MELBOURNE INSTITUTE
Applied Economic & Social Research

Working Paper Series

Implications of Partial Information for Applied Macroeconomic Modelling

Adrian Pagan

Tim Robinson

Working Paper No. 12/19
October 2019

Implications of Partial Information for Applied Macroeconomic Modelling*

Adrian Pagan
School of Economics, University of Sydney,
CAMA, Australian National University

Tim Robinson
Melbourne Institute of Applied Economic and Social Research
The University of Melbourne

Melbourne Institute Working Paper No. 12/19
October 2019

* Earlier versions of this appeared under different titles as Pagan and Robinson (2016) and (2019); it also contains material from the working paper Pagan (2017). We thank Efrem Castelnuovo, Mariano Kulish, Viet Nguyen, Giovanni Pellegrino and Sarantis Tsiaplias, participants at the 2019 Macroeconomic Developments Conference at Deakin University, the 2016 Padova Macro Talks, and seminars at Monash University and the Universities of Adelaide and Tasmania for comments on earlier versions. All errors are our own. Research supported by ARC Grant DP160102654.

Melbourne Institute:
Applied Economic & Social Research
The University of Melbourne
Victoria 3010 Australia
T +61 3 8344 2100
F +61 3 8344 2111
E melb-inst@unimelb.edu.au
W melbourneinstitute.unimelb.edu.au

Abstract

Implications of partial information for applied macroeconomic modelling along four dimensions are shown, and analysis provided on how they can be addressed. First, when permanent shocks are present a Vector Error-Correction Model including latent, as well as observed, variables is required to capture macroeconomic dynamics. Second, the assumption in Dynamic Stochastic General Equilibrium models that shocks are autocorrelated provides identifying information usable in Structural Vector AutoRegressions. Third, estimating models with more shocks than observed variables must yield correlated estimated structural shocks. Fourth, including measurement error, as commonly specified, implies a lack of co-integration between variables, even when actually present.

JEL classification: E37; C51; C52.

Keywords: SVAR; Partial Information; Identification; Measurement Error; DSGE.

1 Introduction

The presence of partial information raises numerous practical issues that an applied macroeconomic researcher must address. For example, when modelling the macroeconomy as a system not all of the variables may be readily observed, or they may be thought to be measured with error. The researcher may neither know the appropriate identifying restrictions nor have enough observed data to accurately estimate all the structural shocks. In this paper we study the implications of partial information for applied macroeconomic research along four dimensions and show how the issues arising from this can be practically addressed.

The first dimension we study is the implications of not all variables being readily observed. To analyse this it is necessary to have a model that is taken to represent the actual economy, which we term the Representative Model (RM). A wide variety of modelling strategies are now used in macroeconometrics for policy analysis. These models differ in their scale, the extent to which they impose strong theoretical ideas upon the data, and the particular purpose for which they are constructed. Examples include: large-scale Dynamic Stochastic General Equilibrium (DSGE) models - e.g. the Federal Reserve's EDO model (Chung et al. 2010) and the Multi-Sector Model of Rees et al. (2016); macroeconometric models, such as FRB/US and MARTIN of the Reserve Bank of Australia; and Structural Vector AutoRegression (SVAR) models, the latter being oriented more towards the data than the theory.[1] To study the implications of not all variables being observed we will take some of the more theory-orientated models as the RM, and examine how well their impulse responses are matched by SVARs and structural Vector-Correction Models (VECM) in just the observed variables.

Comparisons between the impulse responses from a RM and SVARs are common in the literature. SVARs provide economic interpretations of the shocks hitting the economy, but have relatively loose theoretical underpinnings compared to the other structural modelling

[1] There have always been a large range of models used for forecasting, but we are concerned here with models constructed to enable the examination of policy options.

approaches, and consequently may be more aligned with the data. Hence their impulse responses are often used as a check on whether more tightly-specified, larger models are capturing the data adequately. An example is Brayton et al. (2014, p.4), who say, in the context of the Federal Reserve's FRB/US model:

> The responses of the output gap and inflation to a permanent increase in multi-factor productivity are also in general accordance with estimates from the VAR literature of the effects of technology shocks.

Whether SVARs can be used effectively for such comparisons depends crucially on whether they can closely match the impulse responses of the RM. Having only partial information when estimating a SVAR is a potentially important factor influencing the ability to make such a match.

This first source of loss of information that we examine is essentially due to scale - there may be more variables in the RM than in the SVAR, as SVARs typically only include measured variables. Although the solution to most RMs is a finite-order VAR in *all* of the variables in the model, only a sub-set of these variables may be observed. The solution in the observed variables alone may instead be a Vector AutoRegression Moving Average (VARMA) process. Fernández-Villaverde et al. (2007) and Ravenna (2007) present conditions whereby ignoring the unobserved variables could still result in a finite-order VAR underlying the SVAR, i.e. for there to be a lack of *truncation bias* in the VAR.[2]

Satisfaction of these conditions for a finite-order VAR representation is likely to be rare in practice and the relevant issue is then how good the approximation is. Using as the RM an economy-wide model that was a miniature version of the Bank of England's model in the 2000s and assuming that the contemporaneous responses were known, Kapetanios et al. (2007) found a very high-order SVAR to be necessary in order to precisely replicate the RM's impulse response functions. Pagan and Robinson (2016) and (2019) looked at many DSGE models and found that generally a VAR(2) was a reasonable approximation to the

[2] A survey of this literature is Giacomini (2013).

correct VAR even if it was of infinite-order, i.e. the missing variables could be captured by the observables and a finite number of their lags. Exceptions were when there was a stock of foreign debt in small-open economies (as was true in Kapetanios et al.'s case) and the use of concepts in RMs such as flexible-price equilibria in the RM, as in Smets and Wouters (2007). Liu et al. (2018) found that the large multi-sector DSGE model of Rees et al. (2016) could be well approximated with a VAR(2).

An increasingly relevant case in which truncation issues arise is when the RM contains a mixture of transitory and permanent shocks and therefore some of the observed variables are I(1). RMs typically transform these into I(0) variables in the model by dividing by the level of (latent) technology. Ravenna (2007) found that the VAR in observable variables involving growth rates for the I(1) variables exhibited a truncation bias, impacting considerably on the estimated impulse responses. Poskitt and Yao (2017) showed the same outcome and provided a decomposition of the bias. In Section 2 we explain that a source of the bias is a mis-specification of the VAR. The VAR in the observed variables does not capture the fact that there is a further I(1) latent variable in the system that needs to be accounted for - typically the log level of technology. Consequently, a process in the observable variables alone has structural equations that omit an unobservable error-correction (EC) term. An important aspect to note is that the latent process - the log level of technology - is strictly exogenous in most RMs, and therefore can be accommodated using standard state-space filtering methods. We show that by fitting a latent-variable VECM the bias in estimating the impulse responses can be substantially reduced.

The presence of a latent EC term in the VECM has a number of further implications for applied macroeconomic modelling. In particular, it points to difficulties in using either the Blanchard and Quah (1989) or the Shapiro and Watson (1988) methods of imposing long-run restrictions as well as suggesting problems in the construction of DSGE-VARs, which use DSGE models as a source of prior information for the VAR.

Unobserved variables are not the sole source of partial information. A second one is that SVARs often do not fully use the restrictions that are embedded in the RM. One such restriction is that structural shocks in DSGE models are typically assumed to be univariate autoregressive processes. We demonstrate that this is an important identifying assumption for the SVAR representation of the RM as it produces a common factor in the lag polynomials. Hence Section 3 explores what happens if this assumption about the statistical nature of shocks in RMs is exploited when estimating a SVAR. A small scale New-Keynesian model is used as the RM to examine the importance of these common-factor restrictions. We find that they can be very important in the identification of the SVAR and deserve to be used more extensively.

A further aspect of partial information is when there is not enough information to estimate all the structural shocks in the RM. Section 4 examines this. Denoting the observable-variables VAR shocks as η_t, and ε_t as the more numerous structural shocks, Ravenna (2007) says "..it will not be possible to map η_t into a higher-dimension vector of orthogonal shocks". In this situation the estimated ε_t end up being correlated, so as to reduce the rank of the covariance matrix of ε_t down to that of the covariance matrix of η_t. The structural shocks being non-orthogonal is not inconsequential; one can't use these shocks in variance decompositions and it also makes impulse responses difficult to interpret. Often this situation arises from the specification of the RM itself, e.g. in Christiano et al. (2014) there are 12 observed variables but 20 exogenous innovations, including structural shocks and news about risk, so their estimated versions must be correlated. Another example is that of unobserved component models where there are more shocks than observables. We look at a simple case of this and show that the common belief that the Kalman filter can separate the shocks is incorrect; the estimated shocks will be correlated.

A common concern in applied macroeconomic research is that the observed data may be measured with error. Measurement error is another source of partial information, although in this case it is initiated by the researcher, and Section 5 examines its consequences. We demonstrate that, when there are permanent shocks, and therefore some of the observed variables are I(1), the common way that measurement error is allowed for in estimation may have unintended consequences, as it implies a lack of co-integration between variables, regardless of whether this is true or false. We present alternative specifications of the nature of the measurement error that can be easily applied and preserve any co-integrating relationships.

In summary, the contribution of this article is to discuss four dimensions of partial information that are relevant for applied macroeconomic modelling and to present ways by which these issues can be practically addressed.

2 VAR Truncation Bias Due to I(1) Variables in the RM

2.1 Analysis

It has become common for DSGE models to include permanent shocks. An example is if the log-level of technology, a_t, is assumed to follow a unit root. As the economy in this case will have a long-run growth path, a normalisation has to be used prior to log-linearisation of the model for estimation. For example, consider the consumption Euler equation with log utility, namely $C_t^{-1} = \beta E_t C_{t+1}^{-1} R_{t+1}$, where C_t is the level of consumption, β the discount factor and R_t is a real interest rate. The appropriate normalisation would be $A_t \equiv \exp(a_t)$, and the Euler equation would become $(\frac{C_t}{A_t})^{-1} = \beta R^* E_t(\frac{C_{t+1}}{A_{t+1}})^{-1}(\frac{A_{t+1}}{A_t})^{-1}\frac{R_{t+1}}{R^*}$, with R^* the steady-state interest rate. Then, after log-linearisation, we have

$$c_t - a_t = E_t[c_{t+1} - a_{t+1} + \Delta a_{t+1}] - E_t r_{t+1} + r^*, \tag{1}$$

where the lower case letters represent the logs of the upper case ones.

It is apparent that the inclusion of a permanent technology shock results in some variables in these models being I(1) and co-integrated. An example of such co-integration is that between c_t and a_t, since $c_t - a_t$ is I(0) in Equation (1). All the model variables which are I(1) are expressed as deviations from a_t, a process often referred to as "stationising". These "stationised" variables are $I(0)$ and represent error-correction terms.

Now the solution to the RM will be a VAR in the core variables, z_t. These are the smallest set of variables needed to summarise the RM. They may be smaller than the total number of endogenous variables since typically some can be substituted out using, for example, identities. Some of these core variables will be the *stationised variables*, such as $c_t^s = c_t - a_t$. These are only partially observed. Thus, when Δc_t is the observed data on consumption growth, the relation to the model variable c_t^s is $\Delta c_t = \Delta(c_t - a_t) + \Delta a_t = \Delta c_t^s + \Delta a_t$.

2.2 Origin of the Potential Truncation Bias

Poskitt and Yao (2017) and Ravenna (2007) studied the extent of truncation bias for an RM that was a simple Real Business Cycle (RBC) model with non-stationary technology. They have two core observed variables - log of output, y_t which is I(1) and hours, which is $I(0)$.[3] After stationising output we have y_t^s and h_t. Using their parameters we simulated their RBC model and found by regression that a VAR representation of the solution to the RBC model in the two variables y_t^s and h_t is [4]

$$\Delta y_t^s = -.01h_{t-1} - .18y_{t-1}^S - 1.56\varepsilon_t^h - .037\varepsilon_t^a \tag{2}$$

$$h_t = .93h_{t-1} - .24y_{t-1}^S - 2.4\varepsilon_t^h + .48\varepsilon_t^a, \tag{3}$$

where ε_t^a and ε_t^h are the innovations in the two shocks of the model, technology and labour supply. Because Δy_t^s is not observed we can replace it with $\Delta y_t - \varepsilon_t^a$, and Equation (2)

[3] As the latent capital stock, consumption and investment can be substituted out from their RM, the only core endogenous latent variables are output and hours. In some DSGE models utilization rates of capital are present and one cannot eliminate the capital stock.

[4] These equations are identities as they show the solution in terms of the lagged variables *and* the shocks. The R^2 from the regressions (which include the shocks as explanatory variables) is unity. When a finite-order VAR representation exists, in large models it is simplest to use this regression-based technique to find the VAR relations. Alternatively, in this simple case one could derive it analytically.

becomes

$$\Delta y_t = -.01h_{t-1} - .18y^S_{t-1} - 1.56\varepsilon^h_t + .963\varepsilon^a_t. \quad (4)$$

Note that the lagged error correction term y^s_{t-1} is in both equations of the VAR. It cannot be eliminated. If one just fits the equations in the observable variables, Δy_t and h_t, there will be mis-specification and so a bias in the coefficients. Omitting y^s_{t-1} from Equation (4) yields an estimate of the coefficient of h_{t-1} of -.097 (compared with -0.01), while for Equation (3) it produces .83, rather than .93. Consequently, the impact of shocks on Δy_t will correctly die out very quickly, even though there is a bias in the VAR parameter estimates. In contrast, the bias in the parameter estimates for h_{t-1} in Equation (3) results in a very different decay rate for the impulse responses of h_t, when compared to those from the RBC model.

Figure 1 shows the divergence between the impulse response of hours to a technology shock obtained from a VAR using the observed and all RBC model, i.e. including y^s_{t-1}. This difference is much the same as reported in Poskitt and Yao (2017). As discussed above the divergence is due to bias in the h_{t-1} coefficient estimate.

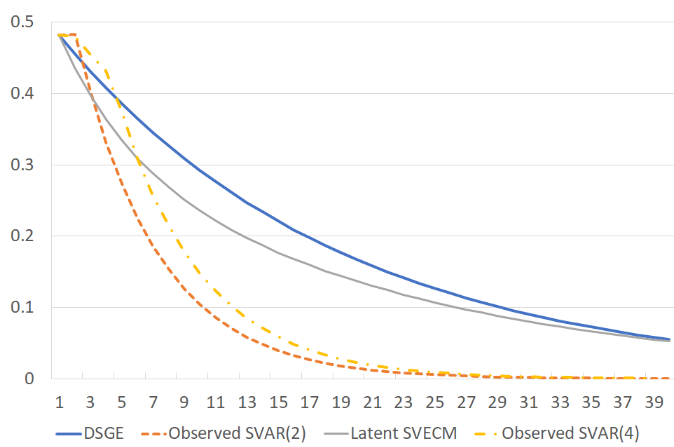


Figure 1: Impulse Response of Hours to a Technology Shock Using: (i) the RM of Poskitt and Yao (2017) (DSGE); (ii) a VAR(2) in the Observed Variables; (iii) a Latent-variable VECM and (iv) a VAR(4) in the Observed Variables

Because the issue is one of bias due to mis-specification, y^S_{t-1} needs to be included in some way in the regression. Adding in lagged observables $\{h_{t-j}, \Delta y_{t-j}\}_{j=1}^M$ is a way of doing this. This involves increasing the order of the observables VAR. Regressing y^S_{t-1} upon

$\{h_{t-j}, \Delta y_{t-j}\}_{j=1}^{M}$ we find that, when $M = 2$, the R^2=.67 ; when $M = 20$ it becomes .85 and when $M = 50$ it is .996. Hence a very high-order VAR would be needed to reconstruct the latent variable from the observed variables.

Differing conclusions have been found over the importance of truncation bias in the literature; for example, Chari, Kehoe and McGrattan (2005) argued that SVARs could not capture the impulse responses of the RM, whereas Christiano, Eichenbaum and Vigfusson (2007) were more optimistic. Differing RMs were used in these cases, and our analysis suggests that whether they include a permanent shock is likely to be important, as it introduces a latent variable which needs to be accounted for in the SVAR, and not just by imposing long-run restrictions. Indeed, Chari et al. (2005) have such an RM, whereas the more optimistic Christiano et al. (2007) used a stationary model. To illustrate that it can make a substantial difference, Figure 2 is the same as Figure 1, but is now for a variant of the Ravenna (2007) and Poskitt and Yao (2017) model where the stochastic component of technology, which had a unit root, is replaced with a persistent, but stationary first-order autoregressive process. The autocorrelation coefficient is set to 0.9; all other parameters are unchanged. It is apparent that truncation bias is much less important in the stationary case; the impulse responses to the labour supply shock from a SVAR(2) are are also a good match to those from the RBC model.

Returning to the case when the RM includes permanent shocks, since adding lags to the VAR is unlikely to deliver a good match without using an impractically high lag length, what other strategies are available? One is to recognise that the system is a *latent-variable VECM*, and to augment it with an assumption about the latent exogenous technology process (as is done in DSGE models). Then the latent VECM can be estimated as a state-space model using the Kalman filter. We assume that technology follows an I(1) process, consistent with the RM that Ravenna (2007) and Poskitt and Yao (2017) use. Note that more generally specifying the process for a latent VECM is straightforward as it is a strongly exogenous permanent shock, so the only decision is whether to allow for persistence in its growth rate.

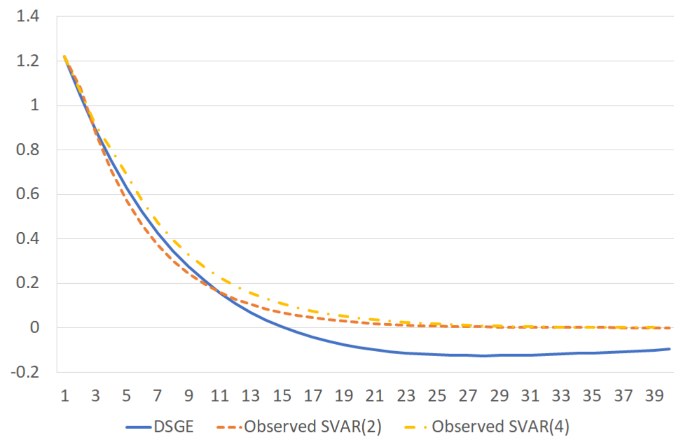


Figure 2: Impulse Response of Hours to a Technology Shock Using: (i) a variant of the RM of Poskitt and Yao (2017) (DSGE) with stationary technology; (ii) a VAR(2) in the Observed Variables and (iii) a VAR(4) in the Observed Variables

The implied state-space model in this instance is provided in Appendix A. As the focus is on the truncation bias, we impose the contemporaneous impulse responses, C_0, as in Ravenna (2007).

Figure 1 shows the impulse response of hours worked to a technology innovation from this latent variable VECM, which closely matches the impulse response from the RM. So estimating a latent-variable VECM seems to be a potentially useful strategy when the RM is believed to have non-stationary technology.

Often one sees a VECM being used that incorporates only *observed* EC terms. An example is Del Negro, Schorfheide, Smets and Wouters (2007), who introduce an I(1) technology shock, a_t, into the Smets and Wouters (2007) model and estimated an observed-variable VECM based upon it. As before, the normalised variables include the EC terms $(y_t - a_t), (c_t - a_t)$ and $(i_t - a_t)$. These can be written equivalently as $(y_t - a_t), (c_t - y_t)$, and $(i_t - y_t)$, with the last two being observed. There will *be one unobserved EC term missing from any observed-variables VECM.*[5] In Del Negro et al. (2007) it would be $(y_t - a_t)$, as output cannot be normalised by itself.[6]

[5]The presence of an unobserved EC term was also discussed in Liu, Pagan and Robinson (2018) in the context of the Rees, Hall and Smith (2016) model.

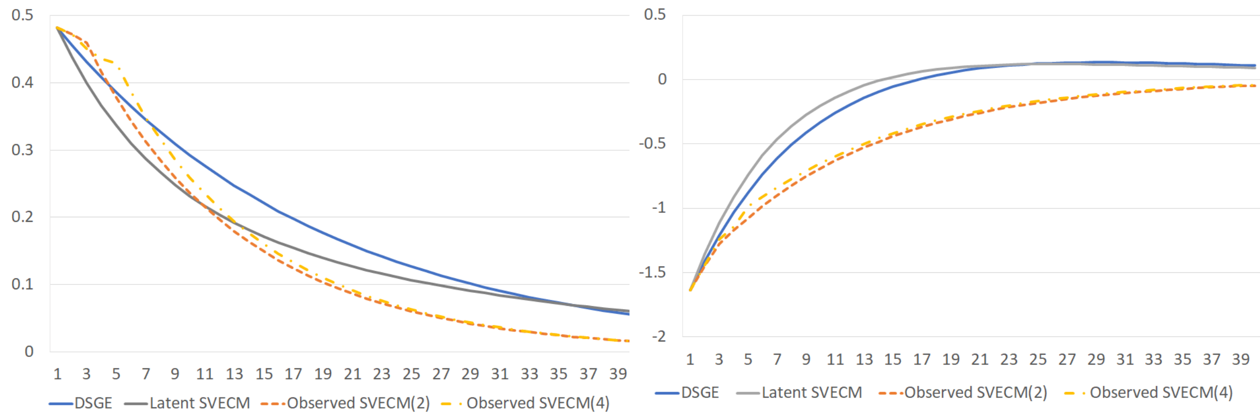
[6]It is not possible to eliminate this term using the resource constraint.

To assess the consequences of using an observed, rather than latent, variable VECM we add to the Ravenna (2007) model an extra shock and observed variable. These are a preference shock to the period utility function, b_t, and an assumption that investment growth is observed. These choices mean that there is now an observed EC term, $(i_t - y_t)$, as well as a latent one, $(y_t - a_t)$. As is standard, it is assumed that the preference shock follows a first-order autoregressive process. We parameterise it to match the same data moments as Ravenna (2007) used to calibrate the original model. This gives $\rho_b = 0.9$ and a standard deviation of its innovation of 0.006.

In some instances, such as the responses of all variables to a labor supply shock, the omission of the latent error-correction term is of little consequence. However, Figure 3 compares impulse response of hours to a technology shock (left-hand side) and a preference shock (right-hand side) when the observed and latent-variable VECMs are used. The latent-variable VECM provides a superior match, particularly for the response to a preference shock. This result reinforces our contention that latent-variable VECMs are a class of models which are likely to be useful to applied researchers when working with systems that contain a mixture of I(1) and stationary variables. A practical question faced in implementing a latent-VECM is how many latent, strictly exogenous, permanent processes to include. The DSGE literature here provides a guide - where typically one, or at most two, permanent shocks are included.

2.3 Further Implications for Applied SVAR Research

The presence of the latent EC term has further implications for applied SVAR research. Returning to the original RBC model, Ravenna noted that it has the same format as that of Blanchard and Quah (1989). Thus one might think about getting impulse responses to the shocks by imposing the long-run restriction that there is one permanent and one transitory shock. However to do this one needs to get good estimates of the VAR coefficients B_j, so as to form their sum $B(1)$, from which the long run restrictions can be derived. Because of the



(a) Hours to a Technology Shock

(b) Hours to a Preference Shock

Figure 3: Impulse Responses of Hours to a Technology or Preference Shock Using: (i) the RM of Poskitt and Yao (2017) (DSGE); (ii) a VECM(2) in the Observed Variables; (iii) a VECM(4) in the Observed Variables and (iv) a Latent-variable VECM

mis-specification the estimated B_j will be biased and so will be $B(1)$. Hence the restrictions being imposed would be incorrect unless one had a very high order approximating VAR.

A similar issue applies to what is normally an identical way of imposing long-run restrictions, the method due to Shapiro and Watson (1988). In their context this would involve first fitting the model

$$\Delta y_t = a_{12}^0 \Delta h_t + a_{11}^1 \Delta y_{t-1} + \varepsilon_t, \tag{5}$$

using h_{t-1} and Δy_{t-1} as instrumental variables, and then using the residuals $\hat{\varepsilon}_t$ to estimate the second structural equation for h_t. Now the exact SVAR equation for Δy_t is

$$\Delta y_t = .135\Delta h_t - .150 y_{t-1}^s + \varepsilon_t, \tag{6}$$

from which one can see that it also has a term y_{t-1}^s. This would be missing in Equation (5) when applying Shapiro and Watson's method. Omitting it yields

$$\Delta y_t = .523\Delta h_t + .006\Delta y_{t-1}, \tag{7}$$

so there is a very significant bias in the coefficient of Δh_t, which impacts upon estimates of the contemporaneous impulse responses.

3 Partial Information and Identification: Estimating Contemporaneous Impulse Responses

3.1 Analysis

There are two concerns when making a match of the RM and SVAR impulse responses. Structural impulse responses from any linear model, C_j, will be generated by $C_j = B_1 C_{j-1} + .. + B_p C_{j-p}$ if there is a VAR(p) with coefficients B_j. Hence to estimate these accurately one has to be able to estimate the contemporaneous responses C_0. In Section 2 we largely set C_0 to that from the RM, since we wanted to look at any bias in the B_j arising from mis-specification. Here we investigate a second complication that can arise when there is partial information, namely when not all of the information embedded in the RM is used in the estimation of C_0 within a SVAR.

Some of the assumptions in the structural equations of RMs - e.g. exclusion restrictions, come from an economic model. Alternatively, others are purely statistical, e.g. that structural shocks are uncorrelated and the shocks are univariate ARs. Researchers estimating a SVAR in the observed variables often seek to have a great deal of flexibility in the dynamics, which can create issues when trying to estimate the contemporaneous responses that are consistent with the RM, as they are not exploiting some of the statistical assumptions of those models. As we will see such restrictions could be used in the identification of the SVAR, although this has rarely been done in practice.

A typical structural equation in a RM, say for the variable y_t, has the form

$$y_t = \alpha_1 E_t(y_{t+1}) + \alpha_2 x_t + \alpha_3 y_{t-1} + u_{yt}, \tag{8}$$

where x_t are other variables and u_{yt} the structural shock. The issue is how to handle

$E_t(y_{t+1})$. The solution of the RM is taken to be a VAR(2) in its endogenous variables z_t, so $E_t(z_{t+1}) = B_1 z_t + B_2 z_{t-1}$. Letting $y_t = S z_t$, where S is a selection matrix, Equation (8) can be expressed as

$$y_t = \alpha_1 S B_1 z_t + \alpha_1 S B_2 z_{t-1} + \alpha_2 x_t + \alpha_3 y_{t-1} + u_{yt}, \tag{9}$$

and this constitutes a SVAR equation. The question then becomes whether we have enough instruments to estimate the parameters of this equation. That there may be enough comes from the assumption often used in RMs that u_{yt} is a univariate AR(1), since this introduces non-linear restrictions between the parameters of the equation to be estimated.

To illustrate this we take as the RM a simple small New-Keynesian (NK) model of the form

$$\begin{aligned} y_t &= E_t(y_{t+1}) - (r_t - E_t(\pi_{t+1})) + u_{yt} && (10)\\ \pi_t &= \beta E_t(\pi_{t+1}) + \kappa y_t + u_{\pi t} && (11)\\ r_t &= \rho_r r_{t-1} + (1-\rho_r)(\gamma_y y_t + \gamma_\pi \pi_t) + \delta \Delta y_t + \varepsilon_{rt}, && (12) \end{aligned}$$

where u_{yt} and $u_{\pi t}$ follow AR(1) processes with parameters ρ_y and ρ_π and have innovations ε_{yt} and $\varepsilon_{\pi t}$. All the innovations are assumed to be white noise processes and uncorrelated with each other. This is the external sector of the Reserve Bank of Australia's Multi-Sector Model (MSM), set out in Rees et al. (2016).[7] We assume that π_t and r_t are observed. y_t in the MSM is actually a stationised variable, however, we will abstract from that complication and think of y_t as being an $I(0)$ observable output gap, in order to focus on the problems of matching the estimated C_0^{SVAR} with C_0^{RM}.

Consider the New-Keynesian Phillips Curve in Equation (11). Now because the MSM solution is a VAR(1) in all variables expected inflation will be

$$E_t(\pi_{t+1}) = b_{21}^1 y_t + b_{22}^1 \pi_t + b_{23}^1 r_t \tag{13}$$

[7] We have dropped the price of commodities from their external sector for simplicity; it does not influence the other variables.

where b_{ij}^1 $(i = 1, ., 3; j = 1, .., 3)$ denotes the elements in B_1. Using the estimated parameter values in Rees et al. (2016) this would be

$$E_t(\pi_{t+1}) = .108y_t + .269\pi_t + .123r_t. \tag{14}$$

Once we replace $E_t(\pi_{t+1})$ in Equation (11) with Equation (13) and re-arrange, the equation will have the form

$$\pi_t = a_{21}^0 y_t + a_{23}^0 r_t + v_{\pi t}, \tag{15}$$

where $a_{21}^0 = \frac{\beta b_{21}^1 + \kappa}{1 - \beta b_{22}^1}, a_{23}^0 = \frac{\beta b_{23}^1}{1 - \beta b_{22}^1}$, $v_{\pi t} = \rho_\pi v_{\pi t-1} + \varepsilon'_{\pi t}$ and $\varepsilon'_{\pi t} = \frac{\varepsilon_{\pi t}}{1 - \beta b_{22}^1}$. $v_{\pi t}$ is just the original cost-push shock re-scaled. Re-writing the shock process with the lag operator gives $v_{\pi t} = (1 - \rho_\pi L)^{-1} \varepsilon'_{\pi t}$, and so Equation (15) becomes

$$\pi_t = a_{21}^0 y_t + a_{23}^0 r_t + (1 - \rho_\pi L)^{-1} \varepsilon'_{\pi t}, \tag{16}$$

and hence

$$(1 - \rho_\pi L)\pi_t = (1 - \rho_\pi L)a_{21}^0 y_t + (1 - \rho_\pi L)a_{23}^0 r_t + \varepsilon'_{\pi t}. \tag{17}$$

It is evident that the coefficients on π_{t-1}, y_{t-1} and r_{t-1} all involve the same parameter ρ_π. Consequently, there is a *common factor* $(1-\rho_\pi L)$ in the three separate lag polynomials arising from a *statistical assumption* about the autoregressive nature of shocks. This common factor (COMFAC) structure was investigated by Hendry and Mizon (1978).

3.2 Using COMFAC Restrictions from the RM in SVAR Analysis

A standard equation for inflation in a SVAR(1) containing y_t, π_t and r_t would be

$$\pi_t = a_{21}^0 y_t + a_{23}^0 r_t + a_{22}^1 \pi_{t-1} + a_{21}^1 y_{t-1} + a_{23}^1 r_{t-1} + \varepsilon_{\pi t}. \tag{18}$$

There are five parameters to be estimated in the conditional mean of this but only three instruments, namely y_{t-1}, r_{t-1} and π_{t-1}. When COMFAC restrictions are applied, Equation (18) becomes Equation (19) (which is equivalent to Equation 17), and there are only three parameters in that conditional mean, i.e. it is exactly identified.

$$\pi_t = a_{21}^0 y_t + a_{23}^0 r_t + \rho_\pi u_{\pi t-1} + \varepsilon_{\pi t}. \tag{19}$$

To demonstrate how important the COMFAC restrictions may be to estimation, note that the parameters from Rees et al. (2016) imply the SVAR would have coefficient values of $a_{21}^0 = 0.196$, $a_{23}^0 = .168$ and $\rho_\pi = .32$. Estimating Equation (19) using simulated data from the MSM (10,000 observations) and the instruments π_{t-1}, y_{t-1}, and r_{t-1}, we obtain the estimates $\widehat{a_{21}^0} = 0.192$, $\widehat{a_{23}^0} = .164$ and $\widehat{\rho_\pi} = 0.31$, which are an excellent match to the true values, enabling the recovery of the MSM shock $\varepsilon_{\pi t}$.

In contrast, if one was to instead identify the SVAR by assuming that it is recursive, this would require assuming that the interest rate does not appear in the inflation equation (i.e. $a_{23}^0 = 0$). While it is true that r_t is absent from the MSM's Phillips curve, it should appear in the SVAR equation for inflation because of expectations (the implied $a_{23}^0 = .168$). So, for the inflation equation, the COMFAC restriction appears preferable to assuming it is recursive. It should be noted, however, that the COMFAC restriction does not necessarily produce strong instruments.

3.3 Implications for Applied SVAR Research

The analysis and illustration show the difficulties that a SVAR can experience in capturing the C_0 from a DSGE model often stem from the fact that the traditional estimation methods for SVAR models seek to avoid imposing statistical restrictions (such as COMFAC), exclusion restrictions (in the interest rate equation π_{t-1} does not appear in the MSM), and other constraints where coefficients are prescribed (for example that on $E_t(y_{t+1})$ in the MSM output equation). In many ways SVARs are about assembling information concerning the dynamics and contemporaneous interactions between variables in the macroeconomy in a way which, while identified, imposes less structure than is included in RMs. Traditionally this flexibility has been achieved by using exactly-identified SVARs, rather than the over-identified structural equations of the DSGE approach. There are, of course, common restrictions between the two approaches, such as the assumption that the structural shocks are uncorrelated.

As we have demonstrated, the COMFAC restrictions, being statistical in nature, can be implemented in a SVAR. Given their widespread usage in DSGE modelling they deserve to be used more widely in SVAR analysis. A caveat, however, is that applying multiple COMFAC restrictions can result in the SVAR being over-identified. To see this, suppose one imposes COMFAC restrictions on the SVAR model $A_0 y_t = A_1 y_{t-1} + u_t, u_t = \Phi u_{t-1} + \varepsilon_t$, where Φ is diagonal and $var(\varepsilon_t) = I$. When there are no other restrictions on A_j there are $2n^2 + n$ parameters to estimate. Since the model implies a VAR(2) there are $2n^2 + \frac{n(n+1)}{2}$ parameters in it. Hence the COMFAC restrictions result in an over-identified SVAR as $n < \frac{n(n+1)}{2}$. The system would be exactly identified if Φ was triangular.

4 Partial Information and Structural Shock Estimation

Estimating the structural shocks is a key component of modern macroeconomics. However, having only partial information may impede this. In particular, there are models where the number of structural shocks exceeds the number of observed variables, and as Ravenna (2007) noted, in this case the shocks cannot all be recovered. At most we can recover the same number of shocks as observed variables (or that many linear combinations of them). This is because the matrix connecting the vector of observed variables and the vector of shocks is not square. Canova and Ferroni (2018) study a similar issue, which they term aggregation, namely when the RM has more shocks than the estimated SVAR, and demonstrate that the estimated impulse responses are combinations of the actual impulse response functions in the RM.

To understand the consequences of partial information for estimating the structural shocks we look at a RM with I(1) variables placed in a State-Space Form (SSF) as follows

$$
\begin{aligned}
z_t &= D_1\psi_t + D_2\psi_{t-1} + Ru_t && (20) \\
\psi_t &= M\psi_{t-1} + Cu_t, && (21)
\end{aligned}
$$

where Equations (20) and (21) are the observation and state equations respectively. z_t are the observed variables, ψ_t the model variables and u_t the shocks; while these may include both structural shocks and measurement error we focus on the former as the implications of measurement error are studied in the next section. For example, the RBC model of Section 2 in this SSF is

$$
\begin{aligned}
z_t &= \begin{bmatrix} \Delta y_t \\ h_t \end{bmatrix}, \psi_t = \begin{bmatrix} y_t^S \\ h_t \end{bmatrix}, D_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, D_2 = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \\
u_t &= \begin{bmatrix} \varepsilon_t^a \\ \varepsilon_t^h \end{bmatrix}, R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.
\end{aligned}
$$

Nimark (2015) derives the Kalman filter for a SSF with a lag of the state in the observation equation, as in Equation (20). We use this purely to simplify the exposition; one could alternatively expand the state vector and use the standard recursions. He shows that the filtered estimate of ψ_t given the data to time t, $E_t\psi_t$, evolves as

$$
\begin{aligned}
E_t\psi_t &= \Phi_t E_{t-1}\psi_{t-1} + K_t z_t && (22)\\
K_t &= [MP_{t-1|t-1}\Psi' + CC'D_1' + CR'][\Psi P_{t-1|t-1}\Psi' + \Lambda\Lambda']^{-1} && (23)\\
P_{t|t} &= P_{t|t-1} - K_t[\Psi P_{t-1|t-1}\Psi' + \Lambda\Lambda']K_t' && (24)\\
P_{t+1|t} &= MP_{t|t}M' + CC' && (25)\\
\Psi &= D_1 M + D_2, \Lambda = D_1 C + R, \Phi_t = M - K_t\Psi, && (26)
\end{aligned}
$$

where K_t is the gain of the Kalman filter (Equation 23). Equation (22) therefore is the updating equation, and $P_{t|t}$ the corresponding Mean Squared Error matrix (Equation 24). Equation (25) provides the one-step-ahead prediction of the latter.

DSGE models exist with more exogenous innovations than observed variables - an example being Christiano et al. (2014). A much simpler instance where this often occurs is the unobserved components models that are used to measure output gaps; see, for example, Orphanides and van Norden (2002). Such models decompose a series y_t as $y_t = y_t^p + y_t^T$, where y_t^p is a permanent component of y_t and y_t^T a transitory component. Assumptions have to be made about how these evolve and we look at the simplest set:

$$
\begin{aligned}
\Delta y_t^p &= e_t^p\\
y_t^T &= e_t^T\\
\Delta y_t &= e_t^p + \Delta y_t^T,
\end{aligned}
$$

where e_t^p and e_t^T are $n.i.d.(0,1)$ and independent of one another. This model implies that Δy_t is a MA(1), $\Delta y_t = (1+\alpha L)\varepsilon_t$, and so there are two parameters that can be estimated from the data - α and σ_ε^2- to give estimates of the variances of e_t^p and e_t^T.

Now, to estimate the two shocks e_t^p and e_t^T from the one observed variable Δy_t we put the model into its SSF. So $\psi_t = \begin{bmatrix} \Delta y_t^p \\ y_t^T \end{bmatrix}$, and the matrices are $M = 0, D_1 = \begin{bmatrix} 1 & 1 \end{bmatrix}, D_2 = \begin{bmatrix} 0 & -1 \end{bmatrix}, C = I, R = 0$. Using Equation (22), the estimated filtered shocks evolve as

$$E_t\psi_t = -K_t D_2 E_{t-1}\psi_{t-1} + K_t z_t.$$

Equation (23) gives the filter gain

$$K_t = D_1'[D_2 P_{t-1|t-1} D_2' + D_1 D_1']^{-1},$$

and

$$(D_2 P_{t-1|t-1} D_2' + D_1 D_1')^{-1} = (P_{22,t-1|t-1} + 2)^{-1},$$

so this is a scalar c_t. Therefore $K_t = \begin{pmatrix} c_t \\ c_t \end{pmatrix}$. Now $K_t D_2 = \begin{bmatrix} 0 & -c_t \\ 0 & -c_t \end{bmatrix}$ so that $E_t\psi_t$ changes as

$$\begin{aligned} E_t\psi_{1t} &= c_t E_{t-1}\psi_{2t-1} + c_t z_t \\ E_t\psi_{2t} &= c_t E_{t-1}\psi_{2t-1} + c_t z_t. \end{aligned}$$

Consequently $E_t\psi_{1t} = E_t\psi_{2t}$, which is not true of the shocks e_t^p and e_t^T.[8] While the Hodrick-Prescott (1997) filter assumes that the cyclical component is white noise, many unobserved-components models allow for some persistence. Doing so, and repeating the analysis above, still yields estimates of the shocks that correlated when the true shocks are not - see Appendix B.

To examine the consequences in a DSGE context, we return to Equation (22). In a

[8]Morley (2011) discusses different interpretations of the Beveridge-Nelson filter. He notes that when the Beveridge-Nelson filter is interpreted as providing an estimate of the unobserved trend, if the transitory shock has a positive variance it is an unobserved components model and the innovations to the trend and transitory components will be (imperfectly) correlated.

stationary DSGE the initial states $E_0\psi_0$ are set to zero, so that $E_1\psi_1 = K_1 z_1$ and, given that K_1 has no more rank that $\dim(z_t)$, it must be that a linear combination of the estimated states $E_1\psi_1$ equals z_1, and so the shocks must be correlated. Once this is true of the first estimated state it is true of all later ones. When there are just two shocks there is perfect correlation but, when there are more than two, the correlations need not be unity, although they must be non-zero. It is the fact that we cannot separate the estimated shocks that leads to the problems with their use in variance decompositions and impulse responses, as both assume the shocks are uncorrelated. Alternatively, if the intended use of the model is prediction then correlated shocks are inconsequential. There are two obvious practical solutions - either use more observed data in estimation or revisit the design of the model.

5 Unintended Consequences of Measurement Error

Partial information can also occur by design. This happens in the literature when a researcher *makes the assumption* that the data and model variables are not the same. This could reflect concerns that the data are measured with error, or that there is a conceptual difference between the model variables and their observed counterparts. These divergences are typically addressed by including auxiliary equations which are assumed to connect the model variables and data with the inclusion of additional shocks, a process which is often referred to as "adding measurement error". That term, however, seems to imply that the "theory is ahead of data", and that the model is correct and the data is wrong; a more neutral description of these additional shocks is that they provide a *reconciliation* between the model variables and the data. An aspect to note is that while in the previous section we focused on structural shocks, if the total number shocks - structural or measurement - exceeds the number of observed variables then the estimated shocks must be correlated. Leaving potential correlation aside, the key issue is the appropriate specification of these reconciliation shocks, and we show that considerable care must be taken when the model variables are I(1).[9]

5.1 Model Variables are $I(1)$

A common approach to introducing measurement error in the case where variables are I(1) is to write $\Delta z_t^D = \Delta z_t^M + \eta_t$, where η_t is $I(0)$. A recent application in this vein is Aruoba et al. (2016). They considered that different measures of GDP growth could be regarded as deviating from true GDP growth, and that these deviations would be described as measurement errors. Many factor models make a similar assumption. In terms of models with a greater economic emphasis, DSGE models often proceed in this way, e.g. Guerron-Quintana (2010), including those used at policy institutions, e.g. Chung et al. (2010) and Rees et al. (2016). In this case the model variables are the first difference of the stationised variables, together with the technology growth, i.e. $\Delta z_t^D = \Delta z_t^s + \Delta a_t + \eta_t$. This wide usage motivates our analysis of the implications of this dimension of partial information.

We start with a simple situation which parallels that in Aruoba et al. (2016). Let Δz_t be the true growth rate in GDP and Δz_{jt} $(j = 1, 2)$ be two noisy measures of it. Then we have $\Delta z_{jt} = \Delta z_t + \eta_{jt}$, where η_{jt} are said to be the measurement errors, which Aruoba et al. (2016) assume to be white noise. Now consider what this means for the relation between the level of GDP and its measures. Selecting the first data series we have

$$\begin{aligned} z_{1t} - z_t &= \sum_{k=1}^{t} (\Delta z_{1k} - \Delta z_k) \\ &= \sum_{k=1}^{t} \eta_{1k}, \end{aligned}$$

assuming there is no measurement error in the initial period. Consequently, under the assumption that measurement errors are white noise it is clear that the data, z_{1t}, and model

[9] In Liu et al. (2018) we discuss how these results apply to the multi-sector model of Rees et al. (2016), drawing on one of the working papers of this article, Pagan (2017).

variable, z_t, do *not* co-integrate. Moreover, the difference

$$\begin{aligned} z_{1t} - z_{2t} &= \sum_{k=1}^{t} (\Delta z_{1k} - \Delta z_{2k}) \\ &= \sum_{k=1}^{t} (\eta_{1k} - \eta_{2k}), \end{aligned}$$

would also be $I(1)$.[10] Therefore, unless η_{1k} and η_{2k} are perfectly correlated, there would be no co-integration between either of the data variables z_{1t} and z_{2t}.[11,12]

One possible reconciliation of the observed and model variables is indeed that they are not co-integrated, but intuitively it would seem more satisfactory if at least one of the measured quantities did co-integrate with the true level of GDP. Therefore it doesn't seem sensible to rule this out when choosing the specification of the measurement error, and doing so may be an unintended consequence of using this standard specification.

One can, of course, test if the data z_{1t} and z_{2t} are co-integrated, which provides a check on whether the assumptions being made about the measurement error are reasonable. Aruoba et al. (2016) take z_{1t} to be the expenditure-based measure of GDP, while z_{2t} is the income-based series. We therefore test if they co-integrate; using a VAR(2), Johansen's trace and eigenvalue tests are equal and of value 3.22. As the 5% critical value is 3.84 there does indeed seem to be co-integration between the observed series.[13,14]

How can measurement error be introduced into a model in a way that maintains the co-integrating relationships? One could instead write $z^D_{jt} - z^M_{jt} = \eta_{jt}$ for all j elements of z^D_t, and assume that η_{jt} are $I(0)$. Then $\Delta z^D_{jt} = \Delta z^M_{jt} + \Delta\eta_{jt}$, which clearly results in co-integration between model and data variables being maintained. The only remaining issues

[10] Once again assuming there is no measurement error in the initial period.

[11] If they were perfectly correlated then basically z_{1t} and z_{2t} would be the same series.

[12] This is also true if the I(1) data is filtered to produce I(0) processes that are used in models e.g. as an output gap. In those cases the filtered data will be weighted averages of growth rates in variables so it is an average of growth rates in the data that would be held to deviate from model growth. Thus, the issues we describe in this section also apply when filtered data are used, although the analysis is more complex.

[13] We thank Dongho Song for providing the data.

[14] It is interesting that the co-integrating vector seems to be (1 -.9978).

is to determine an appropriate specification for η_t. One possibility is to assume it is i.i.d. To see another possible definition, note that if we assume that the data and model variables co-integrate with the same co-integrating vectors, we would have equations for the model of the form

$$\Delta z_t^M = \delta\gamma' z_{t-1}^M + C e_t^M, \tag{27}$$

where δ and C are parameters, e_t^M the model shocks, and γ are the common co-integrating vectors.[15] Then, using the relationship between the observed data and the model counterparts, $\Delta z_t^D = \Delta z_t^M + \Delta\eta_t$, we obtain

$$(\Delta z_t^D - \Delta\eta_t) \;=\; \delta\gamma'(z_{t-1}^D - \eta_{t-1}) + C e_t^M,$$

and therefore

$$\Delta z_t^D = \delta\gamma' z_{t-1}^D + \eta_t - (1 + \delta\gamma')\eta_{t-1} + C e_t^M.$$

Now we would want the shocks in this VECM, e_t^η, to be i.i.d., and so

$$\eta_t = (1 + \delta\gamma')\eta_{t-1} + e_t^\eta - C e_t^M. \tag{28}$$

Hence in this second definition the measurement-error shocks η_t follows a VAR process with innovations constructed from e_t^η and the model shocks e_t^M.[16] Often one sees measurement-error shocks specified as univariate AR(1) processes but this leads to complicated VECM processes. A better specification is to utilise the form in Equation (28), which exploits model co-integrating information and the model shocks as well as the measurement error innovation e_t^η.[17]

[15] The model shocks will be assumed to be white noise processes, i.e. they are innovations, although they only need to be $I(0)$ processes for our analysis.

[16] Introducing lags of Δz_t^M into Equation (27) increases the order of the VAR in η_t.

[17] This can be generalised to the case where the co-integrating vectors among the observables are not the same as the model. In this case η_t has to compensate for that difference as well.

5.2 Implications for Applied Research

To implement either definition of η_t, in the state-space model the measurement equations would be $\Delta z_t^D = \Delta z_t^M + \Delta\eta_t$, with η_t being a state variable. The state equations include those for the model variables, namely Equation (27), and those governing the evolution of η_t, e.g. Equation (28).

6 Conclusion

In this paper we have analysed four dimensions along which partial information is important for applied macroeconomic modelling.

SVARs are typically estimated only in measured variables, whereas representative models of the economy, such as a DSGE model, often include variables that are not observed. Thus partial information may limit the ability of a SVAR to capture the dynamics of the economy, that is, to match the impulse responses of the RM, which is known as truncation bias. We focused on the increasingly common case where the RM includes a permanent shock. It was found that in this instance an important source of truncation bias was a mis-specification of the underlying VECM, as the actual generating process involves a latent error-correction term and not just observed variables. Using the Ravenna (2007) and Poskitt and Yao (2017) RBC model it was demonstrated that the quality of the match to the DSGE impulse responses can be considerably improved by estimating a latent-variable VECM.

Our findings suggest that an important factor influencing the divergent findings in the literature about the ability of SVARs to make a match with the impulse responses from DSGE models is whether or not the DSGEs being analysed include permanent shocks.

A second dimension to making a match is how well the SVAR can capture the initial impulse responses, rather than the dynamics. Once again this is influenced by partial information, as the RM may include identifying restrictions not commonly used in SVARs. In particular, our analysis showed that some of the identification conditions used in DSGE models arise from assumptions made about the shock processes. These are statistical, rather than economic, restrictions. They imply that there are common factors in the SVAR. It was demonstrated that using such restrictions when estimating the SVAR model can considerably improve the match of its initial impulse responses with those from the DSGE model.

Estimating structural shocks is a key aspect of applied macroeconomics, and the consequences of not having enough information to do so were examined. More precisely, it was shown that when the number of shocks exceed the the number of observed variables the estimated structural shocks will be correlated, even when the true shocks are not.

The final dimension of partial information studied was the common practice of treating model variables as deviating from the observed data due to measurement error. Our analysis highlighted that this strategy has to treated carefully when the variables are I(1), as the standard approach will imply a lack of co-integration between the data variables, even if they are co-integrated. It was demonstrated how to specify measurement-error processes that would preserve any co-integrating information.

References

Aruoba, B., F.X. Diebold, J. Nalewaik, F. Schorfheide and D. Song (2016), "Improving GDP Measurement: A Measurement Error Perspective" *Journal of Econometrics*, 191, 384-397.

Blanchard, O. J. and D. Quah (1989), "The Dynamic Effects of Aggregate Demand and Supply Disturbances", *American Economic Review*, 79, 4, 655-673.

Brayton, F., T. Laubach and D. Reifschneider (2014), "The FRB/US Model: A Tool for Macroeconomic Analysis", *FEDS Notes*.

Canova, F. and F. Ferroni (2018), "Mind the Gap! Stylized Dynamic Facts and Structural Models", mimeo.

Chari, V. V., P.J. Kehoe and E.R. McGrattan (2005), "A Critique of Structural VARs Using Business Cycle Theory", *Federal Reserve Bank of Minneapolis Working Paper*, 631.

Christiano, L.J., M. Eichenbaum and Robert Vigfusson (2007), "Assessing Structural VARs", in D. Acemoglu, K. Rogoff and M. Woodford (eds.), *NBER Macroeconomics Annual 2006*, 21, 1-106.

Christiano, L. J., R. Motto and M. Rostagno (2014), "Risk Shocks", *American Economic Review*, 104, 1, 27-65.

Chung, H, M.T. Kiley, and J-P Laforte (2010), "Documentation of the Estimated, Dynamic, Optimization-based (EDO) Model of the U.S. Economy", *Finance and Economics Discussion Series*, 2010-29, Board of Governors of the Federal Reserve System.

Del Negro, M., F. Schorfhiede, F. Smets and R. Wouters (2007), "On the Fit of New Keynesian Models", *Journal of Business and Economic Statistics*, 25, 2, 123-143.

Fernández-Villaverde, J., J. F. Rubio-Ramírez, T. J. Sargent and M. W. Watson (2007), "ABCs (and Ds) of Understanding VARs," *American Economic Review*, 97, 1021-1026.

Giacomini, R. (2013), "The Relationship Between VAR and DSGE Models", in T. B. Fomby, L. Kilian and A. Murphy (eds.), *VAR Models in Macroeconomics – New Developments and Applications: Essays in Honor of Christopher A. Sims*, *Advances in Econometrics*, 32, 1-25.

Guerron-Quintana, P. (2010), "What You Match Does Matter: The Effects of Data on DSGE

Estimation", *Journal of Applied Econometrics*, 25, 774–804.

Hendry, D. F. and G. Mizon (1978), "Serial Correlation as a Convenient Simplification, Not a Nuisance: A Comment on a Study of the Demand for Money by the Bank of England," *Economic Journal*, 88, 549-63.

Hodrick, R. and E. Prescott (1997), "Post-war Business Cycles: An Empirical Investigation", *Journal of Money, Credit and Banking*, 29, 1-16.

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

Liu, X., A. R. Pagan and T. Robinson (2018), "Critically Assessing Estimated DSGE Models: A Case Study of a Multi-sector Model", *Economic Record*, 94, 349-371.

Morley, J. C. (2011), "The Two Interpretations of the Beveridge-Nelson Decomposition", *Macroeconomic Dynamics*, 15, 3, 419-439.

Nimark, K. P. (2015) "A Low Dimensional Kalman Filter for Systems With Lagged States in the Measurement Equation", *Economics Letters*, 127(C), 10-13.

Orphanides, A. and S. van Norden (2002), "The Unreliability of Output-Gap Estimates in Real Time", *Review of Economics and Statistics*, 84, 4, 569-583.

Pagan, A. R. (2017), "Some Consequences of Using 'Measurement Error Shocks' When Estimating Time Series Models", *Centre for Applied Macroeconomic Analysis Working Paper*, Australian National University, 2017-12.

Pagan, A.R. & T. Robinson (2016), "Investigating the Relationship Between DSGE and SVAR Models", *NCER Working Paper Series 112*, National Centre for Econometric Research.

Pagan and & T. Robinson (2019), "Implications of Partial Information for Econometric Modelling of Macroeconomic Systems", *Centre for Applied Macroeconomic Analysis Working Paper*, Australian National University, 2019-41.

Poskitt, D. and Yao, W. (2017), "Vector Autoregressions and Macroeconomic Modelling: an Error Taxonomy", *Journal of Business and Economics Statistics*, 35, 3, 407-419.

Ravenna, F. (2007), "Vector Autoregressions and Reduced Form Representations of DSGE

Models", *Journal of Monetary Economics*, 54, 7, 2048-2064.

Rees, D.M., P. Smith and J. Hall (2016), "A Multi-sector Model of the Australian Economy", *Economic Record*, 92, 298, 374-408.

Shapiro M. and M. Watson (1988), "Sources of Business Cycle Fluctuations", in *NBER Macroeconomics Annual*, 3, 111-148.

Smets, F. and R. Wouters (2007), "Shocks and Frictions in US Business Cycles: a Bayesian DSGE Approach", *American Economic Review*, 97, 586-606.

7 Appendix A: State-space Representation of the VECM

The model is given by Equations (2) and (3), coupled with the measurement equation given by Poskitt and Yao (2017) and their assumption that technology growth is uncorrelated. Denoting the observed variables with a superscript *obs*, the state-space form of the latent-variable VECM includes the measurement equations:

$$\begin{bmatrix} \Delta y_t^{obs} \\ h_t^{obs} \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} y_t^s \\ h_t \\ y_{1t}^s \end{bmatrix} + \begin{bmatrix} \epsilon_t^a \\ 0 \end{bmatrix},$$

where y_{1t}^s is the lag of y_t^s.

The transition equations are:

$$\begin{bmatrix} y_t^s \\ h_t \\ y_{1t}^s \end{bmatrix} = \begin{bmatrix} (1+a_{11}) & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1}^s \\ h_{t-1} \\ y_{1t-1}^s \end{bmatrix} + \begin{bmatrix} c_{11} & (c_{12}-1) \\ c_{21} & c_{22} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \epsilon_t^h \\ \epsilon_t^a \end{bmatrix},$$

where a_{ij} denotes the $j'th$ coefficient in the $i'th$ transition equation to be estimated ($i, j = 1$ and 2), and the standard deviations of ϵ_t^h and ϵ_t^a. c_{ij} are the contemporaneous impulse responses, which are fixed to those from the RM, as in Section 2.

8 Appendix B: Unobserved Components Models Allowing for Persistence in the Cycle

Consider the unobserved components model used in Section 4, but now the cycle is assumed to follow an AR(1) with coefficient ρ. Then doing the same analysis as above

$$\begin{aligned}
M &= \begin{bmatrix} 0 & 0 \\ 0 & \rho \end{bmatrix} \Psi = \begin{bmatrix} 0 & \rho - 1 \end{bmatrix}, \\
K_t &= \begin{bmatrix} c_t \\ a_t \end{bmatrix}, \Phi_t = \begin{bmatrix} 0 & c_t(1-\rho) \\ 0 & \rho - (\rho - 1)a_t \end{bmatrix} \\
c_t &= [\Psi P_{t-1|t-1}\Psi' + \Lambda\Lambda']^{-1} \\
a_t &= [1 + (\rho - 1)\rho P_{22,t|t-1}]c_t,
\end{aligned}$$

which gives us

$$\begin{aligned}
E_t\psi_{1t} &= c_t(1-\rho)E_{t-1}\psi_{2t-1} + c_t z_t \\
E_t\psi_{2t} &= (\rho - (\rho - 1)a_t)E_{t-1}\psi_{2t-1} + a_t z_t.
\end{aligned}$$

Now defining $\xi_t = E_t\psi_{2t} - \rho E_{t-1}\psi_{2t-1}$ we can write

$$\begin{aligned}
\xi_t &= a_t[(1-\rho)E_{t-1}\psi_{2t-1} + z_t] \\
&= c_t^{-1}a_t E_t\psi_{1t} \\
&= [1 + (\rho - 1)\rho P_{22,t|t-1}]E_t e_t^P.
\end{aligned}$$

Hence there is a correlation between ξ_t and $E_t e_t^P$, leading to a correlation between the two estimated shocks, even though the true shocks are not correlated.