde{z}_{t+1}) = S[A_1\zeta_t + A_2\zeta_{t-1}]$, where S is a selection matrix selecting $\tilde{z}_t$ from ζ_t. Hence A_1 and A_2 need to be estimated. Although ζ_t is not observable we can estimate the latent VAR by using any program that performs MLE with the Kalman filter, using the knowledge that there is a distinction between the observed variables z_t and the latent variables ζ_t.[9] Once A_1, A_2 are estimated these are used to form the forward expectations in the structural equations.

5 An Example

5.1 The Model

The model we analyze is that in Lubik and Schorfheide (2007) (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 $\overline{y}_t = y_t - a_t$, where a_t is the log level of technology. Following section 3 however, we will work with the deviation of y_t from its permanent component and, since a_t is an AR(1), y_t^p equals a_t^p and not a_t.[10]

$$\begin{aligned}\tilde{y}_t &= E_t\tilde{y}_{t+1} - [\tau + \theta](R_t - E_t\pi_{t+1}) - \alpha(\tau + \theta)\rho_q\Delta q_t \\ &\quad -\frac{\theta}{\tau}(1 - \rho_{y^*})\tilde{y}_t^* \end{aligned} \tag{11}$$

In the IS equation q_t is the observable terms of trade, π_t is the domestic inflation rate, R_t is the nominal interest rate, α is the import share, τ is the inter-temporal elasticity of substitution and $\theta = \alpha(2 - \alpha)(1 - \tau)$.

[8] This derivation follows Kapetanios et al. (2007)

[9] In fact the observation equation actually used is in terms of growth rates i.e. $\Delta\tilde{z}_t = \Delta z_t - \Delta z_t^p$.

[10] We also eliminate $E_t(\Delta y_{t+1}^*)$ and $E_t(\Delta q_{t+1})$ from their IS curve by using the latter assumption that these are exogenous AR(1) processes so that $E_t(\Delta y_{t+1}^*) = (\rho_{y^*} - 1)y_{t-1}^*$ and $E_t(\Delta q_{t+1}) = \rho_q\Delta q_t$.

Their open economy Phillips curve is

$$\begin{aligned}\pi_t &= \beta E_t \pi_{t+1} - \alpha(1-\beta\rho_q)\Delta q_t + \frac{\kappa}{(\tau+\theta)}\tilde{y}_t \\ &+ \frac{\kappa\rho_a}{(\tau+\theta)(1-\rho_a)}\Delta a_t + \frac{\kappa\theta}{\tau[\tau+\theta]}\tilde{y}_t^*, \end{aligned} \tag{12}$$

where we have used $y_t - a_t = y_t - y_t^p + a_t^p - a_t$ and $(a_t - a_t^p) = -\frac{\rho_a}{1-\rho_a}\Delta a_t$. In (12) β is the discount factor and κ is a "price stickiness" parameter.

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

$$\begin{aligned}R_t &= \rho_R R_{t-1} + (1-\rho_R)\left[(\psi_1+\psi_3)\pi_t + \psi_2\tilde{y}_t + \psi_3(\Delta e_t - \pi_t)\right] \\ &+ \psi_2\frac{(1-\rho_R)\rho_a}{(1-\rho_a)}\Delta a_t + \varepsilon_t^R \end{aligned} \tag{13}$$

with ε_t^R being serially uncorrelated with standard deviation σ_R.

The exchange rate equation is

$$\Delta e_t - \pi_t = -(1-\alpha)\Delta q_t - \pi_t^*, \tag{14}$$

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

Exogenous variables evolve as AR(1) processes, where the shocks ε_t are all *i.i.d.* with standard deviations of $\sigma_q, \sigma_a, \sigma_{y^*}$ and σ_{π^*} respectively.

$$\Delta q_t = \rho_q \Delta q_{t-1} + \varepsilon_t^q. \tag{15}$$

$$\Delta a_t = \rho_a \Delta a_{t-1} + \varepsilon_t^a \tag{16}$$

$$\tilde{y}_t^* = \rho_{y^*}\tilde{y}_{t-1}^* + \varepsilon_t^{y^*}, \tag{17}$$

$$\pi_t^* = \rho_{\pi^*}\pi_{t-1}^* + \varepsilon_t^{\pi^*}, \tag{18}$$

The solution to the model will have the form of a VARX(2)

$$\tilde{z}_t = \Gamma_1 R_{t-1} + \Gamma_2 \Delta q_t + \Gamma_3 \Phi \Gamma_3^{-1}(z_{t-1} - \Gamma_1 R_{t-2} - \Gamma_2 \Delta q_{t-1}) + \Gamma_3 \varepsilon_t \tag{19}$$

where $\tilde{z}_t' = \begin{bmatrix} \tilde{y}_t & \pi_t & \Delta e_t - \pi_t & R_t \end{bmatrix}$.

5.2 Constructing the Auxiliary System for LI Estimators

When Limited Information (LI) estimation is performed it is necessary to specify a VARX process for whatever RHS endogenous variables appear in the structural equation and to use it to form future expectations. There are two possibilities. One is to use the VARX coming from LS's Full Information (FI) estimates. The other is to estimate a VARX unrestrictedly using only the constraint that it has the same variables as in the LS model. In order to encompass a wider range of possible system models we allow this VARX to be of second order:

$$\tilde{z}_t = A_1\tilde{z}_{t-1} + A_2\tilde{z}_{t-2} + A_3\Delta q_t + A_4\Delta q_{t-1} + A_5 u_t. \tag{20}$$

We will refer to this as the LI VARX. The VARX coming from the Bayesian FI estimates estimates will be termed the FI VARX.

Now, there are four unobservable shocks in (20) - the same as in the LS model. One of these shocks, technology, is present in both systems and is defined in the same way so as to retain comparability on output gap measures. The other three shocks are unnamed in (20) - they would be combinations of structural shocks in whatever model generated the data. Although it is reasonable to assume that the shocks u_t are serially uncorrelated, due to the fact that the LI VARX is second order, it is necessary to allow these shocks to be contemporaneously correlated, as they are not (necessarily) structural shocks. To satisfy such requirements u_t is taken to be constructed from three e_{jt} that are $n.i.d.(0,1)$ and the technology shock ε_t^a makes up the fourth fundamental shock, leading to u_t having the form

$$\begin{bmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \\ u_{4t} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \gamma_{21} & 1 & 0 & 0 \\ \gamma_{31} & \gamma_{32} & 1 & 0 \\ \delta_1 & \delta_2 & \delta_3 & 1 \end{bmatrix} \begin{bmatrix} e_{1t} \\ e_{2t} \\ e_{3t} \\ \varepsilon_t^a \end{bmatrix}.$$

This structure allows for a general covariance matrix for u_t. The parameters of the LI VARX(2) to be estimated are A_1, A_2, A_3, A_4, $\gamma_{ij}, \delta_j, \rho_a, \sigma_a$ and $\sigma_j^2 = var(e_{jt})$. This is a latent VARX process as $\tilde{y}_t$ are unobserved with the observed data being $z_t' = \begin{bmatrix} \Delta y_t & \pi_t & \Delta e_t - \pi_t & R_t \end{bmatrix}$. Note that Δq_t is observable and exogenous so that the parameters ρ_q and σ_q can be estimated by a least squares regression of Δq_t against Δq_{t-1}[11]. We also need to define to relate the observable and latent variables through

$$\Delta y_t = \Delta \tilde{y}_t + \Delta y_t^p = \Delta \tilde{y}_t + \frac{\varepsilon_t^a}{1-\rho_a}.$$

Estimates of the parameters were then found by applying MLE to this system. Once $\tilde{A}_j$ are determined, $E_t(\pi_{t+1})$ can be computed from the VARX equations.

5.3 Estimation and Evaluation of the Phillips Curve

We will look at estimation of the Phillips curve. The DSGE model parameters present in the equation are α,$\beta, \tau, \kappa, \rho_{y^*}, \sigma_{y^*}, \rho_q, \sigma_q$, β, ρ_q, σ_a, and ρ_a . Some of these are fixed (β, σ_q), and others - such as ρ_a, σ_a - come from either the FI or LI VARX values, depending on which auxiliary system we use. Basically, interest is in the model parameters α, κ, τ. Therefore, Table 1 presents estimates of the means of the posteriors for these parameters, along with 90% confidence intervals for three estimators. The latter are the original LS estimates and two LI estimators that differ depending on whether they use the FI or LI VARX auxiliary systems.[12] Posteriors are presented since LS performed Bayesian estimation. Bayesian methods can be used with the LI estimators as these work with a LI likelihood and so that can be combined with the same priors as used by LS. The estimation was performed in Dynare by taking the LS Phillips curve and then augmenting it with equations for $E_t\pi_{t+1}$, $\tilde{y}_t, de_t$ and R_t that come from either the FI or LI VARX systems.[13] In the case when the LI VARX provides the auxiliary equations the parameters in it are re-estimated with priors set to a normal density for the A_j and quite wide standard deviations. When the FI VARX is used as the set of auxiliary equations its parameters were not re-estimated.

[11] When this is done ρ_q turns out to be negative. LS constrained it to be positive because of their use of a beta density for the prior on ρ_q. We replace the LS prior with a normal density in what follows. The change has a relatively minor effect on the FIML estimates.

[12] Dynare Version 3.065 was used to get these

[13] Because of the way DYNARE works equations describing the evolution of the exogenous variables $\Delta q_t, \Delta a_t$ and $\widetilde{y}_t^*$ are needed as well.

Table 1 Various Estimates of the Phillips Curve				
		Mean		90% Conf Int
α_{LS}		.14		.09-.19
α_{FIV}		.16		.11-.23
α_{LIV}		.24		.16-.32
κ_{LS}		1.78		1.17-2.39
κ_{FIV}		.09		.03-.15
κ_{LIV}		.25		.09-.42
τ_{LS}		.46		.33-.61
τ_{FIV}		.63		.38-.86
τ_{LIV}		.23		.07-.39

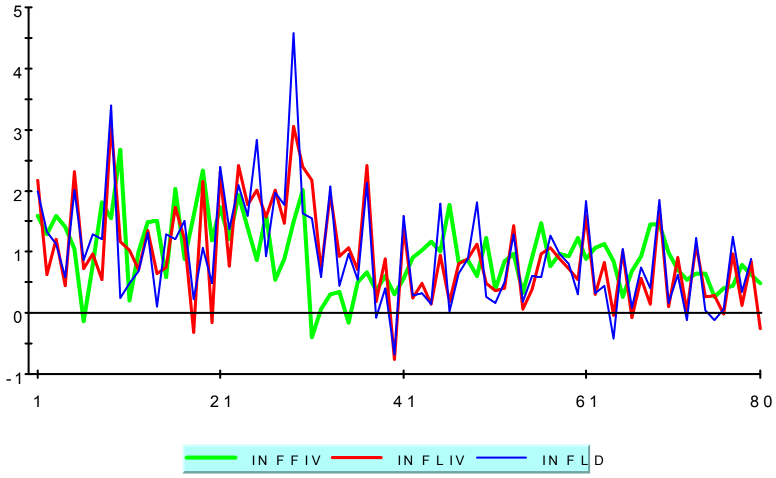
There is one striking difference between the FI LS estimates and those from LI, and that relates to κ - both instances of the latter the posterior mean is far below what is obtained by LS and the estimates lie well outside the 90% confidence intervals produced by the latter.[14] Because κ appears *only* in the Phillips curve this suggests that the assumptions made in estimating the rest of the model have had an enormous impact upon the posterior density for κ and it is very likely that the large effect of the output gap upon inflation suggested by LS's FI estimates is due to mis-specification bias in the complete model.

If we just focus upon the limited information estimates there are also some significant differences between them. To assess the likely reason we note that the major difference between these estimators is the different estimates of $E_t\pi_{t+1}$ that are used. Figure 1 therefore compares these along with the actual inflation rate. It is clear that the estimates from FI VARX give an expected inflation that is very different from the actual one and does not seem to exhibit rational expectations. To formally test this we regress the expectation error from FI VARX, $\pi_{t+1} - E_t\pi_{t+1}$, against π_{t-1}. and find a coefficient of .4 with t ratio of 3.5. This is the same as implementing a parametric encompassing test of the type described earlier.

We finish by looking at whether the Phillips curve has a hybrid structure.

[14] As we noted above this estimate differs from that published by LS because we have fixed ρ_q but the difference is not major.

Fig 1 Estimates of Expected Inflation from
FIML VARX and LIML VARX and Actual Inflation



To do this $\beta E_t(\pi_{t+1})$ is replaced by $\beta[(1-c)E_t(\pi_{t+1})+c\pi_{t-1}]$ and re-estimated with the limited information estimator. If there is a hybrid structure to the Phillips curve then the FI VARX is too low and order but the LI VARX would be high enough. Hence we use it. The estimates of α, τ and κ are very close to the LIV estimates of Table 1 and there is evidence of the need for a hybrid structure since the posterior mean of c is .28 with confidence interval .06-.49.

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 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 and the discrepencies suggest that this model has substantial specification errors.

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.

Brundy, J.M. and D.W. Jorgenson (1971), "Efficient Estimation of Simultaneous Equations by Instrumental Variables", *Review of Economics and Statistics,* 53,207234.

Canova, F., M. Finn and A.R. Pagan (1994) "Evaluating a Real Business Cycle Model" in C. Hargreaves (ed) *Non-Stationary Time Series Analysis and Co-Integration* (Oxford University Press), 225-255.

Fuhrer, J.C. and G.P. Olivei (2004), "Estimating Forward-Looking Euler Equations with GMM and Maximum Likelihood Estimators: An Optimal Instruments Approach", *paper presented to the Conference on Models and Monetary Policy*, Federal Reserve Bank Board of Governors, March

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

Kaiser, R and A. Maravall (2002), "A Complete Model-Based Interpretation of the Hodrick-Prescott Filter: Spuriousness Reconsidered", Working Paper 0208, Banco de Espana

Gourieroux, C., A. Monfort and E. Renault (1993), "Indirect Inference", *Journal of Applied Econometrics*, 8, 85-118.

G. Kapetanios, A.R. Pagan and A. Scott "Making a Match: Combining Theory and Evidence in Policy-oriented Macroeconomic Modeling", *Journal of Econometrics*, 136, 505-594.

Ireland, P.(2004) "Technology Shocks in the New Keynesian Model", *Review of Economics and Statistics*, 86, 923-936.

Kurmann, A. (2007), "VAR-based Estimation of Euler Equations with an Application to New Keynesian Pricing", J*ournal of Economic Dynamics and Control*, 31, 767-796.

Lubik, T.A. and F. Schorfheide (2007) " Do Central Banks respond to exchange Rate Movements? A Structural Investigation", *Journal of Monetary Economics,* 54, 1069-1087.

McCallum, B. (1976), "Rational Expectations and the Natural Rate Hypothesis: Some Consistent Estimates", *Econometrica*, 44, 43-52.

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

Sbordone, A. (2004), "A Limited Information Approach to the Simultaneous Estimation of Wage and Price Dynamics", (mimeo, Federal Reserve Bank of New York)

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

Smith, A.A.(1993) " Estimating Nonlinear Time Series Models using Simulated Vector Autoregressions", *Journal of Applied Econometrics*, 8, S63-S84.