An Examination of Some Tools for Macro-Econometric Model Building

Adrian Pagan*
Australian National University and the University of New South Wales

July 28, 2003

Contents

1 Introduction **2**

2 Modelling Strategies at the Bottom End of the Frontier **4**
2.1 VARs, ECMs and their Problems 4

2.2 The Utility of Non-linear Models 12
2.2.1 A Non-linear Statistical Model of the Business Cycle . 13

3 Moving up the Frontier **18**
3.1 Re-expressing the VAR as an Economic Model 19
3.1.1 The Uncorrelated Shocks Assumption 20
3.1.2 What do Variance Decompositions Tell us about Cycles? 21
3.2 The View from a Position Towards the Top of the Frontier . . 24
3.2.1 Looking Downwards 25
3.2.2 Looking Upwards 28

4 References **30**

*METU Lecture, ERC Conference VII, Ankara, September 2003. An earlier version was presented as an invited lecture to the Eighth African Econometrics Conference, Stellenbosch, July 2003. This research was supported by ESRC Grant No. 000 23-0244

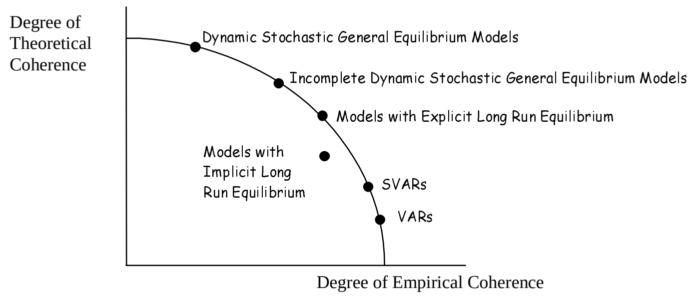

Figure 1:

1 Introduction

I want to start the lecture with a graph. Over the years I have found it useful to think about the activities engaged in by econometric researchers with this representation. It shows the frontier curve that connects up the degree to which our current modelling methods aim to exhibit coherence with the ideas of economic theories and the degree to which they attempt to cohere with the data. At each end of the frontier the coherence is perfect for one of these characteristics and zero for the other. Crudely speaking, we might say that economics has primacy for those modelling strategies located at the top left hand corner while statistics is dominant at the bottom right hand end. For this reason I will refer to the models at the top end as "economic" and those at the bottom as "statistical". Those along the frontier are hybrid models. Another way of expressing this is to say that, at the bottom we have models that simply *summarize* the data, and at the top we have models that aim to *interpret* the data. Along the curve we have work that attempts to trade off the two objectives. When we are inside the frontier it is possible to improve modelling on either one or both dimensions.

Many of the founders of econometrics tended to be people who were at the top left hand part of the curve e.g. Koopmans, Marshak, Frisch are probably names that we associate with pure economic theory today. There were people towards the bottom, of whom the best known was probably Tinbergen. Today, whilst it seems likely that there would be a consensus that we should be located somewhere on the curve between the two polar points, I think it is the case that much academic work in econometrics is closer to the bottom end i.e. often it is concerned with the construction of complex statistical models that are designed to fit increasingly elusive characteristics of the data. In macroeconometrics one sees this fact reflected in a concern about the dependence properties of the data and whether there is non-linear structure to the conditional moments. In microeconometrics we have seen an increasing emphasis placed upon non-parametric methods of estimation that aim to be robust to assumptions made about the density of the random variables that the data is thought to be realizations of.

At the very top of the curve are the models that guide much discussion in macroeconomics historically and today- the IS/LM model, the Fleming-Mundell model, the Real Business Cycle model and the New Keynesian policy model These miniature models are often used for considering the consequences of shocks to a stylized economy. Often they are simply devices to produce qualitative insights into the mechanisms at work and do not relate to any specific economy or body of data. But at other times attempts are made to make them more data coherent e.g. IS-LM in Gali (1992) and the New Keynesian Policy model as discussed by Allsopp and Vines (2000). Because they are miniature models they are often only vaguely designed to represent an actual economy. Nevertheless, they are used to motivate larger models that do have such an objective e.g. the G^3 and MSG2 models of the world economy - McKibbin and Vines (2000)- the QPS model of the Canadian economy - Coletti et al (1996)- and the FPS model of the New Zealand economy- Black et al (1997). Such models represent movements down the curve since there is some coherence with the data, although this may not be achieved in any conventional way i.e. it may not involve selecting parameters by minimizing some criterion which matches the model to the data. Instead, the model parameters are often selected ("calibrated") by presenting simulations to decision makers to see if the results are in accord with their priors concerning the way that the economy would respond to certain shocks. It is rarely the case that the statistical techniques familiar from the bottom of the curve are followed.[1]

Thus there is a gap between the top and the bottom of the curves and it seems important that we close it in some way. One way is to focus upon

[1] The G^3 model does use a data base to estimate many of its parameters and so is further down the curve than its predecssor MSG2.

expanding the miniature economic models so that they are confronted more directly with the data e.g. Smets and Wouters (2002). Another is to transform the statistical models to give them an economic interpretation e.g. the move from a VAR to a structural VAR. There is a good deal of current research that aims at moving us along the curve (and shifting it out), some of which I have mentioned in Pagan (2003). But, in this lecture I have a more limited objective, and that is to ask what lessons there are for statistical modelling from the work that has been done on models towards the top end of the curve and what are the implications for evaluating the statistical models given that we want to use them for answering some question often relating to policy issues. My contention is that the models at the left hand end of the curve should have something to say about the nature of those on the right (and conversely).

So I want to reflect on some aspects of econometric modelling by using this graph. I will start by looking intensively at the bottom right hand end from the perspective of the philosophies embedded in the top. Then I will take a first move up the curve, looking at the main suggestions concerning how this should be done, before finally jumping towards the top and briefly analyzing how one might move in both directions from there. These movements are facilitated by the use of the powerful computational facilities that we have available to us today. Because it is what I know most about I will concentrate upon macroeconometric modelling. I feel that many of the points are as true of microeconometrics, although I do not have the set of ready examples drawn from my own work to deal with the latter topic. There is no sense in which what I give is a comprehensive critique. Rather it is a series of vingettes that are drawn from my own experience.

2 Modelling Strategies at the Bottom End of the Frontier

2.1 VARs, ECMs and their Problems

Let me begin by providing a thumbnail sketch of what I see as the principal statistical models in use in macro-econometrics. We have been able to observe the most important characteristics of macro-economic and financial time series for many years. Graphs show that they tend to evolve relatively smoothly, often they co-evolve and they exhibit cycles, in that they are subject to rises and falls that, whilst not perfectly regular, are recurrent. What we have developed in the past twenty or thirty years are ways of talking about and describing these characteristics through parametric statistical models. In the class of linear models the most commonly used model has been the Vector Autoregressive (VAR) model applied to a vector of n series y_t. This takes the form

$$(I - A_1 L - ... - A_p L^p) y_t = A(L) y_t = v_t, \tag{1}$$

where v_t is an $n \times 1$ vector of shocks and L is a lag operator. The model can be thought of as the equivalent of the reduced form in the simultaneous equations literature. It is embellished in many ways. In particular it is sometimes given more structure by the introduction of latent factors which serve to make the series (or transformations of them) co-evolve. The most common of these is the particular type of factor that comes when the matrix $A(1)$ is rank deficient and the series y_t have a permanent component to them i.e. co-integration. A huge literature has evolved to deal with the estimation of the rank of the matrix $A(1)$ i.e. the number of co-integrating vectors, as well as to propose other types of factors that might drive the transitory part of y_t. Because the model is so popular, it seems appropriate to start by looking at some of the issues raised in this research from the perspective of the economic models.

Let us begin with the question of how one selects y_t and p. Mostly, it is very unclear how the selection of y_t is done, with vague references to "economic theory", but it is important to realize that it can have implications for the magnitude of p and also for the ability to transform the statistical model into an economic model i.e. to move up the curve. I will defer the latter point until later.

The reason for an interaction between p and y_t is that the y_t are chosen from a larger universe of variables that appear in the macro economy i.e. the economy generates data on many more variables than are typically included in y_t. Calling the larger set w_t, and assuming it follows a VAR, then it is a truism that y_t will generally not follow a VAR but be a VARMA process -see Zellner and Palm (1974) and Wallis (1977). Since our interest is rarely in the VAR per se, but rather the impulse responses found from $y_t = D(L)v_t$, in order to accurately approximate the latter may require extremely large values for p, far in excess of the orders of VAR that are typically used. This is a problem that has been little studied since it requires one to have some model that would produce impulse responses for the y_t when the DGP actually involves a larger set of variables w_t. This is where models towards the top of my curve, in particular those being used in central banks, can be very useful. Thus some years back I took the FPS model used at the Reserve Bank of New Zealand, which has many variables (w_t) in it, and then asked whether one could capture the impulse responses from imposing various shocks on it with a VAR that involved a much smaller number of variables than w_t[2]. In particular I looked at what value p would need to be in order to get an accurate representation of the impulse responses from that model.

First, the FPS model was simulated with the following six shocks.

- Interest rate
- Real exchange rate
- Inflation
- Terms of trade
- Domestic demand
- Foreign demand

Then six variables were selected from the w_t and the impulse responses of these to the six shocks above were derived. These variables were chosen to be broadly similar to those the RBNZ had utilized in one of their VAR studies.

- Short term interest rate
- Real Exchange Rate
- Inflation Rate
- Terms of Trade
- Aggregate Demand
- Foreign Demand.

[2] This work was done with John McDermott

We then fitted a number of VAR models with different values of p to these impulse responses in order to gain an appreciation of how well a VAR could capture the true impulse responses of the New Zealand economy (assuming that FPS was a good representation of it). There are a number of ways that this approximation might be done. Here we simply make the approximating VAR(K) match the impulses of the FPS model exactly up to K'th order and then study the approximation error from then on. A selection of the results follow. These look at the ability of the VARs to capture the impact of a transitory interest rate rise and a foreign demand shock. For the interest rate shock VAR(2), VAR(6) and VAR(10)'s were imposed and the fit of the impulse responses from the latter were found to match the equivalent quantities from FPS fairly well. The situation was much less encouraging for the foreign demand shock, where a VAR(15) was also used. In fact, the latter seemed to be far worse than the VAR(10) at approximating the FPS reactions. I feel that this example shows how difficult it will be for VAR's of the typical order and number of variables used in many macroeconometric studies to represent an actual economy. It seems likely that the fit can be improved upon with a careful selection of variables to put into y_t but those given above are fairly typical of VAR studies. It is also the case that the class of VARMA models may prove to provide better approximations. The latter have been hard to estimate but there has been recent progress on this - Kapetianos (2002).

After the VAR is set up analysis generally proceeds to the next step of determining the rank of $A(1)$. If it is rank deficient it is decomposed as $\alpha\beta'$, where β will be the cointegrating vectors. Some analysis makes the strong assertion that the β reflect "long-run relations". Whether this is true or not is hard to assess since it identifies long-run relations with the relation between the permanent components of a series. But even if we accept this we rarely see that β comes entirely from the data. Instead proponents of the view generally have some simple economic model lurking in the back of their mind which tells them that the β should have some particular form (generally based on great ratios) and that these look close to what they observe. Of course we know that the β cannot be unique unless there is a single co-integrating relation and it is often the case that, as the number of variables in the relation grows, it becomes very difficult to decide on what the co-integrating vector should be.

In many cases reasoning from simple models may be quite incorrect. An example that surprised me came from the following experiment reported

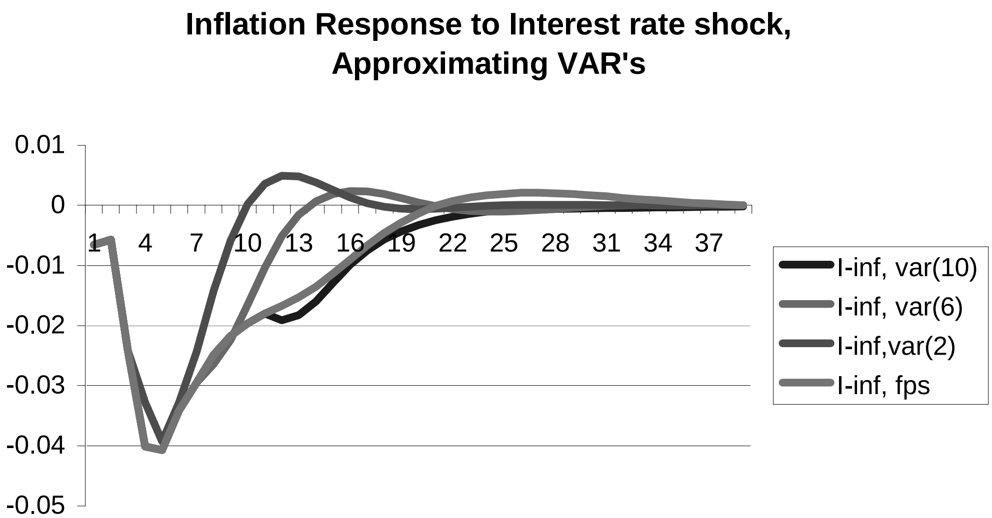

Figure 2:

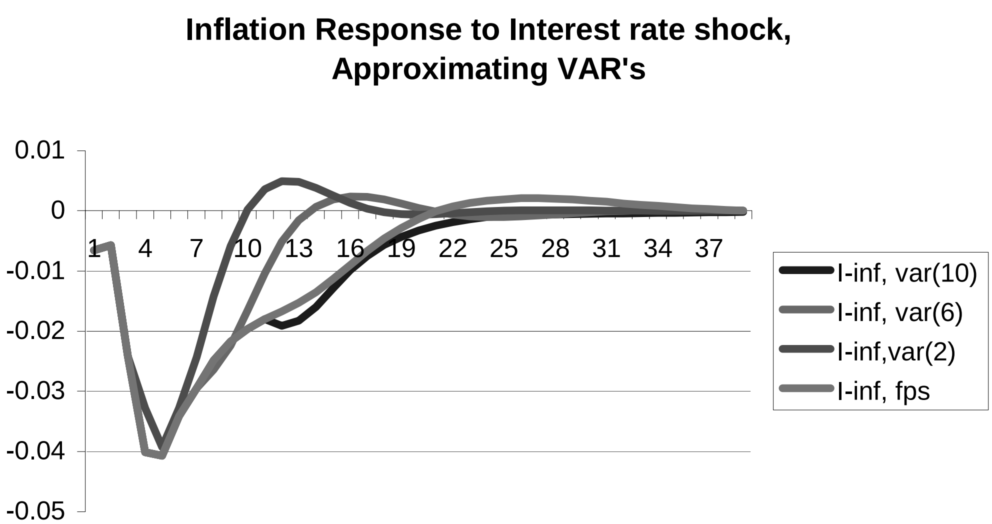

Figure 3:

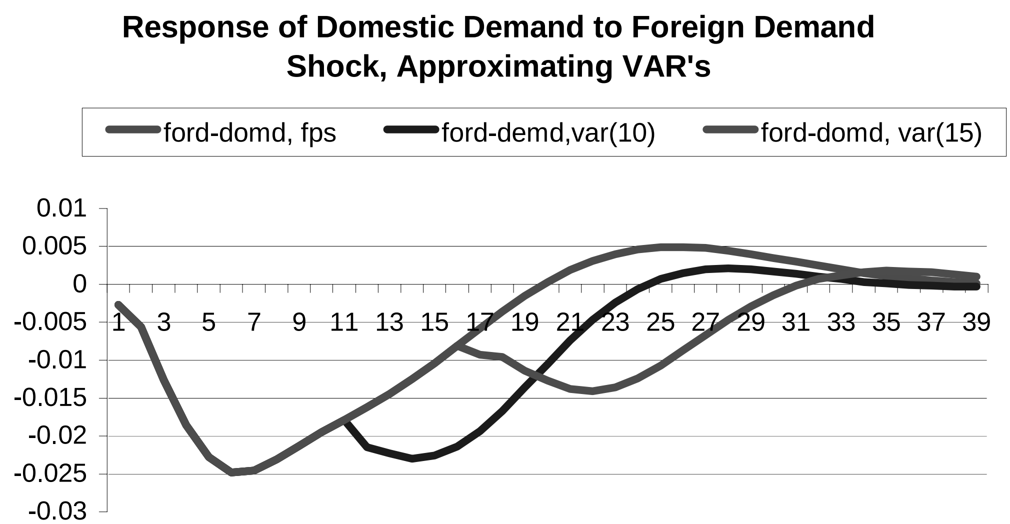

Figure 4:

in McKibbin et al (1998) and Levtchenkova et al (1998), which simulated the long run response of some US variables to four permanent shocks in the MSG2 model documented in McKibbin (1997). The MSG2 model is a multi-country dynamic intertemporal general equilibrium model of the world economy. It has a well determined long run being driven by a Solow-Swan-Ramsey neoclassical growth model, with exogenous technical progress and population growth. In the short run, however, the dynamics of the global economy towards this growth path are determined by a number of Keynesian style rigidities in the goods and labor markets. Households and firms are assumed to maximize intertemporal utility and profit functions subject to intertemporal budget constraints. In the short run some proportion of firms and households use optimal rules of thumb rather than recalculating the entire intertemporal equilibrium of the model. Wages are assumed to adjust slowly to clear labor markets subject to the institutional characteristics of labor markets in different economies. Intertemporal budget constraints are imposed so that all outstanding stocks of assets must be ultimately serviced, and asset markets are efficient, in the sense that asset prices are determined by a combination of intertemporal arbitrage conditions and rational expectations.

The shocks chosen involved 1% permanent increases in the money supply, labour augmenting technical change, oil and LDC lending supply. The variables studied for the effects of the shocks were the log of GDP (ly), the log of the price level (lp), the log of the producer price level (lpp), the log of money (lm), the interest rate (ff) and the log of the trade weighted exchange rate ($ltwi$). The impulse responses to the shocks are the $D(L)$ in $y_t = D(L)v_t$ and the resulting $D(1)$ was:

$$D(1)_{MSG2} = \begin{bmatrix} & Money & Tech & Oil & LDC \\ ly & 0 & .82 & -1.160 & .12 \\ lp & 1 & -.80 & 3.62 & -.22 \\ lpp & 1 & -.82 & 2.83 & -.19 \\ lm & 1 & 0 & 0 & 0 \\ ff & 0 & 0 & .32 & -.15 \\ ltwi & -.8 & .4 & -6.0 & 0 \end{bmatrix}.$$

Some of the elements in $D(1)$ clearly reflect the nature of the experiment; in particular the last three shocks are "pure", in the sense that the money supply is not allowed to increase. But the interesting column is the first, since the exchange rate only changed by -.8, whereas most simple minded approaches would suggest it should be -1. Inspection of the results showed that this was because some countries fix their exchange rates to the US dollar and so, while the effect on many bilateral exchange rates might be a one for one rise with the money shock, it will not be true of the trade weighted index. Now of course one can say that, in the long-run, all exchange rates may be flexible, but imposing such a view is likely to be very inconsistent with historical data. So using simple economic models can be treacherous when producing indicators of the values of co-integrating vectors for actual economies.

Now this example can be pursued a little further since the presence of four permanent shocks among six variables implies two co-integrating vectors. These vectors, β, must satisfy $\beta' D(1) = 0$. Of course such a restriction does not uniquely determine them, and so we choose the first of them such that it is a "real money demand function" and the second as a relation for the exchange rate. With such identifying information one gets

$$
\begin{aligned}
lm &= 1.05ly + 2.03(lpp - lp) - 0.02ff \\
ltwi &= -0.17ly + 5.01(lpp - lp) - 0.8lp + 2.04ff
\end{aligned}
$$

Now these look quite sensible although they imply that the twi is affected by nominal quantities (as we have just observed) and the coefficient of ly is not unity, as is commonly imposed. Furthermore, relative price shifts have an impact upon the holdings of real money balances and this effect needs to be allowed for.[3]

So what would happen if I found the data analogues of the variables in the model and used these to determine co-integrating vectors. Generally data collection is straightforward, except for the money supply, where we used M2. Monthly data from 1974/1-1996/8 was employed and a VAR(6) was fitted. Applying Johansen's estimator and then using the same identification strategy as above i.e. the normalizations on lm and $ltwi$ and the variable exclusions in each case, gives

$$
\begin{aligned}
lm &= 1.52ly + 1.42(lpp - lp) + .05ff \\
ltwi &= 1.26ly - 1.37(lpp - lp) + .4lp + .01ff.
\end{aligned}
$$

It seems clear to me which of these models a decision maker would prefer to work with. In the case of the first co-integrating vector the Johansen results bear some resemblance to the MSG2 ones, but the second one is very different. Given that the latter presumably summarize the data it is not surprising that the MSG2 restrictions are rejected - a $\chi^2(8)$ value of 36.2. But to accept the co-integrating vectors implied by the Johansen estimator would lead one to conclude that a permanent money shock (and the concomitant price level rise) would actually increase the nominal exchange rate. I believe that this conflict between the theory and a relatively unstructured modelling methodology like co-integration is very common. In realistic cases where one is working with more than a few variables one can often find results like this. Too often the examples that are meant to demonstrate the power of statistical methods like Johansen's are rather simplistic, often centering around the estimation of the demand for money, a function that has ceased to be of great moment to many policy makers in the past decade. The biggest

[3] This method has also been applied to the COMPACT model of the UK economy by Jacobs and Wallis (2003).

challenge we face is what to do when we see conflicts like that described above. But often it is only if you have a strong economic theoretical perspective that you would even be aware of it.

2.2 The Utility of Non-linear Models

Statistical models rarely give us a framework that can be used for economic analysis. What attracts people (particularly policy makers) to the top part of the curve is that the models here tell a consistent story that can be readily understood by the group of people making the decisions. Moreover, the questions they want answered are often direct outputs of those models. The fact that we often want to learn about some particular phenomenon directs our attention to the fact that, when assessing the quality of any statistical model, at some point we should design tests that aim to shed light on what the model has to say about the phenomenon of interest. Our philosophy should be to recognize that models are designed to answer specific questions and these should influence our choice of testing procedures.

Now a huge literature has emerged on the need for non-linear models in forecasting and for capturing certain features of the data. From this perspective the VAR model would be an inadequate characterization of the data since it is a linear one. Sometimes the two have been put together e.g. one may have a VAR with a latent factor and the factor is allowed to evolve as a non-linear process. It is interesting to observe that the emphasis upon statistical methods is much stronger when dealing with these non-linear models than was true of VAR's e.g. choices between different non-linear models and whether there is a non-linearity are simply based on likelihood ratio tests or tests of significance (or the equivalent in Bayesian terms). "Economic significance" rarely enters into the picture. By this I mean an assessment of whether the non-linear structure is quantitatively important for the phenomenon being studied. There is nothing wrong with statistical tests. I have been a stalwart advocate of the need for these over the years. But I also feel that we should clearly ask whether whatever non-linearity is claimed to be have been detected is quantitatively important for the question we wish to answer.

2.2.1 A Non-linear Statistical Model of the Business Cycle

I will illustrate the argument above by looking at an example drawn from a literature that I have worked on, viz the value of statistical models of the business cycle. The material is an expanded version of that in Harding and Pagan (2002) and Breunig et al (2002).

We begin with the issue of how to measure the *business cycle* through its turning points. This cycle is sometimes described as the *classical cycle.* It is the definition used by the NBER and refers to the determination of turning points in the *level* of economic activity. Taking a single series Y_t as summarizing the level of economic activity, its turning points would then be the local maxima and minima in its sample path. It is convenient to work with the turning points in $y_t = \ln(Y_t)$ rather than Y_t. Since these turning points are identical the transformation loses no information.

Cycles are divided into *phases*- periods of time in which the economy expands and contracts. Sometimes the latter is described as being a "recession". Common usage of the word "recession" identifies it with a sustained decline in the *level* of economic activity. Consequently, a contraction is initiated by a *peak* in the level of activity and an expansion by a *trough,* so that we need some rule to recognize when a peak (or trough) occurs. Visualizing a peak in a series leads one to the idea that a local peak in y_t occurs at time t if y_t exceeds values y_s for $t-k<s<t$ and $t+k>s>t$, where k delineates some symmetric window in time around t. One can define a trough in a similar way. By making k large enough we also capture the idea that the level of activity has declined (or increased) in a sustained way. Of course we need to limit the size of the window which is used when performing the test. It is this simple idea that is the basis of the NBER procedures summarized in the Bry and Boschan (1971) dating algorithm. In that program, designed for the analysis of monthly data, $k=5$. However, because much analysis is conducted with quarterly data, we will take y_t to be a quarterly series and therefore set $k=2$ as an analogue. One can make the appropriate adjustments, if monthly data are being examined for turning points.

Based on the above discussion we will define turning points in the *business cycle* in the following way.

$$
\begin{aligned}
peak\ at\ t &= \{(y_{t-2,}y_{t-1}) < y_t > (y_{t+1}, y_{t+2})\} \\
trough\quad at\ t &= \{(y_{t-2,}y_{t-1}) > y_t < (y_{t+1}, y_{t+2})\}.
\end{aligned}
\tag{2}
$$

These definitions could be re-expressed as

$$
\begin{aligned}
peak\ at\ t &= \{(\Delta_2 y_t, \Delta y_t) > 0, (\Delta y_{t+1}, \Delta_2 y_{t+2}) < 0\} \\
trough\quad at\ t &= \{(\Delta_2 y_t, \Delta y_t) < 0, (\Delta y_{t+1}, \Delta_2 y_{t+2}) > 0\}
\end{aligned}
\tag{3}
$$

where $\Delta_2 y_t = y_t - y_{t-2}$. In words, a recession occurs if the level of economic activity declines for two quarters and an expansion if it increases for the same interval. In practice, the Bry and Boschan algorithm also applied some extra *censoring* procedures to the dates that emerged from applying the above rule. In particular the contraction and expansion phases must have a minimum duration of six months and a completed cycle must have a minimum duration of fifteen months. We emulate this by imposing two quarter and five quarter minima to the phase lengths and complete cycle durations respectively. Further details on the algorithms that are used to find turning points in this manner can be found in Harding and Pagan (2002) where the computer program which implemented the above rules was termed BBQ[4]. Recently Harding (2003) and Artis et al (2002) have devised useful alternative algorithms for performing the dating; the former is written in GAUSS and the latter in Ox.

As an illustration of the fact that the algorithm simply detects what is visually apparent in the data, figure 5 shows the log of Euro Area GDP over the quarters 1991/2-1993/3. The first quarter of 1992 is marked out as a peak in the business cycle, since it has GDP being larger than for any of the two quarters before or after it. Notice that, whilst 1992/2 was a quarter of negative growth, this was not true of 1992/3, so that we are not defining a recession as two periods of negative growth, although over the recession two periods of negative growth do need to be recorded. The graph also indicates that the recession clearly terminates in the third quarter of 1993. The contraction phase is four quarters long, while the expansion starting in

[4] This program was written in GAUSS and is available at http://www.ecom.unimelb.edu.au/iaesrwww/people/dharding/gcode.html

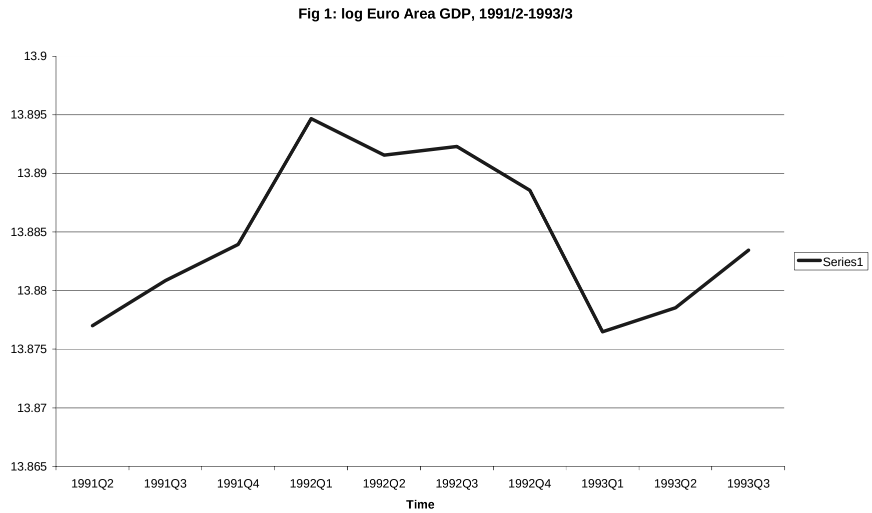

Figure 5:

1993/1 must be at least two quarters long, so that the phase and complete cycle length restrictions are satisfied by the chosen dates for turning points.

The algorithm described above tends to replicate NBER dates quite well -see Harding and Pagan (2003). Thus we conclude that this dating rule seems to be a useful way of constructing business cycle information. It is a very simple algorithm to apply and it is very transparent. It is highly robust in that the dates would not change as one changed the sample of observations. It might not, of course, be robust to major changes in the window width indexed by k. But when one considers that we are choosing k so as to clearly define a turning point it is generally the case that one can make a sensible choice of it in any particular context.

Applying the rules to the level of seasonally adjusted Canadian real GDP from 1947/1 to 1997/2 gives the following statistics on a recession (peak to trough, (PT)) and an expansion (trough to peak, (TP)). These are the durations (in quarters) and the amplitudes (in terms of the percentage change in GDP) across each phase.

Table 1 Business Cycle Characteristics of Models for Canada

	Data	R Walk	MS Model
PT Dur	3	2.7	3
TP Dur	23	23	15.1
PT Amp	-2.1	-1.8	-2.7
TP Amp	29	26.6	20.6

The table also contains the implied recession and expansion characteristics from two statistical models, these being found by applying the BBQ algorithm to data simulated from their calibrated versions. One of these models is the simplest possible linear model for the log of GDP - that it follows a random walk with drift:

$$y(t) = \mu + y(t-1) + \sigma\varepsilon_t,$$

where ε_t is $n.i.d.(0,1)$. We use the sample data to estimate that $\mu = .97$ and $\sigma = 1.24$ and then simulate observations from the process. As one can see from the table the linear model reproduces the business cycle characteristics quite well.

Now, using the same data, Bodman and Crosby (2000) estimate a Markov Switching (MS) model of the type popularized by Hamilton (1989) in which growth Δy_t switches between two values depending upon the realization of a binary random variable z_t. The random variable z_t takes either a value of zero or of one, and it evolves as a Markov Chain. Their model is in (4). They estimate the parameters of this model with MLE and then compare it to a linear model, effectively (4) with $\mu_1 = \mu_0$, using various statistical tests. They conclude that "These results for GDP growth indicate statistical evidence of non-linearity" (p624). Moreover diagnostic tests performed on the model residuals suggest that there was no inadequacy in the fit.

$$\Delta y_t = \mu_0 + z_t\mu_1 + \sum_{i=1}^{4} \phi_i \left(\Delta y_{t-i} - \alpha_0 - z_{t-i}\alpha_1\right) + \sigma e_t. \tag{4}$$

Even if one accepts these tests one has to ask whether the non-linearity is important for the business cycle, which was the motivation for studying the series. It is clear from Table 1 that it isn't. In fact it produces a much worse description of the business cycle than that obtained with the simplest linear model. Indeed, if one decided to evaluate the model by asking how well it can reproduce the business cycle outcomes, one would almost certainly reject it, despite the fact that it has passed a battery of standard statistical tests e.g. the presence of serial correlation in the residuals. The point of the example then is that one should always think about devising at least one test of a model that addresses the economic issue we are really interested in analyzing and not just rely on general purpose tests that are imported from statistics.

Parametric Encompassing Tests In response to the conclusion above one might ask whether the durations of expansions in the MS model are really different to those in the data. Since we will often be interested in considering such questions i.e. we want to know whether the model under investigation encompasses some characteristic we have selected, we will test that proposition.

Let $\hat{\gamma}$ be a parameter that has been estimated from the data and which is to be the basis of an encompassing test e.g. the average duration of expansions. As well we can think of the comparable quantity implied by the statistical model. This will be designated as $\gamma_M(\hat{\theta})$, where $\hat{\theta}$ is the MLE of the parameters θ associated with the statistical model. In most instances we will need to find $\gamma_M(\hat{\theta})$ by taking the point estimate $\hat{\theta}$, simulating data from the statistical model, and estimating $\gamma_M(\hat{\theta})$ from the simulated data. We will assume that we have performed enough replications that the error made in estimating $\gamma_M(\hat{\theta})$ by simulation methods is negligible. Hence $\gamma_M(\hat{\theta})$ is random solely due to the fact that θ has been estimated. Like most of the simulation literature a scaling factor of $(1+\frac{T}{M})$ could be applied to the variance of any test statistic to make an allowance for the effect of the simulation error upon the variance of an estimator, where M is the number of replications and T the number of observations in the sample.

Now consider the statistic based on a comparison of the two quantities:

$$\hat{\tau} = \hat{\gamma} - \gamma_M(\hat{\theta}).$$

The problem is often to find the asymptotic variance of $\hat{\tau}$. One is rarely given the information needed to compute this, namely the asymptotic variance of $\hat{\theta}$. To get around this we compute the test statistic

$$R = \hat{\tau}'[var(\hat{\gamma})]^{-1}\hat{\tau},$$

since it can be shown - as explained in Breunig et al (2002)- that this is less than or equal to the true Wald test statistic

$$R^* = \hat{\tau}'[var(\hat{\tau})]^{-1}\hat{\tau}.$$

This is a simple application of the Newey (1985) and Tauchen (1985) framework for specification testing. Thus it is a conservative test, in that $R \leq R^*$. Consequently, if we reject the null hypothesis using R we must do so more strongly using R^*. Using R is advantageous in that the only thing we need to know is $var(\hat{\gamma})$ and this can be computed from the sample data.

Perhaps the main difficulty we encounter is that of consistently estimating $var(\hat{\gamma})$. Ideally we want to do that under the null hypothesis that the statistical model is correct (and characterized by parameters $\hat{\theta}$) and so we might therefore simulate data from the statistical model, finding what the $var(\hat{\gamma})$ is from the simulations. Alternatively, we might use asymptotic theory and compute a robust estimator of $var(\hat{\gamma})$ that would be compatible with many alternative models. Mostly we use this latter strategy since we measure $\hat{\gamma}$ through a regression and it is a very simple operation in most regression programs to produce robust standard errors. We compute a standard error that is robust to both heteroskedasticity and serial correlation using the Newey-West type formula in Hamilton (1994, p283, eq (10.5.21)) with q set equal to 9.

Testing a Specific MS model of the Business Cycle We now use the encompassing test above to see if the MS model of Bodman-Crosby is needed to produce an explanation of the business cycle. In this case $\hat{\gamma}$ is the estimated average completed duration of expansions. The value of $\sqrt{R} = 13.3$ indicates an emphatic rejection of the model in terms of its ability to explain the business cycle. It should be emphasized that this model had passed many statistical tests for model adequacy. This shows the advantage of focussing directly upon a characteristic that specifically addresses the phenomenon we want to explain, in this case the business cycle, rather than relying upon generalized statistical measures.

3 Moving up the Frontier

It is now time to take some steps away from the bottom of the frontier and move up it towards the economic theoretical models. To do this we need to re-express the statistical model in a form that enables us to tell an economic story. Once done, the new representation is often used to make claims about outcomes such as the business cycle. Thus we need to look carefully at the methods that have been used in this step in order to understand their limitations and the pitfalls that can be present in their use.

3.1 Re-expressing the VAR as an Economic Model

The first move up the frontier towards economic models involves the transformation of a VAR into a structural VAR (SVAR) of the form

$$(B_0 - B_1 L - ... - B_p L^p) y_t = \varepsilon_t, \tag{5}$$

where the shocks ε_t are now given economic names such as "demand", "supply" etc and these are assumed to be uncorrelated with one another. The v_t in the reduced form VAR are nameless and are allowed to be correlated. Thus it is the naming of the shocks which essentially introduces economic content. The VAR and SVAR models are generally made observationally equivalent by ensuring that the structural system of equations is exactly identified, in the sense that one can recover all the parameters in B_j and $cov(\varepsilon_t)$ uniquely from the A_j and $cov(v_t)$. There are a number of ways of doing this involving making B_0 triangular, imposing restrictions between B_0 and $\{B_1, ..., B_p)$ (generally called long-run restrictions), or placing sign restrictions upon the impulse responses C_j i.e. the response of y_{t+j} to a unit change in ε_t.

There is little that one can generally say about the utility (and validity) of restrictions on B_0 and B_1. Whether it is reasonable to assume that B_0 is recursive for example will depend on the context. However, one can sound some warnings e.g. long-run restrictions can often result in weak instruments being used since they effectively use variables x_{t-1} as instruments for Δx_t, and this can distort inference a good deal- see the pictures of the densities from models such as Gali (1992) that use long-run restrictions in Pagan and Robertson (1998). These can be multi-modal and far from normal even if the sample size is large by macro-economic data standards. But it is the assumption that the ε_t are uncorrelated that I want to look at more closely, as it is often invoked with little thought.

3.1.1 The Uncorrelated Shocks Assumption

A difficulty with SVAR's that is often glossed over is the assumption that the shocks ε_t are uncorrelated with each other. One might argue that this is not unreasonable if the system y_t is of large enough dimension but, in other cases, it may well be that a factor that is common to the explanation of more than one member of y_t has been omitted. In this case the factor would be included in the error terms and, being common, will mean that the assumption of uncorrelated errors is invalid. Consequently, imposing an incorrect restriction will generally bias estimators of the parameters of the SVAR.

To think about the likelihood of that happening one might return to the economic models higher up on the curve. A popular miniature model has been the New Keynesian policy model of the form

$$\begin{aligned}
y_t - y_t^* &= a_1(i_{t-1} - \pi_{t-1}) + a_2(y_{t-1} - y_{t-1}^*) + \varepsilon_{dt} \\
\pi_t - \overline{\pi} &= b_1(\pi_{t-1} - \overline{\pi}) + b_2(y_{t-1} - y_{t-1}^*) + \varepsilon_{st} \\
i_t - \pi_t &= c_1(y_t - y_t^*) + c_2(\pi_t - \overline{\pi}) + c_3(y_{t-1} - y_{t-1}^*) + c_4(\pi_{t-1} - \overline{\pi}) + \\
&\quad c_5(i_{t-1} - \pi_{t-1}) + \varepsilon_{mt}
\end{aligned} \tag{6}$$

where the equations are the IS curve, the Phillips curve and the interest rate rule respectively. The variable $y_t - y_t^*$ is an output gap, y_t^* is potential output, π_t is the inflation rate, $\overline{\pi}$ is the target inflation rate, i_t is the interest rate and $\varepsilon_{dt}, \varepsilon_{st}$ and ε_{mt} are demand-, supply- side and monetary policy shocks. In the interest of simplicity, I have abstracted from any expectation terms and have made the responses in the first two equations lagged ones rather then contemporaneous. Now the SVAR literature would imply that the first two equations can be estimated by OLS and, with the assumption that $\varepsilon_{dt}, \varepsilon_{st}$ are uncorrelated with ε_{mt}, one can use the residuals from estimating those equations as instruments in the third one to estimate c_1 and c_2.

Now there is a potential catch here - the SVAR literature has invariably worked with equations in y_t and not the output gap, and so it is the following system that is actually estimated rather than (6).

$$\begin{aligned}
y_t &= a_1(i_{t-1} - \pi_{t-1}) + a_2 y_{t-1} + \{\varepsilon_{dt} + y_t^* - a_2 y_{t-1}^*\} \\
\pi_t - \overline{\pi} &= b_1(\pi_{t-1} - \overline{\pi}) + b_2 y_{t-1} + \{\varepsilon_{st} - b_2 y_{t-1}^*\} \\
i_t - \pi_t &= c_1 y_t + c_2(\pi_t - \overline{\pi}) + c_3 y_{t-1} + c_4(\pi_{t-1} - \overline{\pi}) + \\
&\quad c_5(i_{t-1} - \pi_{t-1}) + \{\varepsilon_{mt} - c_1 y_t^* - c_3 y_{t-1}^*\}
\end{aligned} \tag{7}$$

In (7) the errors (in curly brackets) are clearly correlated. Therefore the residuals from estimating the first two equations in (7) cannot be used as instruments in the last equation. If they are then one will get inconsistent estimates of the parameters, leading to biases in the estimated impulse responses for the money shock. Giordani (2002) shows that this is an explanation of the "price puzzle" in many SVAR studies, wherein a contractionary monetary policy shock leads to a rise in the price level. Thus this is an example of how using one of the models at the top end of the curve can be very informative about how one should proceed when trying to move up the curve. There is nothing incorrect about the fit of the VAR in this instance, but rather it is how one moves closer to an economic model (the SVAR) that creates the difficulty.

3.1.2 What do Variance Decompositions Tell us about Cycles?

One of the advantages claimed for SVAR models is the ability to address questions related to macroeconomic fluctuations, in particular to shed light on questions such as what factors drive the business cycle? The technique that is most commonly used to answer the latter question is the decomposition of the variance of multi-step forecast errors into the contributions from each of the shocks in the system. Now the problem with such exercises is that the connection between multi-step forecasts and the cycle is never clear. Consequently, in this sub-section we use an example to show that there is no connection and that one needs other modes of analysis when it comes to answering questions about the cycle.

Suppose that a model has been fitted which enables one to write down a decomposition of y_t into a sum of past shocks:

$$y_t = \sum_{i=1}^{K}\sum_{j=0}^{\infty} C_{ij}\varepsilon_{it-j},$$

where ε_{it} are uncorrelated shocks and the impulse responses of y_{t+j} to a unit rise in ε_{it} are C_{ij}. The forecast variance of $y_{t+L} - y_t$, the L step ahead forecast error, can be written in terms of the sums of C_{ij}^2 and therefore the fraction of the variance explained by each of the shocks can be determined. Often L is set to the "business cycle horizon" and the dominant contributor to the variance of $y_{t+L} - y_t$ is regarded as the "cause" of the business cycle. A recent paper that does this is Altig et al. (2002) who find that 46% of the forecast variance at $L = 12$ and 63% at $L = 30$ is explained by technology shocks.

What has this variance decomposition got to do with the business cycle? Although oft-repeated, the connection seems to be more one of assertion than coming from any precise analysis. We will demonstrate this by using the example in Altig et al (2002).[5]

In this model there are n variables (y_t) and n shocks, where $n = 9$. One of the variables is the growth in GDP. The model is a structural VAR as in (5) and, associated with it is a "reduced form" VAR, as in (1). The shocks in the two representations are connected as

$$v_t = \Gamma\varepsilon_t = \Gamma_1\varepsilon_{T,t} + \Gamma_2\varepsilon_{M,t} + u_t, \tag{8}$$

where we follow their paper and focus upon only two shocks - to money $\varepsilon_{M,t}$ and to technology $\varepsilon_{T,t}$. The other shocks will not be specifically identified and are grouped into u_t. The money and technology shocks are identified by imposing some restrictions upon the SVAR, namely

- Long-run restrictions to identify the technology shock.
- Short-run restrictions to identify the monetary shock.

Imposition of these restrictions provides enough instruments to estimate the shocks in the SVAR system that can be regarded as technology and money. Let us designate these by $\hat{\varepsilon}_{T,t}$, $\hat{\varepsilon}_{M,t}$. To perform the volatility decomposition it is necessary to find the impulse responses of output to the money and technology shocks i.e. the C_{ij}. To do this we first compute the impulse responses showing the effect on y_t of v_t. These are $\hat{\Pi}^v_j$ and can be found from the VAR. Second, since it is assumed that $\varepsilon_{T,t}$ and $\varepsilon_{M,t}$ are uncorrelated with all remaining members in ε_t, (8) shows that we can regress $\hat{v}_t$ on $\hat{\varepsilon}_{T,t}$ to get $\hat{\Gamma}_1$ and $\hat{v}_t$ on $\hat{\varepsilon}_{M,t}$ to get $\hat{\Gamma}_2$. It then follows that the required impulse responses can be recovered from those of v_t using $\hat{\Pi}^M_j = \hat{\Pi}^v_j\Gamma_2, \hat{\Pi}^T_j = \hat{\Pi}^v_j\Gamma_1$. Note that the impulse responses to the other shocks in ε_t cannot be determined with the limited information being used. Once these

[5] The material here was first used when I was a discussant of that paper. Subsequently Altig et al have plotted the implied values of GDP from their model over the sample period to which their SVAR was fitted in order to show that technology and money shocks fail to explain the US business cycle. Our aim here is to address the utility of variance decompositions rather than saying anything about the U.S. business cycle. By performing the following analysis however one also gains some understanding of the reasons why the variance decomposition is of little use for cycle questions.

impulse responses have been found one can perform the variance decompositions to determine the relative contributions of money and technology to the L step ahead variance of output.

To see if this variance decomposition is really informative about the factors driving the business cycle we simulate data on GDP from the estimated VAR of their paper in two situations: when technology and money shocks are present and when they are absent. Using the simulated data we can then see how the business cycle changes when these shocks are present and when they are absent. This is done by defining

$$\hat{u}_t = \hat{v}_t - \hat{\Gamma}_1 \hat{\varepsilon}_{T,t} - \hat{\Gamma}_2 \hat{\varepsilon}_{M,t},$$

and performing the following two experiments:

- When all shocks are present we utilize the $cov(\hat{v}_t)$ estimated from the VAR to generate shocks for the simulations.
- When technology and money shocks are removed we use the $cov(\hat{u}_t)$ estimated from the $\hat{u}_t$ above to simulate data.

Each of these simulated data sets is then passed through the BBQ program to date the cycles that are associated with them, Results are given in Table 3.

Table 3 U.S. Business Cycles with and Without
Money and Technology Shocks

	All shocks	Tech and Mon removed
Dur		
Contrac	4.2	4.3
Expan	18.2	20.5
Amp		
Contract	-.1.6	-.1.7
Expan	27	29

It is clear from Table 3 that technology and money shocks explain little of the business cycle, which is in striking contrast to the long horizon variance decomposition which claims that technology is the major contributor to the cycle. Thus this example demonstrates that the long horizon variance decompositions do not provide any useful information on what are the important drivers of the cycle. To understand why this is so in the Altig et al case we return to the simple model for GDP (where y_t is now a scalar and is the log of GDP)

$$\Delta y_t = \mu + \rho \Delta y_{t-1} + \sigma e_t.$$

In this situation the cycle in y_t only depends on μ, ρ, σ - see Harding and Pagan (2002). Hence one can gain some insight into the results of Table 3 by fitting this simple model to the simulated data from the experiments above and seeing what differences there are in the values of these three parameters. Doing this I found that μ and ρ essentially remained the same in both experiments and it was only σ that differed. Thus the changed business cycle characteristics in Table 3 depend directly upon how much σ was changed. In the first experiment $\sigma = .0064$, while in the second it declines to .0051. Consequently, because σ does not change very much when technology and money shocks are removed from the system, it is clear that the business cycle will also change very little. This leads us to conclude that there is no short-cut available from impulse responses that will tell us about determinants of the business cycle. One needs to have a mechanism whereby one can perform experiments in which factors (shocks) are removed and the business cycle implications directly assessed. Use of a dating algorithm like BBQ enables this to be done.

3.2 The View from a Position Towards the Top of the Frontier

We now want to leap up to that part of the frontier which has the models that are becoming popular in many central banks around the world. At such a position we want to look in both directions - downwards towards the SVAR and VAR world and upwards towards the miniature models such as the New Keynesian Policy model. Our view will be facilitated by asking what the models in this position would imply for those further up or down the curve. It is often useful to perform such a conversion, either to understand the model residing at the current point or for providing guidance on the nature of models in other locations Thus we would often like to know what VAR (or SVAR) would be expected from one of the models that is towards the top of the curve. In the case of extracting VAR's from (say) the New Keynesian policy model it is easy to do, often analytically. But, as the models get bigger, it becomes virtually impossible to do that. So we seek a general method whereby one can downsize a bigger model to a smaller model. Kapetianos et al (2003) have discussed this in the context of a large scale model whose theoretical base is similar to that of QPS and FPS and which contains the *core theory* for a new model under development at the Bank of England. We will refer to this core model as BE. It is not the model actually used in forecasting and policy analysis but is the base for such a model The key to our analysis will be the fact that models like BE can be subject to shocks such as productivity, monetary policy etc and the responses of a vector of variables y_t to these shocks ε_t can be measured.

3.2.1 Looking Downwards

Let us look at how BE might be converted to either univariate or SVAR models in a set of variables y_t. An issue that has to be addressed is whether we want to understand the nature of the processes in BE or to utilize it to say something about what a VAR should look like when faced with actual data. The first is really an issue about the intrinsic dynamics of the model. It is useful to think about this in the context of a simple forward looking model which summarizes the Euler equations underlying the model. These would have the stylized (and simplified) form

$$y_t = Ay_{t-1} + BE_t(y_{t+1}) + Cu_t$$

where u_t are some shocks. Binder and Pesaran (1995) point out that the solution to this model is

$$y_t = P(A,B)y_{t-1} + \sum_{j=0}^{\infty} \Psi^j(P,B)CE_t(u_{t+j}), \tag{9}$$

where $P\,(A,B)$ indicates dependence on A, B etc. Now to complete the solution some assumption needs to be made about how u_t evolves. If it has the form

$$u_t = \Phi u_{t-1} + \varepsilon_t$$

where ε_t is white noise, then the final solution will have the form

$$y_t = Py_{t-1} + G(P,C,\Phi)u_t$$

Thus the solution has parameters that partly depend on the nature of the economic model (P) and others that also depend also upon the nature of the shocks u_t. The *intrinsic dynamics* in y_t (P) depends only upon the model parameters, A, B. But there may also be a contribution to the dynamics of y_t from the nature of u_t itself since the system is really a VAR of the form

$$y_t = Py_{t-1} + G\Phi u_{t-1} + G\varepsilon_t,$$

pointing to the fact that there are extra or *extrinsic dynamics*. Moreover, when shocks are permanent u_t will be an $I(1)$ process, making y_t also $I(1)$, and the representation for y_t will now potentially be an ECM rather than a VAR. When confronting the data we will need to determine the nature of the shocks u_t i.e. whether they are transitory or permanent and, if they are not the latter, how persistent are they? In this lecture we will restrict ourselves to studying the intrinsic dynamics in BE by considering its reaction to transitory shocks alone. Kapetianos et al (2003) discuss the issues when shocks are permanent.

We begin by simulating BE with unit shocks (separately) and computing the j'th impulse responses of y_t to the vector of shocks ε_t. These are designated as C_j and they are used to generate synthetic data on the y_t in the following three steps:

- Draw $N(0, V)$ random numbers for ε_t
- Choose R and compute pseudo-data for y_t using $y_t = \sum_{j=0}^{R} C_j \varepsilon_{t-j}$. R can be chosen to be very large after inspecting how quickly the impulses die out. A very large number of observations on y_t can be generated.
- Fit a model to the y_t e.g. a VAR using standard econometric methods

For our exercise we can set $V = I$ since the intrinsic dynamics are not affected by this aspect of the shocks. However if one was trying to determine an implied VAR that would fit a given data set it would be necessary to determine V. Since we know that

$$V_y = E(y_t y_t') = \sum_{j=0}^{\infty} C_j V C_j'$$

and, since the LHS can be estimated using data on y_t, while the C_j are known, it is therefore possible to calibrate V. Specifically, since

$$vec(V_y) = (\sum_{j=0}^{\infty} C_j \otimes C_j) vec(V),$$

we have

$$vec(V) = (\sum_{j=0}^{\infty} C_j \otimes C_j)^{-1} vec(V_y).$$

Of course one may only want to recover the diagonal elements of V if it is to be assumed that the shocks have zero correlation.

Now we look at what BE would imply about small models consisting of only five variables- GDP, inflation in the consumption deflator, an output gap, a real exchange rate and a real interest rate. Five transitory shocks are used to compute the impulse responses- these are shocks to foreign demand, government expenditure, the inflation target, technology and the foreign exchange risk premium.

As a first illustration of what might be learned from this procedure Table 4 contains the value of the AR(1) coefficient when an AR(1) is fitted to the pseudo-data generated on the four series output, inflation, the real interest rate and the real exchange rate from the BE model using $R = 140$.

Table 4 Persistence of Variables in BE
Measured by AR(1) Coefficients

Output	.52
Inflation	.89
Real Int Rate	.56
Real Ex Rate	.28

It's clear that there is strong persistence to inflation in the model but only moderate amounts for the other variables. Thus, while the intrinsic dynamics seem to be of reasonable magnitude, in order to explain observed data it is clear that one would need to consider the introduction of permanent shocks, since output has to behave more like an I(1) variable and there should be much greater persistence in the real exchange rate.

Now we could also fit an SVAR to the synthetic data in order to perform a similar study to that done earlier with the FPS model. The variables in the

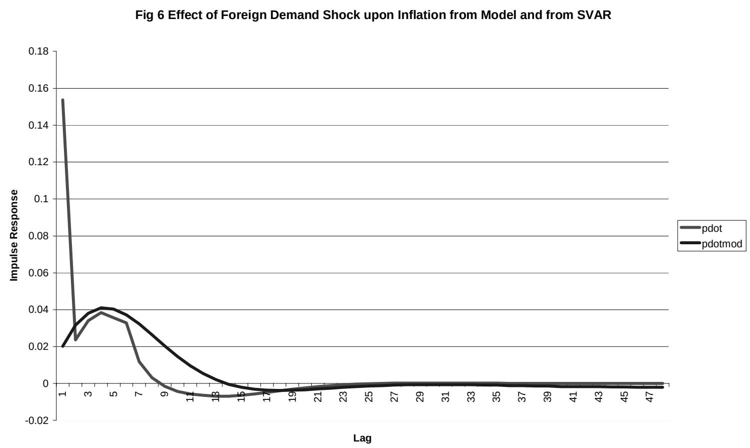

Figure 6:

SVAR were ordered as above (with foreign demand as a fifth variable that is first in the ordering) and then a recursive SVAR is fitted. Figs 6-8 show the true impulse responses of inflation and the real interest rate to a foreign demand shock versus those found from the SVAR. An SVAR(6) was fitted, as that is the order typically used in estimating quarterly SVAR's of this type. Data was generated using 140 lags of the impulse responses and 40,000 observations were used in estimation. The lesson from this experiment is somewhat similar to that drawn from FPS - although there are differences early on, the ability to accurately estimate the responses deteriorates markedly after the maximum order of the VAR.[6]

3.2.2 Looking Upwards

We finally turn briefly to looking how one might downsize the BE model into a model that is a familiar one in recent monetary policy analysis i.e. the

[6]Obviously this technology could also be used to estimate the parameters of the larger model. One would choose (say) a VAR as an auxiliary model and then use indirect estimation methods to match the parameters of the VAR estimated from the data and from the model. It is not necessary that the chosen auxiliary model be the DGP.

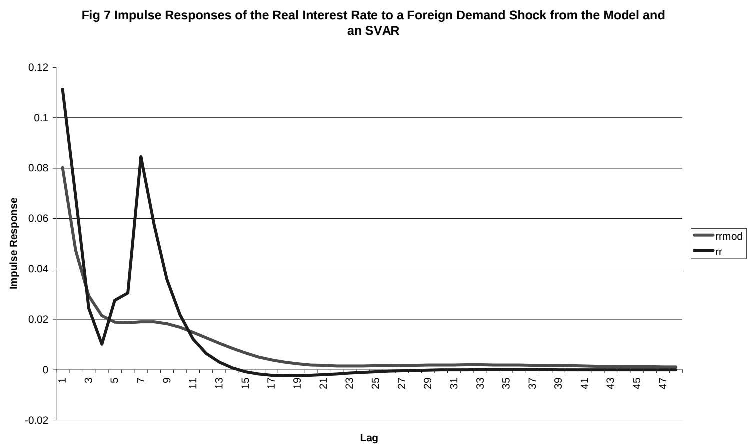

Figure 7:

New Keynesian model as set out in (6), but now with expectation effects added back into the Phillips curve and contemporaneous variables in each equation. To perform the downsizing we take the data simulated as described in the previous sub-section and then just fit such a model. It is necessary to be careful here however as one needs to utilize instruments to overcome the presence of expectations and contemporaneous variables. Since there are few variables in the system it is inevitable that one will resort to lagged values of them as instruments, although, because foreign demand is exogenous, it can sometimes be used for that purpose. As one can see from Table 4 above there is a reasonable degree of persistence in the endogenous variables but not that large as to make one comfortable in using (say) a second lag as an instrument for a contemporaneous value. Generating 40000 observations is useful however, since, even if the instrument is fairly weak, it is likely that it can still be useful in such sample sizes. Of course one can generate as many observations as desired, and sometimes we have found it necessary to generate much larger samples in order to produce reliable estimates of the parameters, but this is not so for the current version of the BE model.

Estimates of the small system are given below[7]

$$\begin{aligned}
y_t - y_t^* &= -.66(i_t - \pi_t) - .06rer_t + 2.54y_t^f + .29(y_{t-1} - y_{t-1}^*) + \varepsilon_{dt} \\
\pi_t - \overline{\pi} &= .54(\pi_{t-1} - \overline{\pi}) + .45E_t(\pi_{t+1} - \overline{\pi}) + .01(y_t - y_t^*) + \varepsilon_{st} \qquad (10) \\
i_t - \pi_t &= .02(y_t - y_t^*) + .33(\pi_t - \overline{\pi}) + .66(i_{t-1} - \pi_{t-1}) + \varepsilon_{mt}
\end{aligned}$$

Initially the nominal exchange rate was placed in the Phillips curve but the coefficient was minute and easily passed a test that it was zero. Hence it is excluded from the equation above. Based on theoretical arguments McCallum and Nelson (2000) argued for such a restriction on Phillips curves. The BE model doesn't have a Phillips curve in it so it is interesting to see what type of curve is implied. In particular it is noticeable that there is an even balance between forward and backward looking effects in expectations. There is very little effect of an output gap in the interest rate rule although the rule actually used in the BE model has a stronger effect that recovered above (and the impact of inflation is smaller as well). But we wouldn't expect to be able to recover that function exactly due to the fact that the system has been downsized. Nevertheless we do recover the essence of the rule. It is also of interest to just estimate the monetary policy rule with OLS. Then we would actually get the wrong sign on both the inflation and the output gap variables. Above we used instruments that were the lagged values of the variables on the RHS.

4 References

Allsopp, C. and D. Vines (2000), "The Assessment: Macroeconomic Policy", *Oxford Review of Economic Policy*, 16, 1-32.

Altig, D., L.J. Christiano, M. Eichenbaum and J. Linde (2002), "Technology Shocks and Aggregate Fluctuations" (mimeo, Northwestern University)

Artis, M, M. Marcellino and T. Proietti (2002), "Dating the Euro Area Business Cycle", (mimeo, European University Institute)

[7] Although we have written the system as involving an output gap in practice we just use output. Since one knows what y_t^* is in the model one could utilize that, although it would be unknown in practice. In fact in this instance it doesn't lead to the estimated parameters changing much.

Black, R., V. Cassino, A. Drew, E. Hansen, B.Hunt, D. Rose and A. Scott (1997), "The Forecasting and Policy System: The Core Model", *Reserve Bank of New Zealand Research Paper 43.*

Bodman,P.M. and M. Crosby (2000), "Phases of the Canadian Business Cycle", *Canadian Journal of Economics*, 33, 618-633.

R. Breunig, S. Najarian and A.R. Pagan (2002), "Specification Testing of Markov Switching Models" presented at EC2 Conference, Bologna

Bry, G., Boschan, C., (1971), *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs,* New York, NBER.

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

Gali, J. (1992) "How well does the IS-LM model fit postwar U.S. data?",107, 709–738.

Giordani, P. (2001), "An Alternative Explanation of the Price Puzzle", mimeo, Stockholm School of Economics

Hamilton,J.D.,(1989), "A New Approach to the Economic Analysis of Non-Stationary Times Series and the Business Cycle", *Econometrica*, 57, pp. 357-384.

Hamilton, J.D. (1994), *Time Series Analysis* (Princeton University Press)

Harding, D. (2003), *Essays on the Business Cycle* (Ph.D. thesis, Yale University)

D. Harding and A. R. Pagan (2002) "Dissecting the Cycle: A Methodological Investigation", *Journal of Monetary Economics*, 49, 365-381.

Harding, D. and A. R. Pagan (2003), "A Comparison of Two Business Cycle Dating Methods" *Journal of Economic Dynamics and Control*, 27, 1681-1690

Jacobs, J. and K. F. Wallis (2002), "Comparing SVARs and SEMs: Two Models of the UK Economy", mimeo, University of Warwick.

Kapetanios, G. (2002), "A note on an iterative least squares estimation method for ARMA and VARMA models" , mimeo, Queen Mary College.

G. Kapetianos, A.R. Pagan and A. Scott (2003) " A Framework for Relating Economic-Theoretic and Economic-Statistical Models in the Macroeconomic Policy Process" (forthcoming)

S.Levtchenkova, A.R. Pagan and J. Robertson (1998) "Shocking Stories", *Journal of Economic Surveys*, 12, 1998, 507-532.

McCallum, B.T. and E. Nelson (2000), "Monetary Policy for an Open Economy: An Alternative Framework with Optimizing Agents and Sticky

Prices", *Oxford Review of Economic Policy*, 16, 74-91.

W.McKibbin, A.R.Pagan and J. Roberston (1998), "Some Experiments in Constructing a Hybrid Model for Macroeconomic Analysis" *Carnegie-Rochester Series in Public Policy,* 49, 113-142,

McKibbin, W. and D. Vines (2000), "Modelling Reality: The Need for Both Inter-temporal Optimization and Stickiness in Models for Policy-Making", *Oxford Review of Economic Policy*, 16, 106-137.

Newey, W.K. (1985), "Maximum Likelihood Specification Testing and Conditional Moment Tests", *Econometrica*, 53, 1047-1070.

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

A.R. Pagan and J. Robertson (1998) "Structural Models of the Liquidity Effect", *Review of Economics and Statistics*, 80, 202-217.

Tauchen, G (1985), "Diagnostic Testing and Evaluation of Maximum Likelihood Models", *Journal of Econometrics*, 30, 415-443.

Smets, F. and R. Wouters (2002), "An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area", *Working Paper #171, European Central Bank*

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.