



DP2008/11

**Limited Information Estimation and Evaluation of
DSGE Models**

Martin Fukač and Adrian Pagan

July 2008

JEL classification: C13, C39, C51, C52, E39

www.rbnz.govt.nz/research/discusspapers/

Discussion Paper Series

ISSN 1177-7567

DP2008/11

**Limited Information Estimation and Evaluation of DSGE
Models***

Martin Fukač[†] and Adrian Pagan[‡]

Abstract

We advance the proposition that dynamic stochastic general equilibrium (DSGE) models should not only be estimated and evaluated with full information methods. These require that the complete system of equations be specified properly. Some limited information analysis, which focuses upon specific equations, is therefore likely to be a useful complement to full system analysis. Two major problems occur when implementing limited information methods. These are the presence of forward-looking expectations in the system as well as unobservable non-stationary variables. We present methods for dealing with both of these difficulties, and illustrate the interaction between full and limited information methods using a well known model.

* The views expressed in this paper are those of the author(s) and do not necessarily reflect those of the Reserve Bank of New Zealand.

[†] Economics Department, 2 The Terrace, Wellington, New Zealand. E-mail: *first-name.lastname@rbnz.govt.nz*.

[‡] Queensland University of Technology, and University of New South Wales. Research of this author was supported by ARC Grant No. DP0449659.

1 Introduction

Models are a vital ingredient in the process of analyzing policy options. But there is uncertainty about which model to choose. If decisions are to be made, it is likely that most decision makers would prefer a model that is capable of producing a match with the data rather than one that simply illustrates scenarios. For this reason, methods are needed that utilize the data for calibrating model parameters, and assessing the correspondence between what the model implies, and what the data tells us. This challenge has been present ever since the first econometric models were used.

Initially, investigators spent considerable time ensuring that the individual equations of the model were satisfactory, before proceeding on to an evaluation of how well the equations fitted together as a system. To do this, it was deemed preferable to work with limited information (LI) methods for estimation and evaluation, as this philosophy encapsulated the view that some parts of the system were better understood than others. Such uncertainty about the details of the model encouraged the development of techniques that tried to study the individual structural equations in isolation from the remainder of the system. Thus, instrumental variable methods of estimation, such as 2SLS, became the norm. Only after researchers were happy with the single equations did they proceed to investigate how well the equations fitted together as a system.

Today dynamic stochastic general equilibrium (DSGE) models are becoming widely used in both academic and central bank research. In contrast to the earlier tradition, the emphasis in DSGE modelling has been upon the system as a whole. Systems estimators, such as full information maximum likelihood (FIML), have been the preferred approach. Evaluation has emphasized system fit; basically offering a comparison with VAR models. This seems unfortunate. As earlier researchers recognized, the single equation and system measures of fit were complements, and not substitutes. Given uncertainty about the model, it was wise to look at the building blocks as well as the complete model.

In this paper, we seek to apply LI methods to the structural equations of DSGE models. In the standard, simultaneous equations context, LI methods were implemented by augmenting the structural equation of interest with reduced form equations for the “right hand side” endogenous variables in that equation. For two reasons, that solution is not readily available for DSGE models. One is the presence of forward-looking expectations, and the second occurs if there are unit root shocks in the model. In relation to the first, Kurmann (2007) has shown that

following the standard LI approaches without modification may potentially lead to inconsistent estimators. The second problem has been dealt with in one of two ways of removing the unit root in variables - either by a preliminary filtering of the data or by a scaling of the variables to induce stationarity.

Section 2 of the paper briefly states the connection between complete structural (DSGE) models, and the VAR formed from the variables appearing in the model. We will be extensively using both of these constructs in what follows. In section 3 of the paper it is assumed that the data are $I(0)$ variables.¹ We describe a LIML estimator that can be implemented, and which avoids the difficulties identified by Kurmann (2007). Section 4 then turns to the situation when the structural equations originally incorporate $I(1)$ variables. We discuss both off-model and on-model (model consistent) approaches for converting such equations to ones that involve only observable $I(0)$ variables. We argue that it is inadvisable to use the former. Section 5 looks at how to generalize the LIML estimator to estimate DSGE model parameters when there are unit roots in the model. It applies the method to the estimation of the Phillips curve in Lubik and Schorfheide's (2007) DSGE model of a small open economy. We also suggest some methods for determining the goodness of fit of the model. Section 6 concludes.

2 DSGE Model Solution for $I(0)$ Variables

DSGE models have the following stylized representation

$$B_0 z_t = B_1 z_{t-1} + D x_t + C E_t z_{t+1} + G \varepsilon_t, \quad (1)$$

where z_t is a vector of $n \times 1$ endogenous variables, x_t is a set of observable, and ε_t a set of unobservable shocks. Generally z_t will be the log of some levels variables Z_t . 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

¹ We will treat these variables as observable. If not then one simply uses the state space form that distinguishes between model and observable variables, and the likelihood is computed with the Kalman filter. This is a simple adaption. The key is to work out the nature of the process for generating the model variables and so we focus upon that.

equations, and possibly a set of identities. The parameters in the DSGE model will be designated as θ .

The solution to this system has the form²

$$z_t = Pz_{t-1} + \sum_{j=0}^{\infty} \Pi_1^j (\Pi_2 E_t x_{t+j} + \Pi_3 E_t G \varepsilon_{t+j}),$$

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 x_t and ε_t are AR(1) processes with matrices Φ_x , and Φ_ε , this reduces to a Vector Autoregression with Exogenous Variables (VARX) system for z_t of the form

$$z_t = Pz_{t-1} + \bar{D}x_t + \bar{G}\varepsilon_t. \quad (2)$$

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

$$E_t(z_{t+1}) = Pz_t + \bar{D}\Phi_x x_t + \bar{G}\Phi_\varepsilon \varepsilon_t. \quad (3)$$

Hence, one could solve for the conditional expectation of z_{t+1} once P , \bar{D} , and \bar{G} are known. Since these are functions of the model parameters 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 ε_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 ε_t). The FIML estimator normally used with DSGE models maximizes the likelihood of (2).

3 LI Estimation and Evaluation of DSGE Models with I(0) Variables

3.1 Estimation

We consider a single Euler equation from the system in (1). For simplicity, we assume that the forward looking expectation in the equation is of the dependent variable for that equation, and there are no exogenous variables ($\bar{D} = 0$). It is easy to see how to handle the general case once the principles of LI estimation are understood. Therefore, this single equation has the form

$$z_{1t} = z_t^{-1} \beta_{10} + z'_{1t-1} \beta_{11} + E_t z'_{1t+1} c_1 + \varepsilon_{1t}, \quad (4)$$

² We refer to Binder and Pesaran (1995) for the solution method.

where z_t^- are the sub-set of z_t that contains the RHS endogenous variables. Correspondingly, the sub-set of (2) that deletes the equation corresponding to z_{1t} is

$$z_t^- = P^- z_{t-1} + \bar{G}^- \varepsilon_t. \quad (5)$$

The solution to the complete model will be (2), with $\bar{D} = 0$.

Due to the presence of the forward looking expectation in the system, difficulties can arise when performing LI estimation of (4). Initially, McCallum (1976) proposed to replace $E_t z_{1t+1}$ by z_{1t+1} , after which instrumental variables were applied to (4). Instruments were chosen from z_{t-1} (or a sub-set of it). Most frequently, one uses linear combinations of z_{t-1} , e.g. regress z_{1t+1} against z_{t-1} and use the predictions from this as an instrument for z_{1t+1} . Effectively, one is just treating z_{1t+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 were a number of suggestions about this, apart from the 2SLS solution. One proposal by Brundy and Jorgenson (1971) was to use as an instrument the conditional expectation of the endogenous variable to be instrumented, where the conditioning is upon the lagged regressors in the derived reduced form. Fuhrer and Olivei (2005) make the same suggestion in order to produce an instrument for z_{1t+1} . They implement it in the following way. Using some values for β_{10} , β_{11} , c_1 , and with P^- replaced by \hat{P}^- , the OLS estimate of the regression of z_t^- upon z_{t-1} , they solve the system composed of (4) and (5) for z_{1t} . This would give a solution for z_{1t} of the form $z_{1t} = P_1 z_{t-1} + G_1 \varepsilon_t$, so that combinations of z_{t-1} would then be the instrument for z_{1t+1} . Because P_1 is a function of β_{10} , β_{11} , c_1 , and P^- the procedure is iterated, although P^- is held constant at \hat{P}^- .

A different approach, used by Sbordone (2006), is to recognize that P is a function of β_{10} , β_{11} , c_1 , and P^- . To find estimates of these parameters, Sbordone minimizes the difference between \hat{P} (the OLS estimator) and $\tilde{P} = g(\tilde{\beta}_{10}, \tilde{\beta}_{11}, \tilde{c}_1, \tilde{P}^-)$, where $\tilde{\beta}_{10}$, $\tilde{\beta}_{11}$, \tilde{c}_1 , and \tilde{P}^- are trial values. Basically, this approach yields an indirect estimator of the parameters β_{10} , β_{11} , c_1 , and P^- , where the model equations are taken to be (4) and (5), while the auxiliary model is (2). The theory of such estimators is set out in Gourieroux, Monfort and Renault (1993). Smith (1993) developed and applied this indirect estimating method when he estimated the parameters of a complete DSGE model by using the VAR as an auxiliary model.

None of the above estimators would be the equivalent of a LIML estimator of a structural equation with forward-looking expectations. To see why, we note that, in the context of estimating a single structural equation without expectations, the

LIML estimator would be found by jointly estimating the parameters of a pseudo-system that consisted of the structural equation of interest and the reduced form equations for any RHS endogenous variables, z_t^- ; see Pagan (1979) for example. Consequently, it is tempting to think that one could derive a LIML estimator in the case where there are forward looking expectations by setting up the pseudo-system system composed of (4) and (5), and then choosing β_{10} , β_{11} , c_1 , P^- , and \bar{G}^- to maximize the likelihood of this system. Of course, the difficulty is that $E_t z_{1t+1}$ is unknown, but for any given values of the parameters, we might think about using standard solution methods to find an expression for $E_t(z_{1t+1})$ from the pseudo-system. If the shocks are all white noise, this will equal $E_t^{PS}(z_{1t+1}) = F_1^* z_t + G_t^* \varepsilon_t$, where F_1^* , G_t^* are determined by whatever values are set for β_{10} , β_{11} , c_1 , P^- , and \bar{G}^- and E_t^{PS} will be functions of z_{t-1} and the shocks in the pseudo-system.

Kurmann (2007) pointed out that there was a possible problem with this strategy. It arises from the fact that, even when a solution for z_{1t} exists for the true system, if one tries to solve for z_{1t} from the pseudo-system (4) and (5) above (even using the true β_{10} , β_{11} , c_1 , Φ) a solution may not exist. If this happens, then any algorithm will try to find values for the structural equation parameters that does produce a solution for z_{1t} , and hence it will move away from the true values of these coefficients i.e. produce inconsistent estimators of them. Fundamentally, this problem arises due to an incorrect estimate of $E_t(z_{1t+1})$ coming from solving the pseudo-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 a second structural equation. An explicit recognition of this contemporaneous dependence turns out to be crucial for the existence of a solution for z_{1t} , and z_{2t} in his example.

Now, as just noted, the potential inconsistency of the MLE arises from the fact that the correct expected value of z_{1t+1} is $E_t(z_{1t+1})$ from (2) and, in general, this would not equal that found from the pseudo-system, $E_t^{PS} z_{1t+1}$. Hence, if we replaced $E_t z_{1t+1}$ with $E_t^{PS} z_{1t+1}$ in the structural equation it would become

$$z_{1t} = z_t^{-'} \beta_{10} + z_{1t-1}' \beta_{11} + E_t^{PS}(z_{1t+1}) c_1 + \varepsilon_{1t} + c_1 (E_t z_{1t+1} - E_t^{PS} z_{1t+1}). \quad (6)$$

Then the MLE working with the pseudo-system would estimate (6) using instruments for $E_t^{PS}(z_{1t+1})$ such as z_{t-1} . These instruments are uncorrelated with ε_{1t} , but, since $E_t z_{1t+1} - E_t^{PS} z_{1t+1}$ will generally be a function of z_{t-1} , the expectation of this with the instruments will not be zero. Consequently, it is important that

one solves for the correct expectation of z_{1t+1} and utilizes it in the structural equation. Notice that the situation is different if one was using an IV estimator, as in Fuhrer and Olivei (2005). In that instance, one replaced $E_t z_{1t+1}$ with z_{1t+1} , and then instrumented this with some combination of z_{t-1} . This will always produce consistent estimators of the parameters of (4), since any combination of z_{t-1} is uncorrelated with $\varepsilon_{1t} + (E_t(z_{1t+1}) - z_{1t+1})'c_1$.

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. It involves estimating $E_t(z_{1t+1})$ from the correct model rather than the pseudo-system described above. But this is just a matter (at least in large samples) of forming a new pseudo-system that consists of the structural equation to be estimated, (5) to describe z_t^- , and an equation for $E_t z_{1t+1}$ that derives from that for z_{1t} in (2). Basically, (2) is being used to produce instruments for z_t^- and to generate future expectations for z_{1t} , since it incorporates (in large samples) the nature of the remaining structural equations, which are being ignored in solving just (4) and (5). Using the expectations computed in this way in Kurmann’s example produces exactly the same solution as in the correct system. Thus, to validly implement a LIML estimator requires that one estimate by maximum likelihood a system composed of

$$z_{1t} = z_t^{-'}\beta_{10} + z_{1t-1}'\beta_{11} + E_t z_{1t+1}'c_1 + \varepsilon_{1t}, \quad (7)$$

(5), and an equation describing $E_t z_{1t+1}$ (based on the VAR in (2)).³ As before, the equations in (5) and that describing $E_t z_{1t+1}$ could be either estimated jointly with the structural parameters or first estimated separately and then treated as known values. The first of these treatments is the analogue of LIML, while the second is a 2SLS-like estimator. Notice that the errors of (7),(5) and the augmenting equation for expectations must be allowed to be correlated, as the shocks in the latter two are functions of the shocks in (7).

3.2 Evaluation

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

³ Of course the answers using the correct estimator may not be very different from those with the incorrect one. Much will depend on the degree of simultaneity in the system. Kurmann presents cases where there is a major difference but Korenok, Radchenko and Swanson (2008) find little difference when estimating the Phillips curve with the LI estimator we describe here.

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 either the LIML or FIML estimates of the structural model, and it is reasonable to perform them with both sets of estimates. We discuss each of these three testing procedures in turn.

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 pre-determined value, is a natural extension to be checked, and we utilize this in the example below.

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.

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 model-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, Finn and Pagan (1994) for example. Modern versions of it try to map the differences between \tilde{P} and \hat{P} into a scalar 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 (7) and (5), and comparing this to

⁴ Again, we are ignoring the fact that the solution may be a VARX rather than VAR system. We set $\bar{D} = 0$.

\hat{P} . This would be a limited information encompassing test, as it only imposes the restrictions from the structural equation upon the VAR.

4 Handling Permanent Components

4.1 Three Strategies

Some of the series in DSGE models are likely to be $I(1)$. Because estimation is straightforward if the variables are $I(0)$, some transformation needs to be applied to the $I(1)$ variables to produce $I(0)$ variables, after which the system can be re-formulated in terms of the latter. This is an issue of model design. Moreover, in many DSGE models co-integration is implied between variables, and in that case the $I(1)$ variables are driven by a common factor. Bearing this in mind, we find that there have generally been three strategies to handle the $I(1)$ nature of variables.

A. The structural equations of the DSGE model might be re-formulated in terms of the error correction terms and changes in the variables.

B. An alternative decomposition would be to express the system in terms of $I(0)$ variables $z_t - z_t^p$ and the changes in the permanent components, z_t^p , i.e. Δz_t^p .

C. Finally, we could use a set of variables made up of $z_t - z_t^*$, and Δz_t^* , where z_t^* consists of a combination of z_t^p and some $I(0)$ series. This is the standard approach in many DSGE models with a single $I(1)$ shock, such as technology, and it involves scaling the level variables Z_t with the levels of the shock.

Once one has decided on one of these three strategies, it is necessary to measure the $I(0)$ variables. This would involve estimating the co-integrating vectors (to find the error correction terms), or one of the $I(1)$ components z_t^p, z_t^* . To perform this task, there have been two approaches.

1. The $I(0)$ variables are constructed “off-model” by filtering the data, after which estimation is performed as described in section 3. Hence, $I(0)$ variable construction and parameter estimation are done sequentially.

2. The model is used to construct estimates of the variables in one of strategies A to C. In this instance, the construction of the $I(0)$ variables, and the estimation of the model parameters, needs to be done simultaneously.

We argue in the next sub-section that it is inadvisable to do the filtering “off-model”, as this has the potential to lead to shocks with incorrect properties, as well as the possibility of inconsistent estimators of the model parameters.

4.2 An Example

The model, which we will utilize later as an example of the methods, is in Lubik and Schorfheide (LS, 2007). It is a four-equation DSGE model of a small open economy, comprising an IS curve, a Phillips curve, an interest rate rule, and an exchange rate equation. To appreciate the differences between the strategies B and C above, we note that LS’s Phillips curve has the form

$$\pi_t = \beta E_t \pi_{t+1} + \alpha \beta E_t \Delta q_{t+1} - \alpha \Delta q_t + \frac{\kappa}{\tau + \theta} \bar{y}_t + \frac{\kappa \theta}{\tau(\tau + \theta)} \tilde{y}_t^*, \quad (8)$$

where α is the import share, β is the discount factor, κ is a “price stickiness” parameter, τ is the inter-temporal elasticity of substitution, and $\theta = \alpha(2 - \alpha)(1 - \tau)$. π_t is the domestic inflation rate, q_t is the observable terms of trade, a_t is the log level of (unobservable) technology, \tilde{y}_t^* is the foreign (unobservable) output. a_t evolves as $\Delta a_t = \rho_a \Delta a_{t-1} + \varepsilon_t^a$ and $\bar{y}_t = y_t - a_t$, meaning that they are using strategy C above to reduce the $I(1)$ variable y_t to an $I(0)$ form, \bar{y}_t .

In contrast, we use strategy B above, which defines the $I(0)$ variable to be $\tilde{y}_t = y_t - y_t^p = y_t - a_t^p$, giving the Phillips curve the form

$$\pi_t = \beta E_t \pi_{t+1} + \alpha \beta E_t \Delta q_{t+1} - \alpha \Delta q_t + \frac{\kappa}{\tau + \theta} \tilde{y}_t + \frac{\kappa \theta}{\tau(\tau + \theta)} \tilde{y}_t^* + \frac{\kappa \rho_a}{(\tau + \theta)(1 - \rho_a)} \Delta a_t, \quad (9)$$

where we have used the fact that $a_t^p = a_t + \frac{\rho_a}{1 - \rho_a} \Delta a_t$, and

$$\begin{aligned} y_t - a_t = \tilde{y}_t &= y_t - y_t^p + a_t^p - a_t \\ &= \tilde{y}_t + a_t^p - a_t \\ &= \tilde{y}_t - \frac{\rho_a}{1 - \rho_a} \Delta a_t. \end{aligned}$$

After reformulating the remaining equations of LS’s model in terms of \tilde{y}_t , the IS curve will be

$$\tilde{y}_t = E_t \tilde{y}_{t+1} - (\tau + \theta)(R_t - E_t \pi_{t+1}) - \alpha(\tau + \theta) \rho_q \Delta q_t - \frac{\theta}{\tau} (1 - \rho_{y^*}) \tilde{y}_t^*, \quad (10)$$

while the exchange rate equation is

$$\Delta e_t - \pi_t = -(1 - \alpha)\Delta q_t - \pi_t^*, \quad (11)$$

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 becomes

$$\begin{aligned} R_t = & \rho_R R_{t-1} + (1 - \rho_R) [(\psi_1 + \psi_3)\pi_t + \psi_2 \tilde{y}_t + \psi_3 (\Delta e_t - \pi_t)] \\ & + \psi_2 (1 - \rho_R) \frac{\rho_a}{1 - \rho_a} \Delta a_t + \varepsilon_t^R, \end{aligned} \quad (12)$$

where ρ_R is the interest rate smoothing parameter and ψ_1, ψ_2, ψ_3 are the weights that the monetary authority places on inflation, output, and exchange rate stabilization, respectively. ε_t^R is an *i.i.d* monetary policy shock with standard deviation of σ_R .

Exogenous variables evolve as AR(1) processes, where the shocks ε_t^j for $j = \{q, a, y^*, \pi^*\}$ are all *i.i.d*. with standard deviations of σ_j .

$$\Delta q_t = \rho_q \Delta q_{t-1} + \varepsilon_t^q, \quad (13)$$

$$\Delta a_t = \rho_a \Delta a_{t-1} + \varepsilon_t^a \quad (14)$$

$$\tilde{y}_t^* = \rho_{y^*} \tilde{y}_{t-1}^* + \varepsilon_t^{y^*}, \quad (15)$$

$$\pi_t^* = \rho_{\pi^*} \pi_{t-1}^* + \varepsilon_t^{\pi^*}. \quad (16)$$

We can see that switching between the types of representations changes the Euler equations. In some instances, it simplifies them, and in other cases, it makes them more complex. It also sometimes makes some assumptions quite clear, e.g. the IS curve now has one shock in it, and the Phillips curve has two. In this form, the two equations would have correlated error terms since there is a common component \tilde{y}_t^* , and if one thought of the system above as an SVAR, the shocks in the equations would not be uncorrelated.

4.3 Off-Model Extraction of Permanent Components

If we utilize strategy B, how do we extract a permanent stochastic component from a series i.e. perform a decomposition of a series (or a vector of them) into permanent and transitory components? It has to be recognized that there is no unique way of doing this. If we write $z_t = z_t^P + z_t^T$, with z_t^P and z_t^T being the permanent and transitory components, we could always add a transitory component to

z_t^P , and simultaneously subtract it from z_t^T , and we would still have a permanent-transitory decomposition. So one needs to add on some restriction to make the decomposition unique. An early method of doing this was the Beveridge-Nelson (BN) method, where a permanent component in z_t was defined as⁵

$$\hat{z}_t^P = z_t + \sum_{j=1}^{\infty} E_t(\Delta z_{t+j}),$$

with the transitory component being $-\sum_{j=1}^{\infty} E_t(\Delta z_{t+j})$. The model of z_t used in this definition was that Δz_t followed a linear stochastic process, and the constraint to produce the decomposition was that $\Delta z_t^P = e_t$, where e_t is white noise, i.e. $E_t(\Delta z_{t+1}^P) = 0$. Applying the definition to a pure random walk, $\Delta z_t = v_t$, we would get a zero transitory component. If, however, $\Delta z_t = \rho \Delta z_{t-1} + v_t$, with v_t being white noise, we would have $\Delta z_t^P = \frac{v_t}{1-\rho}$, and $z_t^T = -\frac{\rho}{1-\rho} v_t$. Because BN is appropriate for linear processes, it is the preferred definition for DSGE models that have the series being generated in that way.

There are other univariate methods of extracting a permanent component. A familiar one is the Hodrick-Prescott (HP) filter. This gets a unique estimate by imposing a smoothness constraint on the estimates. To understand some of its limitations in our context, we draw on the demonstration in Harvey and Jaeger (1993) and Kaiser and Maravall (2002) that HP can be regarded as the Kalman smoother applied to a two component process for z_t , in which the permanent component z_t^P is an $I(2)$ process. Since this means z_t would also be $I(2)$, it would be rare for a DSGE model to have this implication, and so it is a filter that is inconsistent with such models.⁶

We will see later that $E_t \Delta z_{t+1}^P = 0$ is a requirement for consistent estimation of the parameters of Euler equations when pre-filtered data is used in estimation. This holds for the BN filter, but fails for HP. To see this, note that Singleton (1988) shows the HP estimate of the permanent component (in large samples) to have the form of

$$z_t^P = \sum_{j=-T}^T a_j z_{t-j},$$

where the weights a_j are ($\lambda = 1600$)

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

⁵ Here, we are assuming that $E_t(\Delta z_t) = 0$. This definition of the BN filter applies if z_t is a vector, but then $E_t(\cdot)$ involves using information on all the elements in Δz_t to form $E_t \Delta z_{t+j}$. Thus the transitory component of z_t , often called a "gap", is constructed by using whatever information is valuable for forecasting future growth rates of z_t .

⁶ This is because one normally has a unit root process for technology.

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. 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. To illustrate the effect, we simulated 500 observations from a data generating process 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 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 1.⁷ Based on the unobserved components formulation of the HP filter given above, where z_t^p was an $I(2)$ process, this is to be expected.

4.4 Estimation Problems with Off-Model Filters

Now let us discuss the impact of off-model filtering upon the estimation of DSGE models. The levels of the variables in the DSGE model will be Z_t . If these are $I(1)$ we will need to choose some variables Z_t^* that are $I(1)$ with the property that $\ln Z_t^*$ cointegrates with $\ln Z_t$ and $\tilde{z}_t = \ln(Z_t/Z_t^*)$ is $I(0)$. In the case of a stationary variable Z_t^* would of course be a constant. The original Euler equations typically have the form $Z_t = E_t g(Z_{t-1}, Z_{t+1}, \dots)$ and the function $g(\cdot)$ is such that this can be re-expressed in terms of the ratios of Z_t to Z_t^* . For simplicity we will consider a special case where the Euler equation to be studied has the form $Z_{1t} = E_t g(Z_{1t-1}, Z_{1t+1}, Z_t^-)$, where Z_t^- are $I(0)$ variables. This could for example be an IS curve. Then, after log-linearization, we would get the structural equation

$$\begin{aligned} \tilde{z}_{1t} = & \tilde{z}'_{1t-1} \beta_{11} + E_{t+1} \tilde{z}'_{1t+1} c_1 + z_t^{-\prime} \beta^- + \{E_t (\Delta z_{1t+1}^*)' c_1 \\ & - (\Delta z_{1t}^*)' \beta_{11}\} + \varepsilon_{1t}, \end{aligned} \quad (17)$$

where $\tilde{z}_{1t} = z_{1t} - z_{1t}^*$. Clearly one needs to account for the fact that the errors in the transformed equations are not the original shocks as they consist of the latter plus the term in the brackets $\{\}$, and so will depend upon how the permanent component is measured.

Now let us look at some special cases. First $\beta_{11} = 0, c_1 = 1$. In this case the term in curly brackets is $E_t \Delta z_{1t+1}^*$. Now, if the permanent component was formed using the BN filter this term is zero and so estimation can proceed without any

⁷ Regressing Δz_t^p against $\Delta z_{t-1}^p, \Delta z_{t-2}^p, \Delta z_{t-3}^p$, and Δz_{t-4}^p gives coefficient values of 3.48, -4.57, 2.67, and -0.58, which do sum to 1. However, this regression shows that the process for Δz_t^p is not a pure random walk.

difficulties and the shock ε_{1t} can be accurately measured. If however one uses the HP filter $E_t \Delta z_{1t+1}^* \neq 0$, and so we will now find that any IV estimator of the Euler equation parameters will be inconsistent and the shocks would need to be purged of the $E_t \Delta z_{1t+1}^*$ term. Moreover, because $E_t \Delta z_{1t+1}^*$ seems to very persistent if the HP permanent component is used the shock will be very persistent. 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.

If it is the case that $\beta_{11} \neq 0$ then it is clear that filtering the levels variables z_t with either the BN or HP filters would lead to difficulties due to the presence in the errors of the term $-(\Delta z_t^*)' \beta_{11}$. One needs to take explicit account of the nature of the permanent component that is implied by the model.

5 LIML Estimators for DSGE Models with Permanent Components

5.1 Solution of DSGE Models with Permanent Components

Now let the z_t variables be $I(1)$, and z_t^p be their permanent components. As in strategy B, the structural equations will be transformed to a new form, in which two sets of variables, $\tilde{z}_t = z_t - z_t^p$ and Δz_t^p , are present.⁸ Thus the system of equations of the DSGE model will consist of

$$B_0 \tilde{z}_t = B_1 \tilde{z}_{t-1} + D x_t + C E_t \tilde{z}_{t+1} + F_0 \Delta z_t^p + F_1 \Delta z_{t-1}^p + G \varepsilon_t.$$

We will add to this system equations describing Δz_t^p of the form $\Delta z_t^p = \xi_t$, where ξ_t are the white noise shocks driving the common permanent components of the system. Then, defining $\zeta_t = [\tilde{z}_t' \ \Delta z_t^{p'}]'$ and $\eta_t = [\varepsilon_t' \ \xi_t']'$, 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, $\eta_t = H^+(\zeta_t - P \zeta_{t-1})$, where H^+ is the pseudo-inverse of

⁸ 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.

H, and if η_t is a VAR(1) of the form $\eta_t = \Phi_\eta \eta_{t-1} + \psi_t$, we get

$$\begin{aligned}\zeta_t &= P\zeta_{t-1} + H\Phi_\eta H^+(\zeta_{t-1} - P\zeta_{t-2}) + H\psi_t \\ &= (P + H\Phi_\eta H^+)\zeta_{t-1} - H\Phi_\eta H^+P\zeta_{t-2} + u_t \\ &= A_1\zeta_{t-1} + A_2\zeta_{t-2} + u_t.\end{aligned}$$

Consequently, ζ_t follows a VAR(2).⁹ Therefore, the only difference between the situation where all variables were $I(0)$, and the situation when there are permanent components, is that the VAR now contains some latent variables in the form of Δz_t^p .

Because ζ_t is not completely observable now, it is necessary to estimate a latent VAR. This can be done with any program that performs MLE with the Kalman filter, since that enables us to provide a mapping between the observable random variables z_t , and the latent variables ζ_t .¹⁰ Once A_1, A_2 are estimated, these are used to form the forward expectations $E_t(\tilde{z}_{t+1}) = S[A_1\zeta_t + A_2\zeta_{t-1}]$, where S is a selection matrix, selecting \tilde{z}_t from ζ_t in the structural equations.

5.2 Varieties of Estimators

When all variables were $I(0)$ limited information estimation involved finding an expression for $E_t(z_{t+1})$, using this to eliminate the variable from the structural equation, and then estimating a system in which the structural equations are completed with an auxiliary system describing the endogenous variables in them. The latter will now be the latent VAR of the proceeding sub-section. There are then a number of possibilities for how to proceed. One possibility is to get estimates \tilde{A}_1, \tilde{A}_2 from the MLE applied to the latent variable VAR, and to then use these in place of A_1, A_2 . This would be the equivalent of 2SLS. However, one could also just use \tilde{A}_1, \tilde{A}_2 as starting values, and then estimate A_1, A_2 simultaneously along with the structural equation parameters. This would be the equivalent of LIML and is denoted as LI later. FIML estimation can also be thought of in this way, as one simply needs to estimate the model parameters using the *complete* latent variable VAR, and this is the way estimation is done in DYNARE. Of course, in the FI case one does not need to do a prior computation of $E_t z_{t+1}$ as this would be done numerically in a program like DYNARE.

⁹ This derivation follows Kapetanios, Pagan and Scott (2007).

¹⁰ In fact, the observation equation actually used is in terms of growth rates, i.e. $\Delta \tilde{z}_t = \Delta z_t - \Delta z_t^p$.

The FI and LI estimates of the structural parameters can be compared via a Hausman test, since the LI is consistent under the null that the complete model is correctly specified, but it is not fully efficient. For model evaluation purposes, it is also sometimes useful to set the auxiliary equation parameters A_1, A_2 etc. to their FI values and produce an LI estimator that is conditional on these. We will refer to that as the LI/FI estimator. An alternative way of using the same information is to compute $E_t z_{t+1}$ with the LI and FI estimates. This would be an encompassing test viewed through the lens of expectations, and is often a useful visual device for assessing models.

5.3 The VARX Form of the LS system

We now look at implementing the above ideas for Lubik and Schorfheide's (2007) model (9)-(16). In order to encompass the possibility that the LS model is misspecified, we need to use a VARX that is of sufficient order to incorporate rival models. This VARX will be used to generate $E_t z_{t+1}$ and to be the auxiliary system for any endogenous variables that need to be instrumented. If the structural equations have a hybrid form, then we will need to entertain a second order VARX process of the form

$$\tilde{z}_t = A_1 \tilde{z}_{t-1} + A_2 \tilde{z}_{t-2} + A_3 \Delta q_t + A_4 \Delta q_{t-1} + u_t, \quad (18)$$

where $\tilde{z}_t = [\tilde{y}_t \quad \pi_t \quad \Delta e_t - \pi_t \quad R_t]'$.

There are four unobservable shocks in (18) – the same as in LS's model. Since one of the shocks is technology, and it is viewed as permanent, and the source of the $I(1)$ nature of output, it will be assumed to serve that role in whatever model generated the data. In contrast, the other three shocks (18) are unnamed – of course, they would be combinations of the structural shocks in the correct model. Although, it is reasonable to assume that the shocks u_t are serially uncorrelated, due to the assumption that the chosen VARX is of high enough 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 in the following way from three e_{jt} that are *n.i.d.*(0, 1) plus the technology shock ε_t^a

$$\begin{bmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \\ u_{4t} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \gamma_{21} & 1 & 0 & 0 \\ \gamma_{31} & \gamma_{32} & 1 & 0 \\ \delta_1 & \delta_2 & \delta_3 & 1 \end{bmatrix} \begin{bmatrix} e_{1t} \\ e_{2t} \\ e_{3t} \\ \varepsilon_t^a \end{bmatrix}. \quad (19)$$

This structure allows for a general covariance matrix for u_t .

Model (18) is a latent VARX process as \tilde{y}_t is unobserved, with the observed data being $z_t' = [\Delta y_t \quad \pi_t \quad \Delta e_t - \pi_t \quad R_t]$. We relate the observable and latent variables through

$$\Delta y_t = \Delta \tilde{y}_t + \Delta y_t^p = \Delta \tilde{y}_t + \frac{\varepsilon_t^a}{1 - \rho_a}. \quad (20)$$

The system (18) and (20) is then estimated by applying MLE.¹¹ The parameters to be estimated are $A_1, A_2, A_3, A_4, \gamma_{21}, \gamma_{31}, \gamma_{32}, \delta_1, \delta_2, \delta_3, \rho_a, \sigma_a^2$, and $\sigma_j^2 = \text{var}(e_{jt})$ for $j = 1, 2, 3$. Note that Δq_t is observable and exogenous, so that the parameters ρ_q and σ_q in (13) can be estimated by a least squares regression.¹²

5.4 Estimation, and Evaluation of the Phillips Curve

We will look at estimation of the Phillips curve (9) on UK data provided by LS. The DSGE model parameters present in the equation are $\alpha, \beta, \tau, \kappa, \rho_{y^*}, \sigma_{y^*}, \rho_q, \rho_a, \sigma_q$, and σ_a . Some of these are fixed (β, σ_q), and others – such as ρ_a, σ_a – come from either estimating the latent VARX or setting them to the FI estimates. Basically, interest is in the model parameters α, κ , and τ .

We estimate these parameters using the FI, LI and LI/FI estimators described earlier. Since LS used Bayesian methods in constructing their FI estimates, we combine the likelihoods of the LI and LI/FI estimators with LS's priors in order to get comparable estimates. The LI estimation was performed by taking the LS Phillips curve, and then augmenting it with equations for $E_t \pi_{t+1}, \tilde{y}_t, \Delta e_t$, and R_t . The coefficients of the auxiliary system are either jointly estimated with the structural equation parameters, as with the LI estimator, or fixed at FI values, in the case of the LI/FI estimator.¹³ In the LI case the auxiliary equation parameters are estimated with priors set to a normal density and quite wide standard deviations.

¹¹ DYNARE version 3.065 was used to get the estimates.

¹² 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. For more details, we refer the reader to Fukač and Pagan (2006).

¹³ Since there are two shocks ε_t^a and $\varepsilon_t^{y^*}$ in the structural equation we construct the shocks for the auxiliary model from $\varepsilon_t^{y^*}$ and two other white noise errors. In order to allow for correlation between the auxiliary equation shocks and the structural equation shock we use the Cholesky type decomposition that was adopted for u_t previously. Because of the way DYNARE works, equations describing the evolution of the exogenous variables $\Delta q_t, \Delta a_t$, and \tilde{y}_t^* are needed as well.

Table 1
Estimates of the Phillips Curve on UK data

	α	κ	τ
LS	0.14 (0.09-0.19)	1.78 (1.17-2.39)	0.46 (0.33-0.61)
LI/FI	0.11 (0.07-0.15)	1.19 (0.78-1.62)	0.38 (0.29-0.46)
LI	0.17 (0.10-0.22)	1.26 (0.75-1.77)	0.60 (0.26-0.90)

Note: (LS) Lubik and Schorfheide's (2007) FI estimates. The numbers in italics are 90% confidence intervals.

Table 1 presents means of the posteriors for the three estimators, along with 90% confidence intervals.

There are differences between the estimates of κ and τ for the estimators. The LI/FI and LI estimates lie outside the 90% confidence intervals that LS show ¹⁴ To assess the overall impact of the changes however it is best to compute the impact of the output gap upon inflation as this involves a combination of α , κ and τ . That effect is estimated to be .91 (LS), .80(LI/FI) and .36 (LI). Thus there is a very large decline in the estimated effect of the output gap upon inflation, and the LI estimates seem much closer to what is normally found in the literature. It seems therefore as if the assumptions made in estimating the rest of the model have had an enormous impact upon the estimates of the parameters, 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. The origin of the bias here must come from the fact that the FI estimates of the auxiliary system parameters are very different to the LI ones, as that difference influences the estimate of $E_t \pi_{t+1}$ and, as we observed earlier, one needs a good estimate of $E_t \pi_{t+1}$ in order to avoid biases.

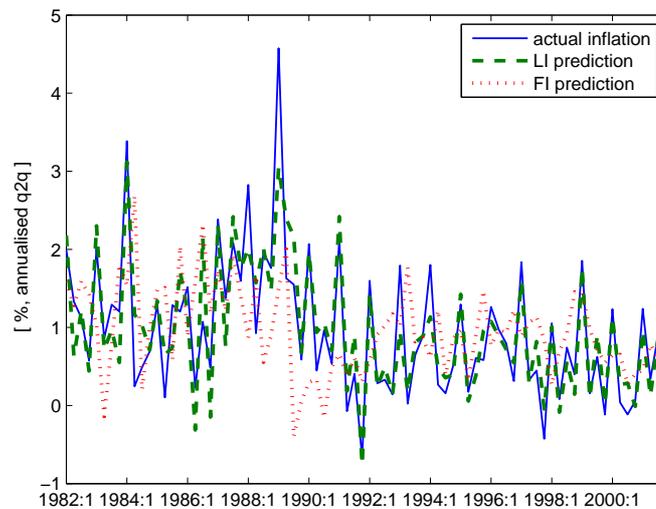
A possible specification error in the LS system is in the Phillips curve itself. It may be that it should have a hybrid structure. To test this $\beta E_t(\pi_{t+1})$ is replaced by $\beta[(1-c)E_t(\pi_{t+1}) + c\pi_{t-1}]$ and re-estimated with the limited information estimator. The estimates of α , τ and κ are very close to the LI estimates of Table 1, and there is clear evidence of the need for a hybrid structure, since the posterior mean of c is 0.41 with confidence interval 0.13-0.70.

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.

5.5 An Evaluation of the FI Estimated Model

We noted earlier that there was probably a substantial difference between the FI and LI estimates of the auxiliary system parameters. A simple way of capturing this difference is to compare the predictions made of inflation using the different values. Because the LI estimates are from an unconstrained VARX process, it is not surprising that $E_t \pi_{t+1}$ and π_{t+1} are close. But, as seen in Figure 1, this is far from true of the $E_t \pi_{t+1}$ constructed with the FI estimates. That figure suggests that LS's estimated model does not produce expectations of the inflation rate that are rational. To formally test this, we regress the expectation error when using the FI values, $\pi_{t+1} - E_t \pi_{t+1}$, against π_t , and find a coefficient of .4 with t ratio of 5.6. This is the same as implementing a parametric encompassing test of the type described earlier.

Figure 1
FI and LI based inflation expectations, and actual inflation



6 Conclusion

We have advanced the proposal that DSGE models should not be estimated and evaluated only with full information methods. These make the assumption that the complete model is correct, and therefore do not recognize the uncertainty that might exist about individual equations within it. Some limited information analysis which does recognize this uncertainty can provide useful complementary information about the adequacy of the model equations in matching the data. 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 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 (2007), and the differences between the LI and FI parameter estimates suggests that this model has substantial specification errors.

References

- Binder, M and M H Pesaran (1995), “Multivariate rational expectations models and macroeconomic modelling: A review and some new results,” in *Handbook of Applied Econometrics*, eds M Pesaran and M Wickens, vol I: Macroeconomics, 139–187, Blackwell Publishing.
- Brundy, J M and D W Jorgenson (1971), “Efficient estimation of simultaneous equations by instrumental variables,” *Review of Economics and Statistics*, 53, 207–234.
- Canova, F, M Finn, and A Pagan (1994), “Evaluating a real business cycle model,” in *Non-Stationary Time Series Analysis and Co-Integration*, ed C Hargreaves, 225–255, Oxford University Press.
- Fuhrer, J C and G Olivei (2005), “Estimating forward-looking Euler equations with GMM and maximum likelihood estimators: An optimal instruments approach,” *Proceedings*, 87–114, Federal Reserve Bank Board of Governors.
- Fukač, M and A R Pagan (2006), “Limited information estimation and evaluation of DSGE models,” *NCER Working Paper*, (6).

- Gourieroux, C, A Monfort, and E Renault (1993), "Indirect inference," *Journal of Applied Econometrics*, 8, 85–118.
- Harvey, A C and A Jaeger (1993), "De-trending, stylized facts and the business cycle," *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 of Banco de Espana*, (0208).
- Kapetanios, G, A Pagan, and A Scott (2007), "Making a match: Combining theory and evidence in policy-oriented macroeconomic modeling," *Journal of Econometrics*, 136, 505–594.
- Korenok, O, S Radchenko, and N R Swanson (2008), "International evidence on the efficacy of new-Keynesian models of inflation persistence," *Journal of Applied Econometrics*, forthcoming.
- Kurmann, A (2007), "VAR-based estimation of Euler equations with an application to New Keynesian pricing," *Journal 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 (1979), "Some consequences of viewing LIML as an iterated aitken estimator," *Economics Letters*, 3, 369–372.
- Sbordone, A (2006), "U.S. wage and price dynamics: A limited information approach," *International Journal of Central Banking*, 2, 155–191.
- 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.