Observations on DeCecio and Nelson "An Estimated DSGE Model for the United Kingdom"*

Martin Fukac† and Adrian Pagan‡

January 11, 2007

The paper by DeCecio and Nelson (DCN) considers the estimation of the parameters of a DSGE model for the U.K. that is virtually the same as that which Christiano, Eichenbaum and Evans (CEE) estimated for the US. The CEE model is much larger than existing academic DSGE models of the U.K. such as Lubik and Schorfheide (2005). It is not as large as the BEQM model of the Bank of England which has both DSGE elements and data imposed dynamics, although, as the latter is to be used for policy, there is a much greater imperative to match the data than found in most academic work. There are a few other DSGE models that have been applied to the UK e.g. Leitemo (2006) but, in general, these are often used to examine some particular question and are also rather restricted in their mode of operation. Often they use a standard open-economy New Keynesian model rather than a straight DSGE model like CEE. Moreover, the authors are often not that familiar with the UK context and data, whereas the current authors are experts in this area, and it certainly shows in their discussion of alternative data sources. So, given the paucity of studies, any new one would be welcome.

Now as the Chinese proverb says a journey of a thousand miles starts with a single step. What we have here is more than single step but well short of a thousand miles. Reading it one longed for a fully fleshed-out model along the lines of Smets and Wouters' work on the US and the Euro area (which is very

*Research supported by ARC Grant DP0449659.

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

‡Queensland University of Technology

similar to CEE) where a complete set of shocks is described and estimated. The fact that only a money shock is identified here means that there are few questions one can ask about the model. So it was disappointing that the authors were not a bit more adventuresome. But we presume that this will be part of a broader piece of research and one looks forward to seeing a more complex model at a later stage. At that time one would like to see some recognition of the open economy characteristics of the U.K. Of course one does have to acknowledge that DSGE models have not had a good record at producing useful models of the open economy. One reason that the authors point to is the prediction of stronger exchange effects than seen in the data. We agree with this and it was a central conclusion about the mini-BEQM model that was calibrated to the U.K. economy in Kapetanios et al.(2007). Moreover, as Justiniano and Preston (2006) argue, it is been vary hard to find much of an influence of the foreign economy upon a small open economy and this is contrary to what evidence we have from VARs. So there is quite a bit to be done both on the broader front of developing useful open economy models and in getting a UK model that is in a more complete state than this one. Because the model is not fleshed out that much we will restrict comments to what DCN do rather than alternatives that might have been tried.

1 Estimation Strategy

The methodology used for estimation is that of CEE. It has four steps

1. Identify monetary policy shocks using a structural equation for the interest rate and a VAR(2) to represent the rest of the system. The identification condition used is that monetary shocks have no contemporaneous effect on any variables of the system except the interest rate. The monetary policy rule depends on all variables in the VAR and these enter the rule both with lags and contemporaneously.

2. Compute the monetary impulse responses C_j^D from this structural VAR (SVAR) (j indexes the $j'th$ impulse response).

3. Choose some parameter values for the DSGE model parameters θ and use them to compute the DSGE model monetary impulse responses C_j^M.

4. Find the value of θ that minimizes $\sum_{j=1}^{M} (C_j^D - C_j^M)'W(C_j^D - C_j^M)$, where W is a diagonal matrix of weights.

They apply this to UK data. The original DSGE model they employ has 15 variables while the SVAR(2) has 6. Estimates of the parameters are presented and some standard errors are given along with plots of the monetary impulses implied by the SVAR and the DSGE model calibrated with the estimated θ.

2 Estimation problems

What could go wrong with this methodology? There are three issues which we discuss in sub-sections below.

2.1 How many impulses to use?

There is a maximum useful choice for M since the C_j^D are simply functions of the SVAR coefficients. Let there be n variables in the SVAR (in DCN $n = 6$ and it is an SVAR (2)). Then the total number of coefficients in the DCN SVAR(2) is 77 - 72 from the 2 lags of the six variables in the six equations plus the five possible coefficients attached to contemporaneous coefficients in the interest rate rule. Since M seems to be 25, that would mean that 150 impulses are used. This is much larger than the number of parameters determining them. Hence there are many redundant impulse responses and the covariance matrix of $C_1^D, ..., C_{25}^D$ must be singular. This might be a problem when one uses the $\delta-$method to compute standard errors. Indeed, the standard errors of $\hat{\theta}$ found by moment matching in DCN seem to be incredibly small. Thus they have an estimate of the markup λ_f parameter of 2.09 with a standard error of .0078. It's hard to believe that one could ever get that degree of precision with just 26 years years of quarterly observations. Since, in only one case (lagged productivity), do the SVAR coefficients have a t ratio above 5 its hard to see how we can end up with t ratios above 400 for $\hat{\theta}$, which are fundamentally derived from the SVAR coefficients.

2.2 Approximating the DSGE Model with a SVAR

There is a generic problem here in that the DSGE model often determines m variables and $m > n$ i.e. the SVAR is fitted to a smaller number of variables than appear in the DSGE model. This is true of DNC where it appears that $n = 6, m = 15$. Now the DSGE model will be an SVAR in the m variables but is unlikely to be an SVAR in the n variables. An old literature, due to Zellner and Palm (1974) and Wallis (1977), has noted that, when a system which is a VAR(p) in n variables is reduced to a smaller system with $m < n$ variables, the smaller system will generally be a VARMA process. Since the CEE procedure involves such compression of variables it might be expected that a VARMA process is needed rather than a VAR, and so the use of a VAR could lead to specification bias. It might be expected that a very high-order VAR could compensate for this mis-specification. This is generally true but the order of VAR needed to deliver a good approximation may in fact be far too high for the data sets one is normally faced with e.g. Kapetanios et al (2007) find that reducing a model that is BEQM-like (but half the size) would require a VAR(50) to capture the effects of some shocks (and this with 30,000 observations). The problem has been analysed in a DSGE context by Ravenna (2006) and Fry and Pagan (2005). We adopt the exposition of the latter.

Suppose that the DSGE model followed a VAR(1) solution (assuming that u_t is *i.i.d.*)

$$z_t = Pz_{t-1} + Gu_t.$$

Now consider what happens if we model only a subset of the variables. We will call the modelled sub-set z_{1t} and the omitted variables z_{2t}. We can decompose the VAR above as

$$z_{1t} = P_{11}z_{1t-1} + P_{12}z_{2t-1} + G_1u_t \tag{1}$$

and we will assume that the following relation holds between z_{1t} and z_{2t}

$$z_{2t} = \bar{D}_0 z_{1t} + \overline{D}_1 z_{2t-1} + \overline{D}_2 u_t.$$

Substituting this in we get

$$z_{1t} = (P_{11} + P_{12}\bar{D}_0)z_{1t-1} + P_{12}\overline{D}_1 z_{2t-2} + G_1u_t + P_{12}\overline{D}_2 u_{t-1}$$

so that the sufficient conditions for there to be a finite order VAR in z_{1t} will be that either $P_{12} = 0$ i.e. z_{2t} does not Granger cause z_{1t}- see Lütkepohl (1993 p 55) and Quenouille (1957, p. 43-44) or $\overline{D}_1 = 0, \overline{D}_2 = 0$ i e. the variables to be eliminated must be connected to the retained variables through an identity and there can be no "own lag" in the omitted variables in the relation connecting z_{1t} and z_{2t}.

This observation looks trivial but, in fact, it happens that many of the problems that have arisen where a finite order VAR does not obtain come from the fact that the variables which are omitted are connected with the retained variables through an identity, but one that contains an "own lag". The classic example is in the basic RBC model where, after log linearization around the steady state, we would get

$$\begin{aligned} l_t &= y_t - c_t \qquad (2) \\ C^* c_t + K^* k_t &= Y^* y_t + (1-\delta) K^* k_{t-1} \qquad (3) \\ c_t &= E_t(c_{t+1} - \alpha\gamma(y_{t+1} - k_{t+1})) \\ y_t &= a_t + \alpha k_{t-1} + (1-\alpha) l_t, \end{aligned}$$

where c_t is the log of consumption, a_t is the log of the technology shock, k_t is the log of the capital stock, l_t is the log of labour input, and y_t is the log of output. An asterisk denotes steady state values and α is the steady state share of capital in output. When a_t is an $AR(1)$ the solution to this system can be made a VAR(1) in c_t, l_t, y_t and k_t. It's clear that we could eliminate any of c_t, l_t or y_t since these do not appear as a lagged variable in the system. Equally clearly k_t cannot be eliminated unless we can find an identity relating it to other variables that does not involve k_{t-1}. Thus the identity (3) shows that this is not possible. Most of the literature that seeks to establish that a SVAR cannot approximate a DSGE model - Chari et al (2004), Erceg et al (2005), Cooley and Dwyer (1995) - substitute out k_t, and so end up with a non-finite order VAR.

The implication of this for DCN's work is that the reduction of the system from 15 to 6 variables might necessitate a very long VAR and not the VAR(2) they adopted. They used statistical criteria to determine the order of the VAR. Kapetanios et al did this as well, and the tests produced a VAR of order six, far below what was needed (50'th order) to produce the correct impulse responses. The reason is that the tests proceed on the assumption that the number of variables in the VAR is correct and it is only the order that needs to be found. So seemed that DCN might be matching impulses that are not strictly comparable. The appropriate procedure would be to simulate a long history of data from the 15-variable DSGE model incorporating just monetary shocks, fit a VAR(2) in just 6 of the variables, and then find the impulse responses from such an approximating VAR (one has to be careful to note that some of the lagged values will be perfectly correlated with others and one will need to combine variables together to overcome that problem). These are then matched up with the empirically observed VAR(2) impulse responses in the six variables. We have assessed this by examining a variant of their model where information is dated at t rather than the combination of t and $t-1$ that is in their paper. However, we used the same model parameters as DCN. Although there are some differences between the true impulse responses and those delivered by a VAR(2), it seems that the approximations are quite good, except at longer horizons. So there does seem to be an approximation issue here, although in any application one should check that there is no problem, as it is not very difficult to do. One should note however that the δ method used by CEE to compute standard errors is only correct if the approximation is satisfactory. Basically the estimator of the DSGE model parameters is an indirect estimator, being derived from functions of the SVAR coefficients represented by the impulse responses. The covariance matrix of such an estimator requires one to compute derivatives of the model-implied VAR impulse responses with respect to the θ parameters, and not the derivatives with respect to the model impulses, as done by CEE. These are only the same if there is no approximation error.

2.3 Multiple Solutions

Ignoring the problem identified in the previous section, estimators such as MLE basically attempt to match the VAR coefficients from the data with those from the model, rather than attempting to match impulse responses. To see the problems you might encounter with the latter let us look at the simple model

$$\begin{aligned} y_t &= \beta E_{t-1}(x_{t+1}) + \varepsilon_{yt}, \\ x_t &= \rho x_{t-1} + \varepsilon_{xt} \end{aligned}$$

The VAR will be

$$\begin{aligned} y_t &= a_1 x_{t-1} + \varepsilon_{yt} \\ x_t &= a_2 x_{t-1} + \varepsilon_{xt}, \end{aligned}$$

where $a_1 = \beta\rho^2, a_2 = \rho$, and we have the impulse responses as

$$c^M_{1,y\varepsilon_x} = \beta\rho^2, c^M_{2,y\varepsilon_x} = \beta\rho^3, c_{1,x\varepsilon_x} = \rho, c_{2,x\varepsilon_x} = \rho^2$$

If we try to find β and ρ by match the first two impulse responses we would be minimizing (assuming that the weights in W are equal)

$$\begin{aligned}&(c^D_{1,y\varepsilon_x} - \beta\rho^2)^2 + (c^D_{2,y\varepsilon_x} - \beta\rho^3)^2 + \\ &(c^D_{1,x\varepsilon_x} - \rho)^2 + (c^D_{2,x\varepsilon_x} - \rho^2)^2\end{aligned}$$

It's clear that this has the problem that it produces a sixth order polynomial in ρ so that we may get multiple solutions. This wouldn't arise if we were matching to the VAR coefficients since then $\hat{\rho} = \hat{a}_2, \hat{\beta} = \hat{a}_1/\hat{\rho}^2$. Bearing in mind the first point as well it seems better to match the VAR to get estimates of θ and then to show the impulse response correspondence.

3 Looking at At Some of the Euler Equations

Now it would seem useful to develop a method that uses the same information as impulse response matching but which is a bit simpler, provides ready ways of computing standard errors, emphasizes the economics, and which can be used to tell us something about the ability of the DSGE model to match the data. Basically the proposal is to work with the Euler equations and to estimate the model parameters directly from them with a single-equation estimator. Of course this is an old idea, but it has fallen out of favour, possibly because of the literature claiming that systems estimators of parameters of the New Keynesian system performed better than the single equation ones due to weak instruments. But, in many DSGE models, enough parameters are prescribed that weak instrument problems are not present, and we will see this in the DCN context.

The Euler equations of DSGE models have the generic form (the dating of expectations here comes from DCN and reflects the assumptions about interest rates having no effect on contemporaneous variables)

$$E_{t-1}z_t = \eta_1 z_{t-1} + \eta_2 E_{t-1}(z_{t+1}) + \eta_3 E_{t-1}w_t.$$

In this equation z_t is the endogenous variable whose coefficient is normalized to unity, w_t are either exogenous or other endogenous variables and the parameters η_j are functions of some of the DSGE model parameters θ. Now this can be written as

$$z_t = \eta_1 z_{t-1} + \eta_2 E_{t-1}(z_{t+1}) + \eta_3 E_{t-1}w_t + \varepsilon_t,$$

and the RHS regressors are uncorrelated with the error ε_t. If we had these conditional expectations we could run a regression. We note that this equation holds for any sub-set of information used by the economic agents. Hence let us define the information used in the estimation as that of the DCN VAR(2) i.e. two lagged values of $y_t, c_t, i_t, y_t - h_t, r_t$ and Δp_t[1]. Call these the vector ζ_{t-1}. Then, if we can estimate $E_{t-1}(z_{t+1})$ and $E_{t-1}w_t$, we could simply fit a regression to this equation and thereby measure η_j. Since the model is linear we can indeed estimate $E_{t-1}(z_{t+1})$ and $E_{t-1}w_t$ as the predictions from the regression of z_{t+1} and w_t against ζ_{t-1}. Basically this estimation method uses the same information as moment matching i.e. the information contained in the VAR. Notice that standard errors are easily found from this by treating it as an IV estimator with z_{t+1} and w_t as variables that need to be instrumented. As we will see later, in most cases the instruments are very good, and so there is no reason to doubt the standard errors of η_j found in this way.

This is a relatively simple way to estimate the η_j. Whether one can estimate the DSGE model parameters θ is a different question as there may be a non-linear mapping between the η and θ and so we may not be able to recover θ uniquely. This shouldn't concern us unduly since fundamentally the impact of monetary policy depends upon the η_j but there may be some cases where we want to think about changing θ and so we would then need to identify it. Ma (2002) pointed out that there was an identification problem like this in strictly forward-looking New Keynesian Phillips curves and we will see that it comes up in the CEE model as well.

Let us look at the above principles in the context of some of the equations in DCN. First we look at the Phillips curve. After normalizing on π_t the Euler equation becomes

$$-E_{t-1}\pi_t + \frac{1}{1+\beta}\pi_{t-1} + \frac{\beta}{1+\beta}E_{t-1}\pi_{t+1} + \frac{(1-\beta\xi)(1-\xi)}{(1+\beta)\xi}E_{t-1}s_t \qquad (4)$$

[1]We work with data that is not deviations from steady state values so will have to include intercepts in any equations we estimate

We can write this as an equation of the form

$$\begin{aligned}\pi_t &= \frac{1}{1+\beta}\pi_{t-1} + \frac{\beta}{1+\beta}\pi_{t+1} + \\ &\quad \frac{(1-\beta\xi)(1-\xi)}{(1+\beta)\xi}E_{t-1}[(\alpha r_t^k + (1-\alpha)(R_t + w_t))] + \varepsilon_t\end{aligned}$$

or

$$\pi_t - \frac{1}{1+\beta}\pi_{t-1} - \frac{\beta}{1+\beta}\pi_{t+1} = \eta_1 E_{t-1}[(\alpha r_t^k + (1-\alpha)(R_t + w_t))] + \varepsilon_t,$$

where $E_{t-1}(\varepsilon_t) = 0$. We note that, since $\beta = .99$ is imposed, we are not trying to estimate the coefficients attached to π_t and π_{t+1}. Now using data one can form $\alpha r_t^k + (1-\alpha)(R_t + w_t)$ as DCN pre-set α to .36, and then regress this against the information in the VAR lagged variables to get $E_{t-1}[(\alpha r_t^k + (1-\alpha)(R_t + w_t))]$.[2] The regression of this variable against ζ_{t-1} (the VAR(2) lagged variables) gives an R^2 of .99 so it is a very good instrument for $\alpha r_t^k + (1-\alpha)(R_t + w_t)$. If we fit a non-linear regression to this equation we get an estimated coefficient for ξ_p of .988, which is reasonably close to the .94 reported in the paper from impulse response matching. But the standard deviation is .092, which is nowhere near the .0003 given in the paper, although if one makes it robust to serial correlation it halves. Clearly the estimate here implies very low frequency of price adjustment, as does DCN's. Whilst this estimate seems implausible the interpretation would seem to be that there are some problems in the specification of the Phillips curve (indeed the serial correlation in the residuals is consistent with that).

Another equation in DCN we could look at estimating would be the production function which has the form

$$y_t = \lambda_f(\alpha k_t + (1-\alpha)l_t)$$

Given DCN prescribe α we can form $\zeta_t = \alpha k_t + (1-\alpha)l_t$ and treat this as a regressor to estimate λ_f. There will of course be technology in this relation

[2] Since DCN conclude that $\sigma_a = \infty$ there is no difference between the capital stock and services. We compute the capital stock recursively but this means the estimates are inaccurate until the initial condition disappears. Since we start the recursion in 1955/2, but only use data after 1979/2, we feel that the effects of the initial condition will have died away, as it will be multiplied by the term $(.975)^{104}$.

and, by treating it as an AR(1) process, the equation will have an AR(1) error term. Because the regressor will generally be correlated with the white noise shock driving the AR in technology we need instruments to estimate λ_f and for this we use y_{t-1} and ζ_{t-1}. We also include a constant to reflect the fact that we are not using variables that are deviations from a constant steady state and that technology should have a constant mean. Then we get an estimate of λ_f of 1.29 with a standard error of .05. This seems more reasonable than the value of 2.0982 that they obtain, although they give a defence of it. Again the standard errors are very different.

The interest rate rule parameter values are somewhat puzzling. Under the assumptions in force here one should be able to fit this rule by OLS regression as it is assumed that the regressors are all uncorrelated with the interest rate shock. If we run the regression they fit we would get

$$R_t = .883R_{t-1} + (1 - .883)(.0001y_t + 1.28\pi_{t-1}) + .05\Delta y_t + .10\Delta\pi_t$$

versus the estimated equation of the paper

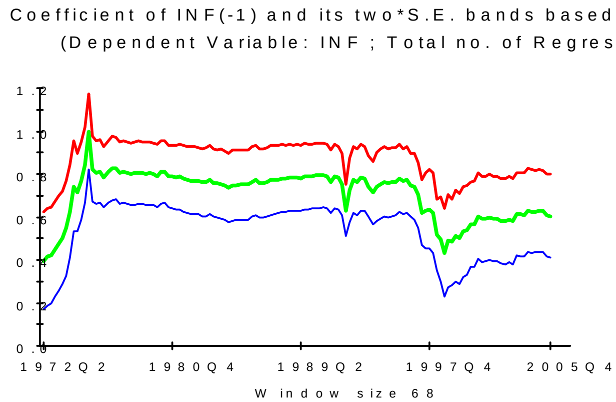
$$R_t = .872R_{t-1} + .(1 - .872)(.348y_t + 1.27\pi_{t-1}) + .43\Delta y_t - .62\Delta\pi_t$$

The standard deviation on y_t from the OLS regression is very small so these estimates are very different. DCN note that the rule they fit is not the one in the VAR(2) as that would include other lags in the variables. But if we just fit a VAR(1) then it should be comparable to what they claim the estimated money rule is. In fact there is not much difference if we add on extra lags. Notice also that the negative sign on $\Delta\pi_t$ that perturbed them has gone. Since this seems a logical way to estimate the money rule given the assumptions made about the structure of the model one is puzzled about the results that come from impulse response matching.

What explains this? One possibility is that the DSGE model implies a particular value for the intercept of the equation, whereas we have just subsumed this into a constant term that is freely estimated. However, the steady state values used in the model for variables seemed quite close to the sample means over the estimation horizon so it would seem that one would get much the same intercepts (provided of course the slope coefficients were correct).

There are some problems with multiple parameter values in both the Phillips curve and the wage equation. Because

$$\pi_t - \frac{1}{1+\beta}\pi_{t-1} - \frac{\beta}{1+\beta}\pi_{t+1} = \eta_1 E_{t-1}[(\alpha r_t^k + (1-\alpha)(R_t + w_t))] + \varepsilon_t$$

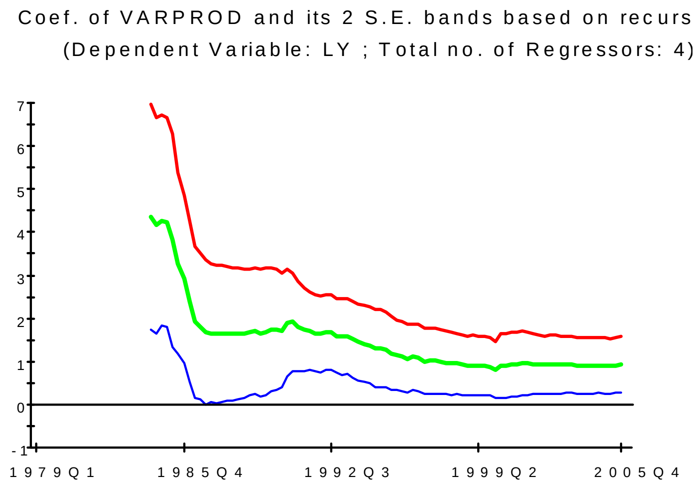


and $\eta_1 = \frac{(1-\beta\xi_p)(1-\xi_p)}{(1+\beta)\xi_p}$ we see that the solution for ξ_p involves a quadratic. There are two estimates of ξ_p which produce exactly the same likelihood. The other one is 1.02. A similar situation exists for the wage equation. Perhaps this is one reason why Bayesian methods might work better in these models as they would impose the restriction that ξ_p and ξ_w lie between 0 and 1.

4 Structural Change in the Model

The authors look at structural change in the SVAR and conclude that there was some back in the 1970s but this was due to industrial issues and not monetary policy regime changes. But it's always difficult learning something about the stability of all the parameters in a VAR. One might also want to ask where one wants to place a possible monetary policy regime change. Is it when Thatcher came in, when inflation targeting was adopted, or when there was a formal change to the institution with the formation of the MPC? In Pagan (2003) it was argued that there had been a change in the level of persistence in inflation in the UK after the formation of the MPC. This is still evident in the data- see figure 1 which gives estimates of the coefficient of π_{t-1} using a rolling horizon of 68 quarters.

So this looks like structural change in the dynamics and perhaps the VAR stability tests should have focussed more around that the post 1997 point, although this means a very short post-break sample. At the end of



the day graphs like this have to make one wonder about applying a constant parameter DSGE model to such data. .It would seem as if one might have to just use the post 1997 period to estimate the DSGE model, although with such small samples one might need to use some sort of Bayesian approach. Perhaps one could use the estimated values of this paper to produce priors.

The issue of structural change prompts one to go back to the production function and look at the estimate of λ_f recursively. This is done in figure 2, which shows that there have been quite large changes, although the relatively small sample makes it hard to be sure of this. It's interesting that early on the estimate was much closer to that found by DCN.

5 References

Chari, V.V., P.J. Kehoe and E. R. McGrattan (2004), "A Critique of Structural VARs Using Real Business Cycle Theory", *Working Paper no 631 Federal Reserve Bank of Minneapolis.*

Cooley, T.F. and M. Dwyer (1995), "Business Cycles Without Much Theory: A Look at Structural VARs", *Journal of Econometrics*, 83, 57-88.

Erceg, C.J., L. Guerrieri and C. Gust (2005), "Can Long-Run Restrictions Identify Technology Shocks?", *Journal of the European Economic Association*, vol. 3, pp. 1237-1278.

R. Fry and A.R. Pagan (2005), "Some Issues in Using VARs for Macroeconometric Research", *CAMA Working Paper 2005/18* , Australian National University

Justiniano, A. and B. Preston (2006), "Can Structural Small Open Economy Models Account For the Influence of Foreign Disturbances", *CAMA Working Paper, 12/2006.*

Kapetanios, G., A.R. Pagan and A. Scott (2007), "Making a Match: Combining Theory and Evidence in Policy-oriented Macroeconomic Modeling", *Journal of Econometrics*, 136, 505-594.

Leitemo, K. (2006), "Targeting Inflation by Forecast Feedback Rules in Small Open Economies", *Journal of Economic Dynamics and Control*, 30, 393-413.

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

Lütkepohl, H. (1993), *An Introduction to Multiple Time Series* (Springer-Verlag, Berlin)

Ma, A. (2002), " GMM Estimation of the New Phillips Curve", *Economics Letters*, 76, 411-417

Pagan, A.R. (2003) "Report on Modelling and Forecasting at the Bank of England". *Bank of England Quarterly Bulletin*, Spring, 1-29

Quenouille, M.H. (1957), *The Analysis of Multiple Time Series,* Griffin's Statistical Monograohs and Course No 1 (Griffin, London)

Ravenna, F. (2006) "Vector Autoregressions and Reduced Form Representations of Dynamic Stochastic General Equilibrium Models" *Journal of Monetary Economics* (forthcoming)

Wallis, K.F. (1977), "Multiple Time Series and the Final Form of Econometric Models", *Econometrica*, 45, 1481-1497.

Zellner, A. and F. Palm (1974), "Time Series Analysis and Simultaneous Equation Econometric Models", *Journal of Econometrics*, 2, 17-54.