Tasmanian School of Business and Economics
University of Tasmania

Discussion Paper Series N 2014-14

VAR(MA), What is it Good For? More Bad News for Reduced-form Estimation and Inference

Wenying YAO
University of Tasmania,

Timothy KAM
Australian National University

Farshid VAHID
Monash University

ISBN 978-1-86295-946-0

VAR(MA), What is it Good For?
More Bad News for Reduced-form Estimation and Inference

Wenying Yao[a,*], Timothy Kam[b], Farshid Vahid[c]

[a] *Tasmanian School of Business and Economics, Private Bag 84, University of Tasmania, Hobart TAS 7001, Australia.*
[b] *Research School of Economics, Australian National University, Canberra ACT 0200, Australia.*
[c] *Department of Econometrics and Business Statistics, Monash University, Clayton VIC 3800, Australia.*

Abstract

It is common practice to use reduced-form vector autoregression (VAR) models, or more generally vector autoregressive moving average (VARMA) models, to characterize the dynamics in observed data and to identify innovations to the macroeconomy in some economically meaningful way. We demonstrate that neither approach—VAR or VARMA—are suitable reduced form guides to "reality", if reality were induced by some underlying structural DSGE model. We conduct such a thought experiment across a wide class of DSGE structures that imply particular VARMA data generating processes (DGPs). We find that with the typical small samples for macroeconomic data, the MA component of the fitted VARMA models is close to being non-identified. This in turn leads to an order reduction when identifying the lag structures of the VARMA models. As a result, VARMA models barely show any advantage over VARs using realistic sample sizes. However, the VAR remains a truly misspecified approximation. The VAR's performance deteriorates, in contrast to the VARMA's, as we enlarge the sample size generated from the true DGPs.

JEL: C15, C52, C32.

Keywords: VARMA, VAR, DSGE, impulse response analysis

*Corresponding author: Wenying Yao, wenying.yao@utas.edu.au, Tasmanian School of Business and Economics, Private Bag 84, University of Tasmania. Tel: +61 3 62267141, fax: +61 3 62267587.

Email addresses: `wenying.yao@utas.edu.au` (Wenying Yao), `tcy.kam@gmail.com` (Timothy Kam), `farshid.vahid@monash.edu` (Farshid Vahid)

1. Introduction

It remains common practice to use reduced-form vector autoregression (VAR) models, or more generally vector autoregressive moving average (VARMA) models, to characterize the dynamics in observed data and to identify policy, supply-side or demand-side innovations to the macroeconomy. Underlying such an approach is a two-pronged presumption: (i) VARs or VARMAs provide a useful way to summarize macroeconomic reality with minimal theoretical or structural impositions; and (ii) they form a "good approximation" of some unknown structure—often interpretable as some dynamic stochastic general equilibrium (DSGE)—underlying the data. Our task in this paper is to demonstrate that neither approach—VAR or VARMA—are suitable reduced form guides to "reality", if reality were induced by some underlying structural DSGE model. As a further contribution to the existing literature, we conduct such a thought experiment across more classes of data generating processes (DGPs), each with increasing layers of dynamic sophistication.

Macroeconomists have recognized that approximate solutions to theoretical DSGE models have a VARMA representation with a non-trivial moving average component (King et al., 1988; Cooley and Dwyer, 1998; Fernández-Villaverde et al., 2007). However, in practice finite lag VAR models are used as reduced-form approximations to locally linear solutions of the DSGE model (see e.g., Christiano et al., 2006; Bagliano and Favero, 1998; Erceg et al., 2005; Pagan and Pesaran, 2008). There is an important issue with using truncated- or finite-order VARs. Chari et al. (2007) and Ravenna (2007) showed that a VAR is incapable of capturing the impulse-response dynamics of the true VARMA representation of the DSGE model solution, because the VAR is only a truncated approximation of the true VARMA DGP. On the other hand, Kascha and Mertens (2009) find that VARMA models do not perform better than VARs when the econometrician is endowed with small sample sizes. They attributed it to the fact that the DGPs used in the previous literature (in particular Christiano et al., 2006; Chari et al., 2007) are "*nearly non-stationary, nearly non-invertible and the correct VARMA representation is close to being not identified.*"

Although VARMA models are the correct specification of the underlying DGPs, Kascha and Mertens (2009) estimate a final moving average representation in their simulation, which is not the most parsimonious VARMA structure. In addition, this strand of literature almost exclusively focuses on simulation results based on simple real business cycle (RBC) models using small sample sizes of no more than 50 years of quarterly data. One might question whether the lack of identification in the VARMA model, in small samples, is an artefact of the RBC class of models. As a result, the ubiquity of the conclusions drawn by the aforementioned research may still be questionable.

Our contribution to this ongoing debate is twofold. First, we further investigate whether VARMA models can outperform VARs in replicating the theoretical impulse responses, when the true DGPs are taken from a wider set of DSGE models. We provide even more resounding—albeit not complete—confirmation of the negative verdict on VARMA models in small samples, previously arrived at by Kascha and Mertens (2009) using the RBC model as the sole experimental DGP. We explore why that is the

We thank Heather Anderson, Mardi Dungey, Christian Kascha, Eric Leeper, Don Poskitt, Shaun Vahey, and participants at 2011 Econometric Society Australasian Meeting, 2011 Australian Conference of Economists and Monash Macro Workshop for their valuable comments. Further, we acknowledge financial support from Australian Research Council DP0984399.

case. We demonstrate that the difficulty in identifying the correct VARMA structure with small samples is caused by the fact that the roots of the AR and MA lag matrix polynomials are always very close to each other. We also show that this is almost always true by varying the true DGPs over a wide space of economically reasonable structural parameters. This implies that the "near cancellation of roots" problem is not merely due to just a particular parametric instance of a DGP. This near cancellation in the AR and MA dynamics leads to an order reduction when identifying the canonical VARMA structure. The same phenomenon has been noted by Cogley and Nason (1993) in the case of scalar ARMA processes.

Second, we also illustrate that the VAR's performance, in terms of approximating the true DGPs' impulse dynamics, deteriorates as the sample size is enlarged. In constrast, the VARMA does better in larger samples. This is not surprising as the VAR, by construction is a truncated-lag mis-specification of the true DGPs' VARMA reduced forms.

Our experimental design follows and extends the work of Kascha and Mertens (2009), Erceg et al. (2005), and Ravenna (2007). We first take the stylized RBC model by Hansen (1985) and derive its VARMA representation as the true DGP in the simulation. This model has been commonly used in related literature (for example Ravenna, 2007); hence, it is easier to compare our results with those of other studies. We attempt to identify the dynamics by fitting VAR and VARMA models to the simulated data. These models are chosen by various diagnostics, for example, information criteria for selecting the lag length for VARs, and the scalar component model (SCM) methodology developed by Tiao and Tsay (1989) and Athanasopoulos and Vahid (2008) for identifying parsimonious VARMA forms. In addition to the small sample ($T = 200$) properties of the estimated VAR and VARMA models, their large sample ($T = 20,000$) properties are also explored here. Simulation results suggest that using $T = 20,000$, a VARMA model is able to generate reliable estimates of the impulse responses whereas a VAR model cannot. However, with samples of $T = 200$ observations, a VARMA model does not exhibit any advantage over finite lag VAR, as the identification of the correct VARMA structure from the pseudo data is problematic.

Although it is not possible to examine the entire universe of known (and unknown) model DGPs, we approach the question from two different angles. Beginning with the RBC model as DGP, following Kascha and Mertens (2009), we employ different parameterization schemes by drawing values of deep structural parameters from reasonable prior distributions, and examine whether the near cancellation in the AR and MA dynamics still persists. Then, we move away from the standard RBC paradigm and consider alternative DSGE models. These alternative DGPs have various different economic frictions and richer internal propagation mechanisms. We conduct the reduced-form VARMA identification procedure by repeating the simulation experiments on these models. We show that the problem of near cancellation is caused by the inherent structure of the DGP, and is not confined to a specific set of parametrization.

The remainder of this paper is organised as follows. Section 2 considers the comparison between VARs and VARMA models within the RBC-as-DGP framework, and discusses the near cancellation in the AR and MA dynamics in the resulting VARMA DGP. Section 3 experiments with several alternative DGPs that come from more complex DSGE models to examine the weak identification of the correct VARMA structure. Section 4 concludes.

2. A Real Business Cycle Environment

Consider the indivisible labor RBC model of Hansen (1985). There are two exogenous structural shocks in the model: a non-stationary technology shock Z_t, and a stationary labor supply shock D_t. The per-period preference of the representative household is given by the quasi-linear utility function: $\ln C_t + \phi D_t(1 - N_t)$, where $\phi > 0$ is a constant that represents the relative importance of consumption and leisure in the utility. The social planner chooses a state-contingent sequence of consumption C_t, capital stock K_t, and labor N_t to maximize the expected value of the discounted lifetime utility

$$\mathbb{E}_0 \sum_{t=1}^{\infty} \beta^t [\ln C_t + \phi D_t(1 - N_t)], \tag{1}$$

subject to the capital accumulation law and a Cobb-Douglas production technological constraint:

$$K_t = X_t - C_t + (1 - \delta)K_{t-1}, \tag{2}$$

$$X_t = K_{t-1}^{\alpha}(Z_t N_t)^{1-\alpha}. \tag{3}$$

In equations (1)–(3), $\beta \in [0,\ 1]$ is the inter-temporal discounting rate, $\delta > 0$ is the capital depreciation rate, and $\alpha \in [0,\ 1]$ is the capital share of income. The labor-augmenting technology level Z_t and the labor supply shifter D_t follow the exogenous stochastic processes:

$$\ln Z_t = \ln Z_{t-1} + \mu_z + \varepsilon_{z_t}; \tag{4}$$

$$\ln D_t = (1 - \rho_d) \ln \bar{D} + \rho_d \ln D_{t-1} + \varepsilon_{d_t}, \tag{5}$$

where μ_z is the drift term for the random walk process $\{\ln Z_t\}$, $\ln \bar{D}$ denotes the long run mean of $\{\ln D_t\}$, $\varepsilon_{i_t} \sim \mathcal{N}(0, \sigma_i^2),\ i = z, d$, and $0 < \rho_d < 1$.

A technology shock has a permanent effect on the level of Z_t, and hence on C_t, K_t, and X_t. Therefore, we define the model in terms of the stationary variables $\{N_t,\ R_t,\ D_t,\ \hat{C}_t = C_t/Z_t,\ \hat{X}_t = X_t/Z_t,\ \hat{K}_t = K_t/Z_t,\ \hat{Z}_t = Z_t/Z_{t-1}\}_{t=1}^{\infty}$, and log-linearize around the steady state. For any variable S_t, we define its log-deviation from the steady state value $\bar{S}$, by the lower case letter $s_t = \ln(S_t/\bar{S})$. Following Blanchard and Quah (1989), the percentage deviations of hours worked n_t and output growth $\Delta \ln X_t = \hat{x}_t - \hat{x}_{t-1} + \hat{z}_t$ are taken as the observable variables, that is, $y_t := (n_t,\ \Delta \ln X_t)'$.

2.1. VARMA Representation of the Observables

This model is parameterized following the RBC literature, in particular, Erceg et al. (2005) and Ravenna (2007). First, for the technology shock $\mu_z = 0.0037$ and $\sigma_z = 0.0148$, and hence $\overline{\hat{Z}} = e^{0.0037}$. The steady state level of labor supply shock is normalized to 1, because it does not affect the model's log-linear dynamics. Its standard deviation is $\sigma_d = 0.009$, and the first order autocorrelation coefficient is set to $\rho_d = 0.80$, which indicates a relatively strong persistence of the labor supply shock. The total capital share α is 0.35. The quarterly depreciation rate for installed capital δ is assumed to be 2%, and let $\beta = 1.03^{-0.25}$. Normalizing the total labor endowment to 1, the parameter ϕ in the utility function in equation (1) is chosen such that the steady state level employment is implied to be $\bar{N} = 1/3$, which is full-time work. Given these values of the structural parameters, the solutions for the endogenous variables

$\hat{k}_t$, n_t, and $\hat{x}_t$ are

$$\hat{k}_t = 0.95\,\hat{k}_{t-1} + (-0.95,\ -0.13)\,w_t, \tag{6a}$$

$$\begin{pmatrix} n_t \\ \hat{x}_t \end{pmatrix} = \begin{pmatrix} -0.48 \\ 0.04 \end{pmatrix} \hat{k}_{t-1} + \begin{pmatrix} 0.48 & -2.40 \\ -0.04 & -1.56 \end{pmatrix} w_t, \tag{6b}$$

where $w_t = (\hat{z}_t,\ d_t)'$ is the vector of exogenous state variables, and follows the law of motion

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 - \rho_d L \end{pmatrix} w_t = \varepsilon_t, \quad \text{and } \varepsilon_t = \begin{pmatrix} \varepsilon_{z_t} \\ \varepsilon_{d_t} \end{pmatrix}. \tag{7}$$

We derive the VARMA(1,1) representation for the observable variables $y_t = (n_t,\ \Delta \ln X_t)'$ as the following:

$$y_t = \begin{pmatrix} 0.94 & 1.05 \\ 0 & 0.80 \end{pmatrix} y_{t-1} + u_t + \begin{pmatrix} -0.25 & -0.92 \\ -0.19 & -0.71 \end{pmatrix} u_{t-1}, \tag{8}$$

where u_t is the reduced form disturbance. The VARMA process (8) is strictly stationary and invertible. The eigenvalues of the AR and MA coefficient matrices are given in Table 1.

2.2. Monte Carlo Simulation

We take the reduced form VARMA representation (8) derived from the RBC model outlined above as the data generating mechanism in the Monte Carlo simulation. Assuming that we can only observe simulated pseudo data, we fit the VAR and VARMA structures to each simulated sample path as competing reduced form models. We evaluate whether the identification procedure for VARMA models could successfully detect the correct VARMA structure. Furthermore, we transform both VAR and VARMA models into structural forms using the long-run identification restriction given by Blanchard and Quah (1989), and examine whether they are able to reproduce the impulse dynamics of the true VARMA DGP (8). We consider two cases in this experiment: (i) the pseudo data has a limited number of observations that most practitioners face; and (ii) the number of observations approaches infinity, in which case the asymptotic theory takes effect. This experimental design is repeated in all remaining DSGE models that are used as alternative DGPs.

We simulate 1,000 sample paths from the log-linearized solution for both large ($T = 20,000$) and small samples ($T = 200$). The setting of $T = 200$ corresponds to 50 years of quarterly data, which is similar to the sample sizes examined in previous studies (see, for example, Kascha and Mertens, 2009), and represents a typical sample size for macroeconomic data.

2.2.1. Identifying Structural Shocks

This section presents the identification scheme of Blanchard and Quah (1989) that is used to obtain the responses of observable variables to disturbances in the structural shocks. In the RBC model setting, the only long-run effect exists in the case of the technology shock ε_{z_t} on total production $\ln X_t$. The structural shocks ε_{z_t} and ε_{d_t} can be identified from the additional long-run restriction that the long-run effect of the labor supply shock on output is zero. We denote the VMA representation of $y_t = (n_t,\ \Delta \ln X_t)'$ as $y_t = \Upsilon(L)u_t$, where u_t is the reduced form disturbance with mean zero and a non-singular variance-covariance matrix Σ_u. The matrix polynomial $\Upsilon(L)$ is established by fitting a time series model (VAR or

Table 1: Properties of the true DGPs and simulation results

		RBC	HI[a]	NEWS	MS
Properties of the DGPs					
eig(AR)[b]		0.9459, 0.8	0.99, 0.79, 0.0002, 0[c]	0.98, 0.90, 0.8, 0[c]	0.9386, 0.6, 0.5857, 0
eig(MA)		0.9563, 0	0.91, 0.20[c]	0.92, 0.92[c,d]	0.9425, 0.0087
True VARMA orders		(1,1)	(2,1)	(2,1)	(2,1)
Correct SCM structure		(1,0)∼(1,1)	(1,1)∼(2,1)	(1,1)∼(2,1)	(1,1)∼(2,1)
Simulation results					
$T = 20,000$	% of corr.id. SCM[e]	93.1%	85.9%	97.0%	93.5%
	Median VAR lag[f]	18	49	26	21
$T = 200$	% of corr.id. SCM	4.5%	1.5%	0.2%	1.5%
	Median VAR lag	1	4	1	2
	Most.id. SCM[g]	(0,0)∼(1,0)	(1,0)∼(1,1)	(0,0)∼(1,0)	(1,0)∼(1,1)
	% of most.id. SCM[h]	84.1%	87.8%	79.2%	87.2%

[a] HI stands for the RBC model with habit formation and investment adjustment cost presented in Section 3.1, NEWS stands for the model with news shocks from Section 3.2, and MS denotes for monetary search model discussed in Section 3.3.
[b] eig(AR) and eig(MA) denote roots of the AR and MA characteristic polynomials of the true VARMA processes.
[c] AR and MA roots of the characteristic polynomials for HI and NEWS models are obtained from estimated VARMA models instead of the true DGP.
[d] The MA roots shown here are calculated from the corresponding fundamental representation of the VARMA process.
[e] Represents the percentage of correct identification of the SCM structure when the sample size is $T = 20,000$.
[f] Represents the median of the lag lengths chosen by AIC for VAR using $T = 20,000$. The maximum lag length allowed is set to 50 for $T = 20,000$, and 10 for $T = 200$.
[g] Represents the mostly identified SCM structure when the sample size is $T = 200$.
[h] Represents the percentage of identifying the SCM structure in the row above when the sample size is $T = 200$.

VARMA) to the observed data. Denote the transformation from u_t to structural shocks ε_t as $u_t = A_0\,\varepsilon_t$, then $y_t = \Upsilon(L)A_0\,\varepsilon_t$. The response of $\Delta \ln X_t$ to a labor supply shock ε_{d_t} is associated with the term $[\Upsilon(L)A_0]_{22}$, keeping in mind that the ordering of variables is $y_t = (n_t,\ \Delta \ln X_t)'$ and $\varepsilon_t = (\varepsilon_{z_t},\ \varepsilon_{d_t})'$. The long-run impact of ε_{d_t} on $\ln X_t$ is the sum of all of the coefficients on L, that is, $[\Upsilon(1)A_0]_{22}$. Hence, the long-run restriction implies that $[\Upsilon(1)A_0]_{22} = 0$. The other three restrictions needed to identify the transformation matrix come from the relationship between the variance-covariance matrices of u_t and ε_t that

$$\Sigma_u = A_0\,\Sigma_\varepsilon\,A_0', \quad \text{where } \Sigma_\varepsilon = \mathrm{diag}\{\sigma_z^2,\ \sigma_d^2\}. \tag{9}$$

Given these conditions, A_0 can be identified up to a sign transformation. The last step is to use *a priori* economic intuition to determine the sign of the impacts of structural shocks on the variables of interest; for example, a positive technology shock has a positive effect on the total output, while a positive labor supply shock increases the relative importance of leisure in the utility function, and hence has a negative impact on hours worked. In this paper, we use the impulse responses from the theoretical model as the *a priori* knowledge. The sign of the impulse responses is identified by matching the direction of the long-run impact of the technology shock on output.

The same identification procedure is applied for both the VAR and VARMA models, for both large and small samples. The true impulse responses from the theoretical model are taken as the benchmark that an ideal model is supposed to replicate. The focus is on the impact of the technology shock on $\ln X_t$ and n_t, particularly the response of $\ln X_t$. The effects of the labor supply shock will eventually fade out, and the only permanent effect is that of the technology shock on total output.

2.2.2. Large Sample Simulation

There are two main reasons for studying large sample size simulations. First, the lag order required for a VAR to approximate an infinite VAR process well is likely to be very high, and thus not feasible with small samples. One purpose of our large sample simulation is to see how well the impulse responses based on a sample of $T = 20,000$ observations approximate the true impulse responses. Second, the estimation bias of a VAR decreases as the sample size increases. We do not want to confound the truncation bias, which is caused by using a finite VAR to approximate a VARMA DGP, with the estimation bias arising from the use of small samples.[1]

For each simulated sample path, the structure of the VARMA models is specified using the scalar components model (SCM) methodology developed by Athanasopoulos and Vahid (2008). This method specifies each row of the VARMA model as an SCM with certain orders (p,q), where p denotes the AR lag length and q denotes the MA lag length in that row. The highest SCM orders of the entire system are always the same as the overall VARMA orders. For example, most sample paths simulated from process (8) are identified as a VARMA(1,1) model with SCM(1,1) $\sim$ SCM(1,0). The canonical form for this SCM VARMA structure is

$$\begin{pmatrix} 1 & 0 \\ a_0 & 1 \end{pmatrix} y_t = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} y_{t-1} + u_t + \begin{pmatrix} \theta_{11} & \theta_{12} \\ 0 & 0 \end{pmatrix} u_{t-1}, \tag{10}$$

[1] The terminology "truncation bias" was first used by Ravenna (2007) to represent the bias in the finite lag VAR coefficients.

where the SCM(1,0) (second row) does not have an MA lag. There are certain zero and normalization restrictions imposed on the left-hand side transformation matrix, to obtain a unique identification of the unknown parameters.[2]

Out of the total number of simulated sample paths, the percentage instances of identifying the correct VARMA(1,1) model with SCM(1,0) $\sim$ SCM(1,1) is shown in Table 1 as 95.6%. This is the simplest structure of the underlying true DGP for y_t. In the other 4.4% of the time, the SCM methodology always finds a structure that nests the SCM(1,0) $\sim$ SCM(1,1) with higher orders. The identified canonical SCM VARMA models are estimated using full information maximum likelihood (FIML).

The AR lag of the estimated VARs is selected by the Akaike information criterion (AIC), which chooses a median of 24 lags. We use AIC because it has the smallest penalty for the number of parameters among the popular model selection criteria. Therefore, it should select the model with the longest lag, which may actually be the desirable choice in this case. Most of the chosen lags are higher than 15.

The impulse responses of $\ln X_t$ and n_t to the technology shock ε_{z_t} are plotted in Figure 1 up to 100 periods after a shock occurs. It shows the mean as well as the 2.5 and 97.5 percentiles of the impulse responses estimated from fitted VARs and VARMA models. In each individual period, the average of the point estimates generated from 1,000 simulated samples is taken as the mean impulse response. The distance between the 2.5 and 97.5 percentiles is referred to as the 95 percentile interval throughout the paper. Panels (a) and (c) suggest that the average responses generated from the estimated VARMA models almost overlap with the theoretical ones. The 95 percentile intervals of both responses have reasonable scale. On the other hand, Panels (b) and (d) show that the impulse responses generated from the estimated VARs are systematically biased. Comparing panels (a) and (b), the average response of $\ln X_t$ to ε_{z_t} generated from VARs has a completely different shape from the true response. Moreover, panel (d) shows that the 95 percentile interval of the response of n_t to ε_{z_t} excludes the true response for at least 20 periods in the middle. Even with long lags, VARs are still incapable of mimicking the true dynamics from the theoretical RBC model.

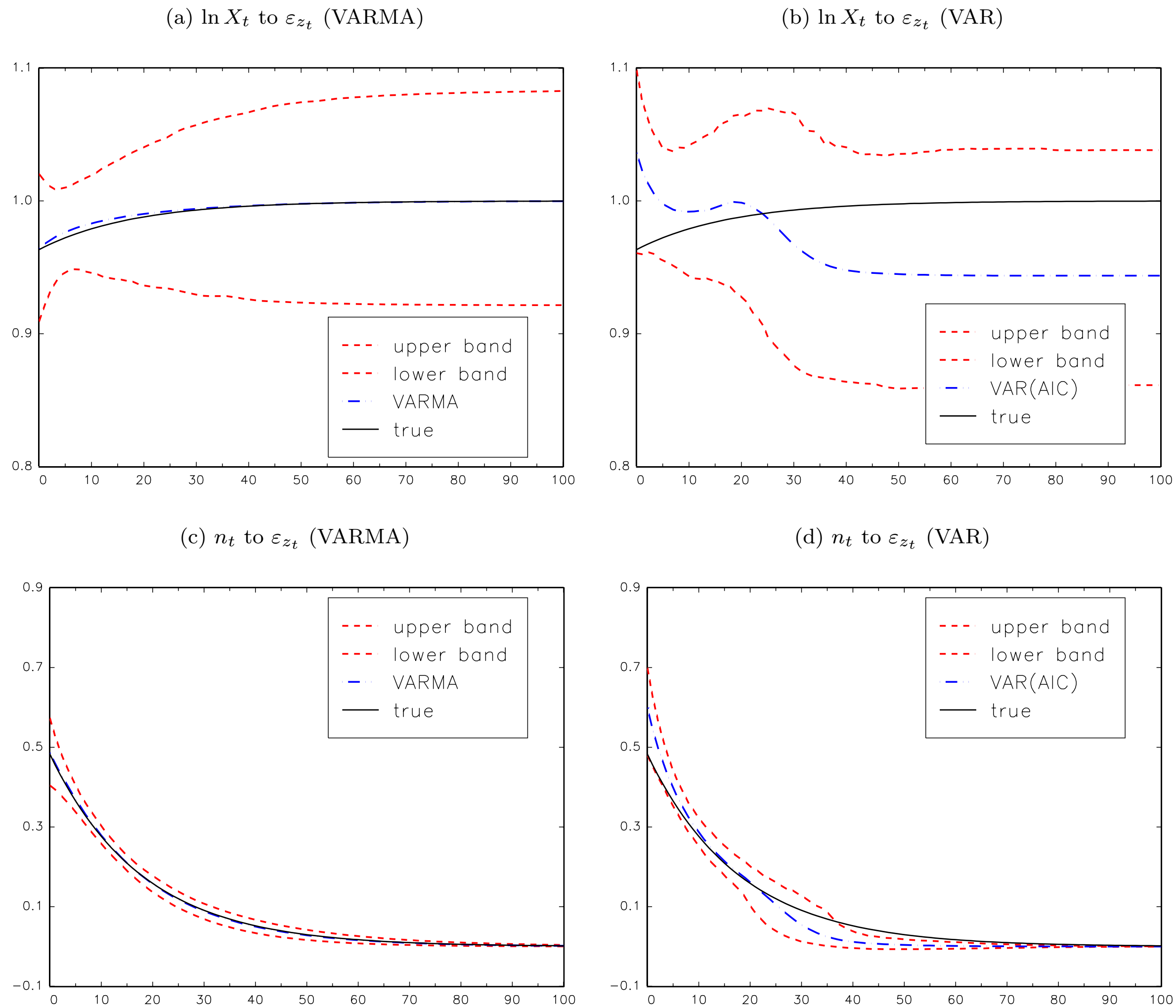
One might suspect that the inability of VARs to approximate the theoretical impulse responses is a result of their use of a model selection criterion for choosing the lag length. In particular, one could suggest that, since we know that RBC models lead to VARMA dynamics, it might be advisable to choose a long lag length, such as $\sqrt{T}$, instead of using model selection criteria. In what follows, we examine the impulse responses produced by one particular sample draw. The AIC chooses 24 lags for this sample path, the Hannan-Quinn information criterion (HQ) chooses 12 lags, and the Bayesian information criterion (BIC) chooses 3 lags. The SCM structure for this sample is correctly identified as SCM(1,0) $\sim$ SCM(1,1). The impulse responses of $\ln X_t$ and n_t to the technology shock generated from large sample estimations of this particular sample path are plotted in Figure 2.

Panels (a) and (c) plot the impulse responses generated from the VAR models with lag lengths chosen by the three information criteria, and the VARMA model is estimated with the identified underlying SCM structure.[3] Evidently, the VARs are incapable of reproducing the true dynamics of the theoretical model, even with a lag length as high as 24. Given the conclusion of Kapetanios et al. (2007) that a VAR

[2]Please refer to Athanasopoulos and Vahid (2008) for more details on the SCM methodology.

[3]The SCM structure does not seem to be crucial in mimicking the impulse dynamics of the theoretical model. The impulse responses generated from estimating a reduced form VARMA(1,1) model without assuming any SCM structure almost overlap with those presented here.

Figure 1: The impulse responses generated from the estimated models (large sample)

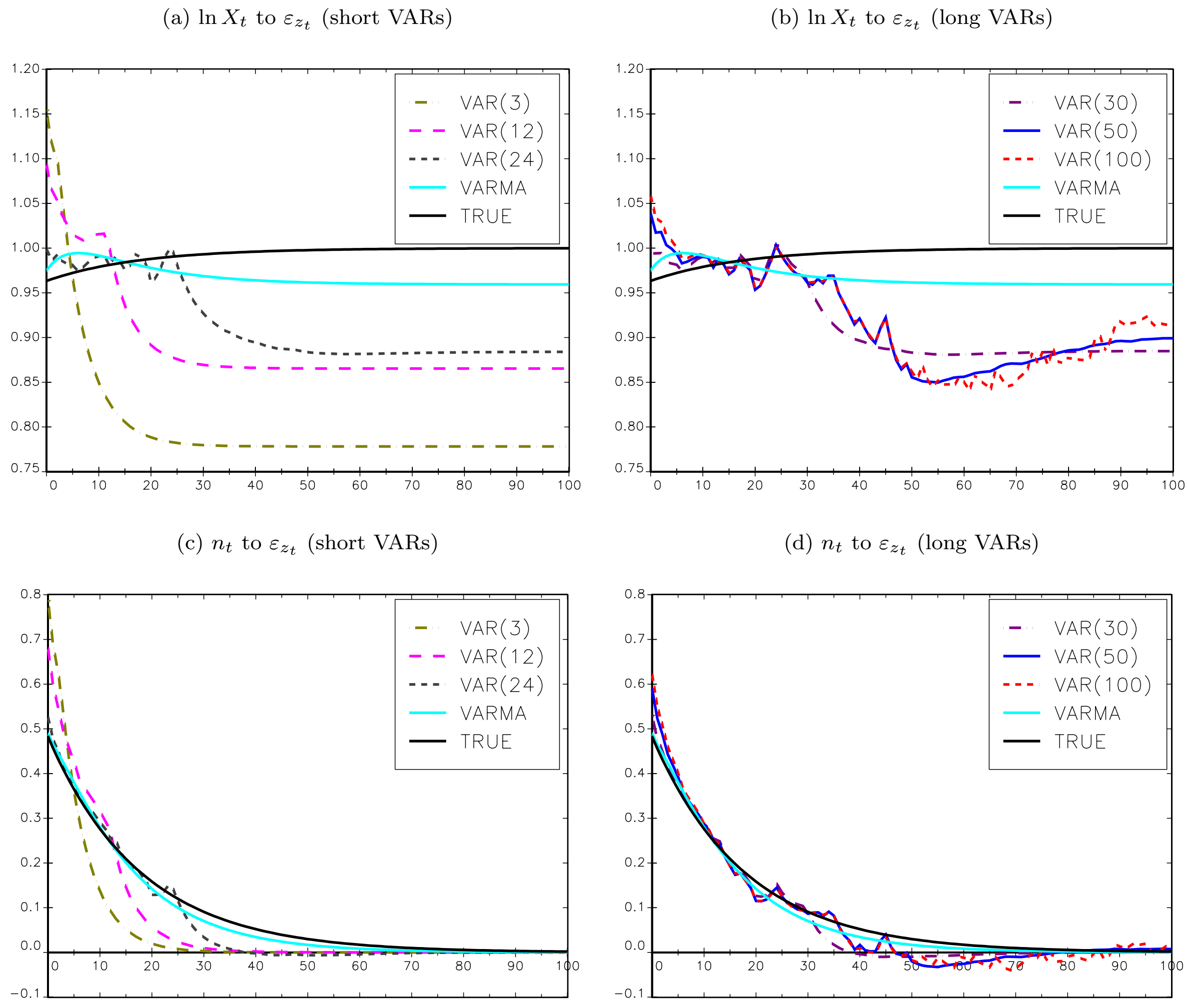


or order 50 is required for a sample of 30,000 observations, it is plausible to expect that longer VARs will be able to capture the effects of technology shock more adequately. However, as panels (b) and (d) of Figure 2 suggest, higher order VARs (e.g. the VAR(100)) contribute nothing other than fluctuations around the estimated impulse responses from the VAR(30). This is consistent with the findings of Poskitt and Yao (2012), that the "approximation error" stems from the difference between the minimum mean squared error VAR approximation, and the true VARMA process converges to its asymptotic limit far more slowly than the asymptotic theory dictates. Consequently, even with considerably large sample sizes and lag lengths, VAR models are likely to exhibit serious errors and behave poorly in practice.

2.2.3. Small Sample Simulation

When working with empirical macroeconomic data, usually there are only a limited number of observations available. Hence, the comparison of VAR and VARMA models based on small samples is crucial for practitioners. Unfortunately, the SCM identification procedure for VARMA models always fails to detect the MA component with sample size $T = 200$, it chooses SCM(0,0) $\sim$ SCM(1,0) 84.1% of the time as shown in Table 1, which is equivalent to a VAR(1). With regard to the estimated VAR models, the

Figure 2: Estimated impulse responses from one particular sample path (large sample)



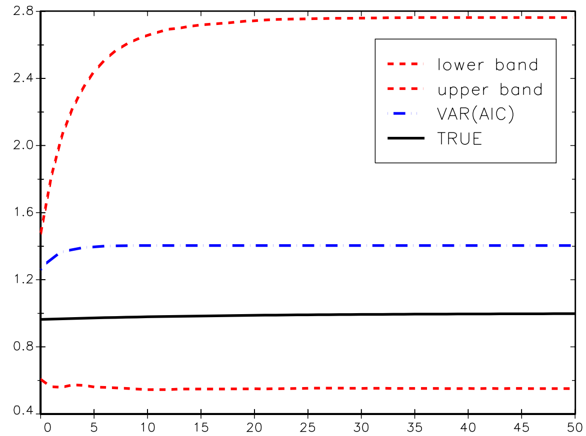
(a) $\ln X_t$ to ε_{z_t} (short VARs)
(b) $\ln X_t$ to ε_{z_t} (long VARs)
(c) n_t to ε_{z_t} (short VARs)
(d) n_t to ε_{z_t} (long VARs)

three information criteria AIC, HQ, and BIC only choose lag one for a majority of the pseudo samples.

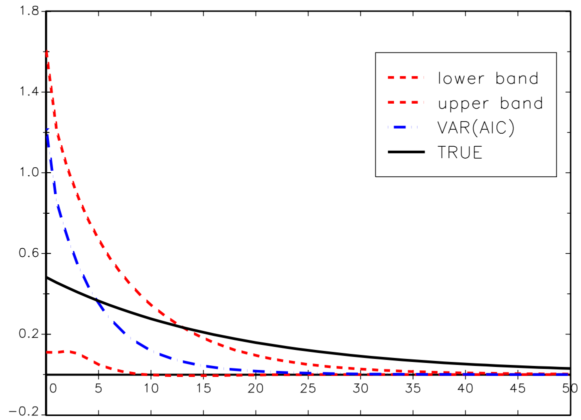
Figure 3 depicts the mean and 95 percentile intervals of the estimated impulse responses with 200 observations. It shows that VAR models based on small samples tend to overestimate the initial impact of the technology shock, and underestimate the initial impact of the labor supply shock. This phenomenon is widely found in all cases, even with a larger sample, or a much higher AR lag length. An important feature of Figure 3 is that the effect of technology shock on labor supply in VARs dies out much faster than the true effect from the theoretical model. This can be attributed to the absence of an MA component, in which case the shocks will appear to be less persistent. Kilian (2011) indicates that finite VAR approximation to the VARMA process may be poor for realistic sample sizes for any feasible choice of lag length, particularly when the VARMA representation has a large MA root. The first column of Table 1 shows that one of the MA roots (0.9563) is very close to the unit circle; thus, finite VARs always fail to produce good approximations of the impulse dynamics in the true VARMA process. We also estimate VARMA models with several different SCM structures to validate this. The resulting impulse responses do not show visible differences from those in Figure 3, which is consistent with the conclusion of Kascha

Figure 3: The impulse responses generated from the estimated VARs (small sample)

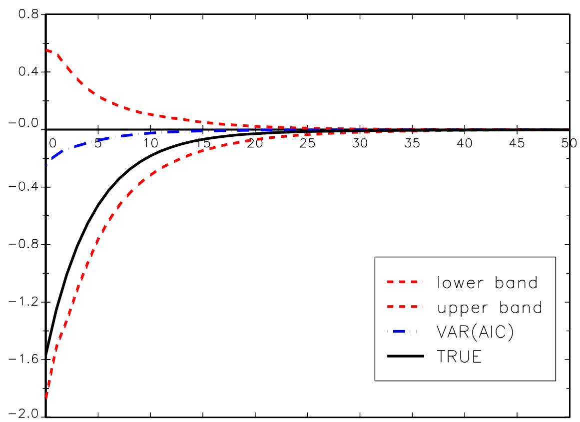
(a) $\ln X_t$ to ε_{z_t}



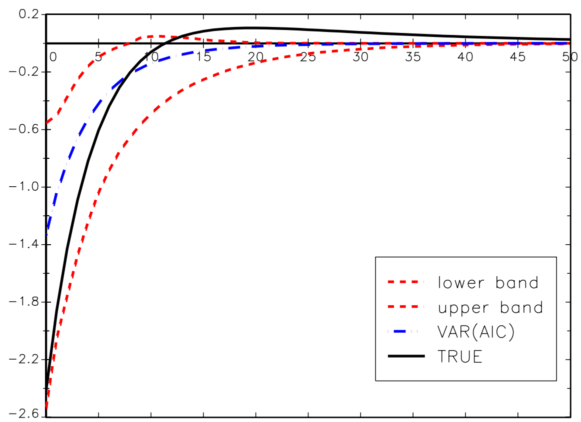
(b) n_t to ε_{z_t}



(c) $\ln X_t$ to ε_{d_t}



(d) n_t to ε_{d_t}



and Mertens (2009).[4]

In their Monte Carlo study, Kascha and Mertens (2009) also find that "*the correct VARMA representation is close to being not identified*". They conclude that the fact that the eigenvalues of the AR and MA parts are very close to each other is the likely explanation of this lack of identification. We face the same problem when identifying the SCM structure using small sample sizes above. The natural question to ask is that, is this always the case for any VARMA processes obtained from the log-linearised solutions of DSGE models? To answer this question, in what follows we first examine different parameterization schemes within the RBC model framework. We demonstrate that the identification difficulty is not caused by a specific choice of the structural parameters, it is rather an inherent feature of the RBC model for almost all reasonable values of structural parameters. We do so by examining the possibility of order reduction in the canonical SCM representations of VARMA DGPs, which are derived from RBC models with different values of the structural parameters. Then we explore alternative DSGE models as

[4]Plots of the impulse responses generated from the estimated VARMA models are omitted here as they are almost exactly the same as those presented in Figure 3.

DGP for the weak identification issue.

2.3. Examining the AR and MA Roots

A two-dimensional VARMA(1,1) process with SCM(1,0) $\sim$ SCM(1,1) obtained from the RBC model will take form as process (10):

$$\begin{pmatrix} 1 & 0 \\ a_0 & 1 \end{pmatrix} y_t = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & 0 \end{pmatrix} y_{t-1} + u_t + \begin{pmatrix} \theta_{11} & \theta_{12} \\ 0 & 0 \end{pmatrix} u_{t-1}. \tag{10}$$

Suppose that the values of the structural parameters are not pre-determined. The unknown coefficients a_0, ϕ_{ij} and θ_{ij}, where i, $j = 1, 2$, need to be estimated. Note that the MA component only appears in the first row of the system of equations. If a fraction of the second row (say, κ) is added to the first row, we obtain

$$\Big[(1+\kappa a_0) - (\phi_{11} + \kappa\phi_{21})\, L\Big] n_t + (\kappa - \phi_{12}\, L)\Delta \ln X_t = u_{1t} + \kappa\, u_{2t} + \theta_{11} u_{1,t-1} + \theta_{12} u_{2,t-1}. \tag{11}$$

The terms on the left-hand side of equation (11) has an MA(1) structure; hence, we can redefine it as $(1 - \gamma\, L)e_t$, where e_t is a univariate error term, and $\gamma < 1$ is the MA(1) coefficient that guarantees the invertibility of this process. If the process u_t is known, γ can be solved analytically given any set of parameter values κ, θ_{11}, and θ_{12}.

From equation (11), the two AR(1) coefficients for n_t and $\Delta \ln X_t$ are

$$AR_n^{(1)} = \frac{\phi_{11} + \kappa\phi_{21}}{1+\kappa a_0}, \quad \text{and} \quad AR_x^{(1)} = \frac{\phi_{12}}{\kappa}. \tag{12}$$

For any fraction κ, there is a corresponding point in the three-dimensional space ($AR_n^{(1)}(\kappa)$, $AR_x^{(1)}(\kappa)$, $\gamma(\kappa)$). In the extreme case where $AR_n^{(1)}(\kappa) = AR_x^{(1)}(\kappa) = \gamma(\kappa)$, (11) degenerates to a static equation with no lagged variables involved, and hence the MA component is not detectable in the system (10). Under looser situations where the set of triplets are close to each other, it still poses a challenge to the identification of the correct VARMA structure, particularly in small samples. Cogley and Nason (1993) come across the same situation with a univariate ARMA process, where the AR and MA lag polynomials have roughly the same factors that almost cancel each other out.

We examine whether the AR(1) and MA(1) coefficients always stay close via a simple simulation exercise. We draw values of the structural parameters in the RBC model randomly from the distributions tabulated in Table 2, and then compute the coefficients in equation (10) given the set of parameter values. κ is chosen such that it makes the two AR(1) coefficients the same, that is, κ is solved by equating (12):

$$AR_n^{(1)}(\kappa) = \frac{\phi_{11} + \kappa\phi_{21}}{1+\kappa a_0} = \frac{\phi_{12}}{\kappa} = AR_x^{(1)}(\kappa). \tag{13}$$

The MA(1) coefficient γ is calculated using this value of κ.[5] This κ may not be the one that minimizes the differences among the set of triplets ($AR_n^{(1)}(\kappa)$, $AR_x^{(1)}(\kappa)$, $\gamma(\kappa)$), but it reduces the problem to two

[5]Generally, equation (13) will yield two different values of κ. We choose the one which generates the smaller distance between the AR(1) and MA(1) coefficients.

dimensions with only one AR coefficient and one MA coefficient.

Table 2: Distributions of the structural parameters used in the simulation

	parameter	distribution	mean	variance
β	discounting factor	$\boldsymbol{U}[0.98, 1]$		
$\bar{N}$	steady state employment	$\boldsymbol{U}[0, 2/3]$		
δ	capital depreciation rate	$\boldsymbol{U}[0, 5\%]$		
α	capital share of income	$\boldsymbol{U}[20\%, 60\%]$		
μ_z	trend in $\ln Z_t$	$\boldsymbol{U}[0, 0.01]$		
ρ_d	AR(1) coefficient in d_t	Beta	0.8	0.022
σ_z	standard error of ε_{z_t}	Inverse Gamma	0.0148	1
σ_d	standard error of ε_{d_t}	Inverse Gamma	0.009	1

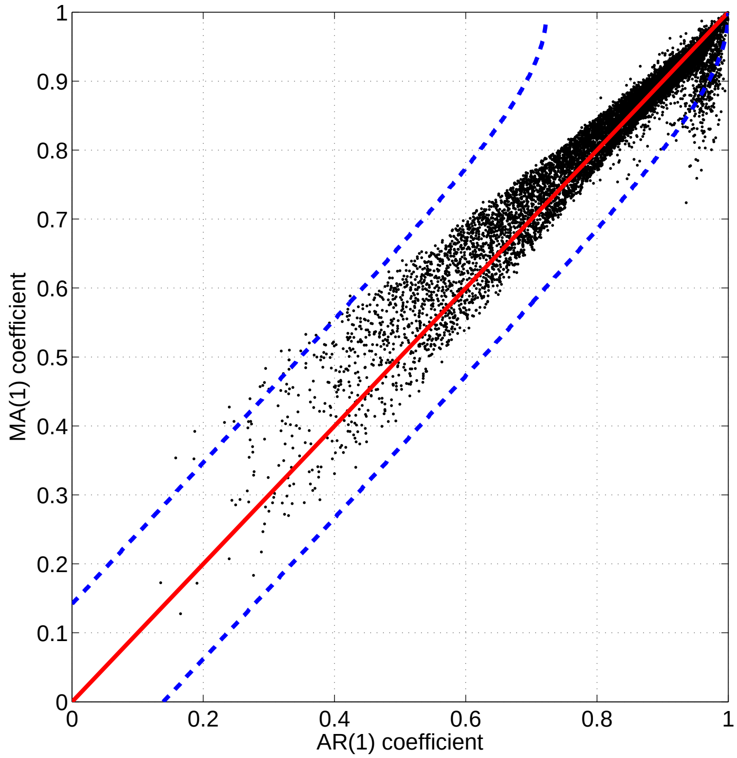
Uniform distributions are used for parameters of which we don't have strong priors, an idea borrowed from the Bayesian DSGE literature. The range of the steady state level of employment is set to be $[0, 2/3]$ so that its mean is the same as is in the previous parameterization. The capital share of income ranges from 20% to 60% according to the estimations given by Valentinyi and Herrendorf (2008). Beta distribution is used for ρ_d because ρ_d is bounded within [0, 1]. The standard errors of the two structural shocks are drawn from Inverse Gamma distributions with unit variance.

For each set of simulated parameters, the AR(1) and MA(1) coefficients in the VARMA representation (10) are computed using the method described above. Figure 4 plots the AR(1) and MA(1) coefficients together with the 45 degree line. The region between the two dashed lines is the parameter space where the true first order autocorrelation coefficient of this ARMA(1,1) process is smaller than $1.96/\sqrt{200}$. That is, given these combinations of AR(1) and MA(1) coefficients, the true first order autocorrelation will not be recognized as statistically significantly different from zero, and hence the process is likely to be identified as a white noise process.[6]

Evidently, the AR(1) and MA(1) coefficients are always close to each other for most simulated values of the structural parameters. From 1,000,000 simulations, the difference between the AR(1) and MA(1) coefficients is greater than 0.1 in absolute value for only 2.5% of the time. The pair of coefficients almost always falls inside the region where the first order serial correlation is statistically insignificant. This conclusively shows that the inherent structure of the RBC model itself always gives rise to a data generating mechanism, in which this type of near cancellation is very likely to occur. The findings of Kascha and Mertens (2009) still hold in our more general setting, that changing the value of structural parameters in the RBC model barely affects the fact that the MA root stays very close to one of the AR roots. This reveals an important challenge for practitioners: even though the RBC model implies a VARMA process, it is difficult to distinguish this VARMA process from a finite order VAR using a small number of observations. However, the VAR also fails to approximate the true VARMA impulse dynamics as we enlarge the sample size.

[6] This is only the tip of the iceberg of the identification problem, because it assumes that first order autocorrelation coefficient can be estimated precisely, which is not the case when AR and MA parameters are close to each other. Andrews and Cheng (2012) provide a comprehensive analysis of this problem.

Figure 4: Simulated distribution of the AR(1)-MA(1) pair of coefficients



3. Near Cancellation in DSGE Models

In the RBC model above, we show that the roots of the AR and MA lag polynomials are always close to each other, which in turn causes an order reduction when identifying the structure of the VARMA models using small samples. More importantly, this problem of near cancellation still remains when we change the values of the structural parameters within a reasonable range. Therefore, we conclude that the similarity in the AR and MA dynamics is an inherent feature of the RBC model itself.

A natural question that arises is whether all VARMA DGPs implied by other types of DSGE models have the near cancellation in the AR and MA dynamics. Although it is not possible to analyze all existing models in the literature, we select a few simple but representative DSGE models, which have richer dynamics induced by various different economic frictions, and study them below. We conduct the SCM identification procedure on these models using both large and small samples, and examining the roots of their AR and MA polynomials. The experimental designs for these DSGE models are the same as in the previous section, and hence are presented in brevity. This exercise will shed some light on the ubiquity of the near cancellation problem.

3.1. Habit Formation and Investment Adjustment Cost

The inclusion of habit formation in the consumer's utility function and investment adjustment cost in the capital accumulation law has become standard in many macroeconomic models. These factors enrich the dynamic interactions among variables and increase the complexity of the model, which may help to resolve the near cancellation in the AR and MA dynamics.

With habit formation in consumption C_t and leisure $L_t = 1 - N_t$, the utility function of a representative household becomes

$$\mathbb{E}_0 \left\{ \sum_{t=1}^{\infty} \beta^t [\ln(C_t - \theta_c C_{t-1}) + D_t (L_t - \theta_l L_{t-1})] \right\}, \tag{14}$$

where $\theta_c > 0$ and $\theta_l > 0$. The agent's utility depends on the current consumption relative to a fraction of the past consumption and leisure. There has been some empirical evidence in support of habit formation in the utility function in the literature (see, for example, Campbell and Cochrane, 1999; Carroll et al., 2000). Furthermore, previous studies (Fuhrer, 2000; Boldrin et al., 2001) have shown that general equilibrium models with utility functions which incorporate habit formation are able to produce a hump-shaped responses of consumption and output to all shocks in the model, and in particular to the monetary policy shock. Such consumption smoothness is more realistic, and is also consistent with empirical data.

Investment adjustment costs are often introduced in RBC models. They help to match the empirical evidence that investment adjusts slowly in response to shocks, and hence have substantive implications for understanding the aggregate dynamics of DSGE models. We augment the previous model with the cost S_t, which is defined by a quadratic loss function

$$S_t = \frac{\varsigma}{2} \left(\frac{I_t}{I_{t-1}} - 1 \right)^2, \qquad \varsigma > 0. \tag{15}$$

Let the total capital stock in the economy accumulate according to

$$K_t = (1 - \delta_t)\, K_{t-1} + (1 - S_t)\, I_t, \tag{16}$$

where δ_t is the time-varying depreciation rate for installed capital, and is a quadratic function of the capital utilization rate U_t,

$$\delta_t = \delta_0 + \delta_1 (U_t - 1) + \frac{\delta_2}{2} (U_t - 1)^2, \qquad \delta_0,\, \delta_1,\, \delta_2 > 0. \tag{17}$$

Here U_t will also affect the production function,

$$X_t = (U_t\, K_{t-1})^{\alpha}\, (Z_t\, N_t)^{1-\alpha}. \tag{18}$$

Equations (14)–(18) outline the additional features in a more complex version of the RBC model.

Values of the additional parameters in this model are set as follows. The habit persistence parameters for consumption and leisure are $\theta_c = 0.7$ and $\theta_l = 0.5$. The investment adjustment cost parameter is $\varsigma = 4$. The steady state level of capital depreciation rate is $\bar{\delta} = \delta_0 = 2.5\%$, and $\delta_2 = 0.11$. Capital stock is fully utilized in the steady state, hence $\bar{U} = 1$.

Simulation results in the second column of Table 1 reveal that with a sample size of $T = 20,000$, the SCM methodology detects a VARMA(2,1) model for $y_t = (n_t,\ \Delta \ln X_t)$ most often, but only identifies VARMA(1,1) in the case of $T = 200$, where the AR and MA orders in the SCM structure are reduced by one. Apparently, the identification difficulty of the correct VARMA structure also exists in this DGP when we use small samples. The fact that AIC always chooses very long lags for finite VAR in the case of $T = 20,000$ suggests that this DGP has strong MA prorogation dynamics—almost all of the selected lag lengths are higher than 40, with median 49 when the maximum permissable lag length is set to 50.

The minimal VARMA representation for true process $y_t = (n_t,\ \Delta \ln X_t)'$ cannot be derived analytically from the log-linearized solution, because there are more endogenous state variables than the observable variables in this model. The exact mechanism of this order reduction in identification using small samples is unknown, as the VARMA coefficients are complicated functions of the structural parameters from the DSGE model. Given the RBC model example, we suspect that one likely reason is the closeness of the AR and MA roots, which leads to similar AR and MA dynamics. To gain some insight into the near cancellation of AR and MA dynamics, we use one simulated sample path with $T = 20,000$ to estimate the identified structure SCM(1,1) $\sim$ SCM(2,1), and examine the roots of the AR and MA characteristic polynomials of the estimated VARMA(2,1) model. FIML estimates of these fitted VARMA models display satisfactory large sample properties, the estimated coefficients and characteristic roots obtained from several different sample paths are very similar. The characteristic roots of the AR and MA lag polynomials tabulated in Table 1 display the same property as in the RBC model example: they are close to each other and near unity. The closeness of the AR and MA roots is likely to be the reason for the order reduction in the identification using small samples. Nevertheless, the VAR also fails to approximate the true VARMA impulse dynamics as we enlarge the sample size.

3.2. RBC Model with News Shocks

Macroeconomic models with anticipated policy shocks have drawn a considerable amount of attention in recent years. This type of model is also appealing to econometricians, because it breaks the conventional information assumption regarding unanticipated shocks in econometric models. These macroeconomic models yield non-fundamental shocks, that is, the information set of the forward-looking economic agents does not match the information set of econometricians. Hence, the space spanned by the structural shocks is larger than the space spanned by current and lagged variables (see Hansen and Sargent, 1991). Mathematically, this will cause the VARMA representation of the log-linearized solution of the economic model to be non-invertible, and the structural shocks cannot be recovered from a VAR(∞) process. In such situations, econometricians can only work with VARMA models, even with an infinite number of observations. Studies of this type of model include Sims (1988); Edelberg et al. (1999), and Leeper et al. (2008).

Ignoring for the moment the problem of non-fundamentalness, we construct a simple RBC model with fiscal foresight based on the work of Yang (2005) in order to examine the identification of the VARMA structure. The main differences between this model and the RBC model in Section 2 are the additional policy variables and the specification of exogenous processes.

The representative household's maximization problem is given by

$$\max \mathbb{E}_0 \left\{ \sum_{t=1}^{\infty} \beta^t [\ln C_t + D_t(1 - N_t)] \right\}, \tag{19}$$

subject to the agent's per-period budget constraint

$$C_t + K_t - (1-\delta)\, K_{t-1} + T_t = (1 - \tau_t^L) W_t N_t + (1 - \tau_t^K) R_t K_{t-1}. \tag{20}$$

Here, T_t is a lump-sum tax, and τ_t^L and τ_t^K are the tax rates on labor and capital income. The representative firm faces the standard profit maximization problem subject to a Cobb-Douglas production technology. The government's per-period budget constraint requires

$$G_t = T_t + \tau_t^L W_t N_t + \tau_t^K R_t K_{t-1}, \tag{21}$$

where the policy variables, government spending G_t, and the two tax rates, evolve according to

$$\ln \tau_t^L = \rho_L \ln \tau_{t-1}^L + \mu_L \ln(X_t/Z_t) + \varepsilon_{L_{t-1}} + r_{K,L}\, \varepsilon_{K_{t-1}}, \tag{22a}$$

$$\ln \tau_t^K = \rho_K \ln \tau_{t-1}^K + \mu_K \ln(X_t/Z_t) + \varepsilon_{K_{t-1}} + r_{K,L}\, \varepsilon_{L_{t-1}}, \tag{22b}$$

$$\ln(G_t/Z_t) = \rho_G \ln(G_{t-1}/Z_{t-1}) + \varepsilon_{G_t}. \tag{22c}$$

The random variables ε_{G_t}, ε_{L_t}, and ε_{K_t}, are the *i.i.d.* exogenous government spending, labor and capital tax shocks, respectively, and $r_{K,L} = 0.26$ allows for a correlation between the two tax processes. Note that, based on the specifications of equations (22a) and (22b), tax shocks occurring at period-t will change the tax rates at period-$(t+1)$. Hence, the agents have one-period foresight. Table 3 presents the values of the structural parameters. Most of them are chosen to be the same as those in the work of Yang (2005).

Table 3: Parameterization of the RBC model with news shocks

	parameter	value
ρ_G	AR(1) coefficient of government spending	0.99
ρ_K	AR(1) coefficient of capital income tax	0.97
ρ_L	AR(1) coefficient of labor income tax	0.99
σ_d	standard deviation of labor supply shock	0.005
σ_G	standard deviation of government spending	0.018
σ_K	standard deviation of capital income tax	0.025
σ_L	standard deviation of labor income tax	0.020
σ_z	standard deviation of technology shock	0.009
μ_K	elasticity of capital tax w.r.t. income	1
μ_L	elasticity of labor tax w.r.t. income	1
α	capital share of total income	0.36
β	inter-temporal discounting rate	0.99
$\bar{\tau}^K$	steady state level of capital tax	0.39
$\bar{\tau}^L$	steady state level of labor tax	0.21

Taking $y_t = (n_t,\ \Delta \ln X_t)'$, column three of Table 1 suggests that identification using 20,000 observations finds that the theoretical DGP is a VARMA(2,1) process with SCM(1,1)$\sim$SCM(2,1). However, y_t is identified to be a VARMA(1,0) process in most cases using 200 observations. We examine the roots of the AR and MA characteristic polynomials from one simulated sample path, where the MA roots are calculated from the corresponding fundamental representation.[7] We also calculated roots from estimated

[7]The VARMA models are usually estimated from the data while imposing the stationarity and invertibility conditions, thus we use the fundamental MA roots instead of the non-fundamental ones. The non-fundamental roots can be readily obtained using Blaschke matrices as introduced by Lippi and Reichlin (1994).

VARMA models using different samples of size $T = 20,000$, they yield very similar estimates of the AR and MA roots. The results presented in Table 1 suggests that the two MA characteristic roots are close to two of AR roots. Hence, this model with anticipated policy shocks still suffers from the problems of order reduction and near cancellation of the AR and MA dynamics. This analysis based on the VARMA DGP resulting from the RBC model with anticipated policy shocks provides additional evidence of the closeness of AR and MA characteristic roots and the identification difficulty using small samples. Further, as in the RBC example earlier, the VAR still fails to approximate the true VARMA impulse dynamics as we enlarge the sample size.

3.3. A Monetary Search-theoretic Model

We turn to a monetary model with searching-matching friction along the line of Aruoba et al. (2008), which builds upon the seminal work of Lagos and Wright (2005). These monetary models have recently been shown to capture US real and monetary (closed or international) business cycle facts rather well. For example, Aruoba (2011), among others, examines consumption, investment, labour productivity, wage, and markups; Gomis-Porqueras et al. (2013) extend the analysis to an international setting and show that it also matches the excess volatility and persistence in real exchange rate. The RBC model is known for matching real data's business cycle facts (i.e. standard deviation, autocorrelation, cross-correlation), and search models provide more micro-foundation above that. This model has distinct dynamics from the RBC models examined in previous sections. These features do not have specific implications on the linearised VARMA solution of the theoretical model, but in general it is expected that richer dynamics could be useful in order to avoid the near cancellation problem encountered with small samples.

At the beginning of each time period t, anonymous agents exist on a continuum $[0, 1]$ and have a common discount factor $\beta \in (0, 1)$. Each $t \in \mathbb{N}$ is composed of two sub-periods, night and day. At night, the agents face a random meeting technology, which determines whether they enter a *decentralized market* (DM) or not. We assume that with probability $\sigma \leq 1/2$ that each agent can access the DM as a buyer of a particular good q^b. With the same probability σ, the agent can access the DM to sell his specific q^s. With probability $1 - 2\sigma$, the agent will leave the DM with no exchange. For the sake of simplicity, we assume that "double-coincidence-of-wants" events (where buyers and sellers in the DM are able to barter) and events where the agent can buy q^b and sell q^s simultaneously, both occur with zero probability. Anonymity and stochastic trading opportunities in the DM imply that an intrinsically worthless money-like object (fiat money) will be the only medium of exchange accepted in these DM trades.[8]

During the day, agents trade in *centralized markets* (CM). The CM resembles a standard neoclassical monetary business cycle model with Walrasian markets. Agents gain utility from consuming the CM general good X, and disutility of work effort N. Hence agents' per-period utility function in the CM is $\varrho \ln(X_t) - \phi N_t$, with the following budget constraint

$$X_t + k_t - (1 - \delta)k_{t-1} = \frac{m_{t-1} - m_t}{P_t} + w_t N_t + r_t k_{t-1} + TR_t, \tag{23}$$

[8] In particular, in the absence of a means of monitoring or communicating between agents, and a lack of ability to punish unilateral deviations from contractual obligations, an equilibrium with credit or private claims as media of exchange cannot exist. The inability to enforce or punish arises naturally as a result of the continuum-of-agents assumption (see e.g., Aliprantis et al., 2007*a*,*b*).

where m_{t-1} and k_{t-1} are stocks of individual nominal money and capital holdings, P_t is the price level of X_t, and TR_t is the lump-sum transfer from the monetary authority. The representative firm solves a standard profit maximization problem facing a Cobb-Douglas production function.

The structural shocks in this model are a money supply shock and a technology shock. We assume that the growth factor of money supply, $\psi_t := M_t/M_{t-1}$, follows a stationary AR(1) process:

$$\ln(\psi_t) = \rho_M \ln(\psi_{t-1}) + \sigma_M \varepsilon_{\psi_t}, \quad 0 < \rho_M < 1, \text{ and } \quad \varepsilon_{\psi_t} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1). \tag{24}$$

Following Ireland and Schuh (2008), we specify the technology stochastic process as an AR(1) in its growth factor, $\tilde{Z}_{t+1} := Z_{t+1}/Z_t$:

$$\ln(\tilde{Z}_t) = \rho_Z \ln(\tilde{Z}_{t-1}) + \sigma_Z \varepsilon_{Z_t}, \quad 0 < \rho_Z \leq 1, \text{ and } \quad \varepsilon_{Z_t} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1). \tag{25}$$

As this model now has two sectors, DM and CM, we define aggregate measure of output and employment for the economy as a whole. The details of the monetary search model are described in Supplementary Appendix A. In terms of the observable variables—percentage deviations of the aggregate employment and the growth of real output—we have $y_t := (n_{tot,t},\ \Delta \ln X_{tot,t})'$.

We parameterize the model according to the monetary model literature (Schlagenhauf and Wrase, 1995; Chari et al., 2002; Heathcote and Perri, 2002; Ireland and Schuh, 2008). The probability of entering DM as a buyer or seller is $\rho = 0.26$. The AR(1) coefficients of the exogenous processes are $\rho_M = 0.5857$ and $\rho_Z = 0.6$, while the standard deviations are $\sigma_M = 0.00397$ and $\sigma_Z = 0.007$. Other calibrated parameters are discussed in Appendix A.5. The minimal VARMA representation of the log-linearized solution is a VARMA(2,1) process with the structure SCM(1,1) $\sim$ SCM(2,1)

$$\begin{pmatrix} 1 & 0 \\ -0.09 & 1 \end{pmatrix} y_t = \begin{pmatrix} 1.56 & -0.02 \\ -0.07 & 0.57 \end{pmatrix} y_{t-1} + \begin{pmatrix} -0.58 & 0 \\ 0 & 0 \end{pmatrix} y_{t-2} + u_t + \begin{pmatrix} -1.26 & 0.40 \\ -1.09 & 0.34 \end{pmatrix} u_{t-1}.$$

Simulation suggests that when $T = 20,000$, the SCM methodology identifies the correct structure 93.5% of the time, but identifies a VARMA(1,1) with SCM(1,0) $\sim$ SCM(1,1) most often when we reduce the sample size to $T = 200$. We encounter exactly the same problem as in the prototype RBC model, that is the correct VARMA structure cannot be identified with only 200 observations. The AR and MA roots shown in Table 1 suggest that once again, some of the roots are very close, and hence the near cancellation of AR and MA dynamics is very likely to occur. However, as in the RBC example earlier, the VAR also fails to approximate the true VARMA impulse dynamics as we enlarge the sample size.

4. Conclusion

Our task in this paper was to demonstrate that neither approach—VAR or VARMA—are suitable reduced form guides to "reality" if reality were induced by some underlying structural DSGE model. As a contribution to the existing literature, we conduct such a thought experiment across more classes of data generating processes (DGPs), each with increasing layers of dynamic sophistication.

Table 1 summarized the properties of the VARMA DGPs and the identification results from both large and small sample simulations for all four DSGE models examined in this paper. These models have very diverse dynamic interactions among variables in the system. However, identification of the correct VARMA (as well as the SCM) structure given small sample sizes is a very challenging task in all cases. It is also evident that for the DGPs' true VARMA representations considered in this paper, their AR and MA components almost always have similar roots.

What we have demonstrated are the following: Large samples are required to identify the VARMA structure of the underlying true DGP correctly. Given a sufficiently large number of observations, we do observe that the VARMA approach outperforms the VAR in approximating the underlying true DGP's impulse dynamics. A finite order VAR induces misleading impulse dynamics of the system, since it is always a truncated approximation of a DGP's VARMA process. The former (VAR) suffers from a model mis-specification problem which worsens as the sample size enlarges. The latter (VARMA) while being closer to the true DGP, suffers from identification problems in small samples. The only saving grace for the VARMA is that with very large samples, it begins to approximate the true DGPs' impulse dynamics well. This poses a conundrum for the reduced form VAR or VARMA practitioner intent on identifying, making inferences about, and quantifying economically meaningful shocks to the economy.

Therefore, if macroeconomic reality can be thought of as being induced by some underlying DSGE structure, then the way forward for the practitioner is to take a stance and proceed with a well specified DSGE model whose solution can potentially be taken directly to the data. There are existing sufficiency conditions that allow one to conclude when a well specified DSGE model is (locally) identifiable, but there are no known results to tell us why a VARMA or VAR fails to identify the correct stochastic process underlying the data.[9] As we have demonstrated, there is no point in taking the intermediate step of working with reduced form VARs or VARMAs to uncover macroeconomic reality.

References

Aliprantis, C. D., Camera, G. and Puzzello, D. (2007*a*), "Anonymous Markets and Monetary Trading", *Journal of Monetary Economics* **54**(7), 1905–1928.

Aliprantis, C. D., Camera, G. and Puzzello, D. (2007*b*), "Contagion Equilibria in a Monetary Model", *Econometrica* **75**(1), 277–282.

Andrews, D. W. K. and Cheng, X. (2012), "Estimation and Inference With Weak, Semi-Strong, and Strong Identification", *Econometrica* **80**(5), 2153–2211.

Aruoba, S. B. (2011), "Money, Search, And Business Cycles", *International Economic Review* **52**(3), 935–959.

Aruoba, S. B., Waller, C. J. and Wright, R. (2008), "Money and Capital", *Journal of Monetary Economics* **58**(2), 98–116.

Athanasopoulos, G. and Vahid, F. (2008), "A Complete VARMA Modelling Methodology Based on Scalar Components", *Journal of Time Series Analysis* **29**(3), 533–554.

[9] Komunjer and Ng (2011) provide sufficient conditions for identifying the structural parameters of a linear DSGE model, but they do not address the identifiability of the DGP's VARMA form, *given* a DSGE as the DGP. Andrews and Cheng (2012) study the properties of standard estimators, tests, and confidence sets for reduced-form model parameters when they are not identified or weakly identified. They also derive methods for inference that remain robust to such identification problems. However, they do not address the open question of why reduced-form VARMAs fail to be identified in small samples induced by known DGPs with VARMA representations.

Bagliano, F. C. and Favero, C. A. (1998), "Measuring Monetary Policy with VAR Models: An Evaluation", *European Economic Review* **42**(6), 1069–1112.

Blanchard, O. J. and Quah, D. (1989), "The Dynamic Effects of Aggregate Demand and Supply Disturbances", *American Economic Review* **79**(4), 655–73.

Boldrin, M., Christiano, L. J. and Fisher, J. D. M. (2001), "Habit Persistence, Asset Returns, and the Business Cycle", *The American Economic Review* **91**(1), pp. 149–166.

Campbell, J. Y. and Cochrane, J. (1999), "Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior", *Journal of Political Economy* **107**(2), 205–251.

Carroll, C. D., Overland, J. and Weil, D. N. (2000), "Saving and Growth with Habit Formation", *The American Economic Review* **90**(3), 341–355.

Chari, V. V., Kehoe, P. J. and McGrattan, E. R. (2002), "Can Sticky Price Models Generate Volatile and Persistent Real Exchange Rates?", *Review of Economic Studies* **69**(3), 533–563.

Chari, V. V., Kehoe, P. J. and McGrattan, E. R. (2007), "Are Structural VARs with Long-run Restrictions Useful in Developing Business Cycle Theory?", Technical report, Federal Reserve Bank of Minneapolis.

Christiano, L. J., Eichenbaum, M. and Vigfusson, R. (2006), "Assessing Structural VARs", NBER Working Papers 12353, National Bureau of Economic Research, Inc.

Cogley, T. and Nason, J. M. (1993), "Impulse Dynamics and Propagation Mechanisms in a Real Business Cycle Model", *Economics Letters* **43**(1), 77–81.

Cooley, T. and Dwyer, M. (1998), "Business Cycle Analysis without Much Theory. A Look at Structural VARs", *Journal of Econometrics* **83**, 57–88.

Duffie, D., Geanakoplos, J., Mas-Colell, A. and McLennan, A. (1994), "Stationary Markov Equilibria", *Econometrica* **62**(4), 745–781.

Edelberg, W., Eichenbaum, M. and Fisher, J. D. (1999), "Understanding the Effects of a Shock to Government Purchases", *Review of Economic Dynamics* **2**(1), 166–206.

Erceg, C. J., Guerrieri, L. and Gust, C. (2005), "Can Long-Run Restrictions Identify Technology Shocks?", *Journal of the European Economic Association* **3**(6), 1237–1278.

Fernández-Villaverde, J., Rubio-Ramírez, J. F., Sargent, T. J. and Watson, M. W. (2007), "ABCs (and Ds) for understanding VARs", *American Economic Review* **97**(3), 1021–1026.

Fuhrer, J. C. (2000), "Habit Formation in Consumption and Its Implications for Monetary Policy Models", *The American Economic Review* **90**(3), 367–390.

Gomis-Porqueras, P., Kam, T. and Lee, J. (2013), "Money, Capital, and Exchange Rate Fluctuations", *International Economic Review* **54**(1), 329–353.

Hansen, G. D. (1985), "Indivisible Labor and the Business Cycle", *Journal of Monetary Economics* **16**(3), 309–327.

Hansen, L. P. and Sargent, T. J. (1991), "Two Difficulties in Interpreting Vector Autoregressions", *Rational Expectations Econometrics, Westview Press, Boulder and London*, 77–119.

Heathcote, J. and Perri, F. (2002), "Financial Autarky and International Business Cycles", *Journal of Monetary Economics* **49**(3), 601–627.

Ireland, P. and Schuh, S. (2008), "Productivity and U.S. Macroeconomic Performance: Interpreting the Past and Predicting the Future with a Two-sector Real Business Cycle Model", *Review of Economic Dynamics* **11**(3), 473–492.

Kapetanios, G., Pagan, A. and Scott, A. (2007), "Making a Match: Combining Theory and Evidence in Policy-oriented Macroeconomic Modeling", *Journal of Econometrics* **136**(2), 565–594.

Kascha, C. and Mertens, K. (2009), "Business Cycle Analysis and VARMA Models", *Journal of Economic Dynamics and Control* **33**(2), 267–282.

Kilian, L. (2011), "Structural Vector Autoregressions", CEPR Discussion Papers 8515, C.E.P.R. Discussion Papers.

King, R. G., Plosser, C. I. and Rebelo, S. T. (1988), "Production, Growth and Business Cycles: II. New Directions", *Journal of Monetary Economics* **21**(2-3), 309–341.

Komunjer, I. and Ng, S. (2011), "Dynamic Identification of Dynamic Stochastic General Equilibrium Models", *Econometrica* **79**(6), 1995–2032.

Lagos, R. and Wright, R. (2003), "Dynamics, Cycles, and Sunspot Equilibria in 'Genuinely Dynamic, Fundamentally Disaggregative' Models of Money", *Journal of Economic Theory* **109**(2), 156–171.

Lagos, R. and Wright, R. (2005), "A Unified Framework for Monetary Theory and Policy Analysis", *Journal of Political Economy* **113**(3), 463–484.

Leeper, E. M., Walker, T. B. and Yang, S.-C. S. (2008), "Fiscal Foresight: Analytics and Econometrics", Working Paper 14028, National Bureau of Economic Research.

Lippi, M. and Reichlin, L. (1994), "VAR Analysis, Nonfundamental Representations, Blaschke Matrices", *Journal of Econometrics* **63**(1), 307–325.

Pagan, A. and Pesaran, M. H. (2008), "Econometric Analysis of Structural Systems with Permanent and Transitory Shocks", *Journal of Economic Dynamics and Control* **32**(10), 3376–3395.

Poskitt, D. and Yao, W. (2012), "Vector Autoregressions and Macroeconomic Modelling: An Error taxonomy", Monash Econometrics and Business Statistics Working Papers 11/12, Monash University, Department of Econometrics and Business Statistics.

Ravenna, F. (2007), "Vector Autoregressions and Reduced Form Representations of DSGE Models", *Journal of Monetary Economics* **54**(7), 2048–2064.

Schlagenhauf, D. E. and Wrase, J. (1995), "Exchange Rate Dynamics and International Effects of Monetary Shocks in Monetary Equilibrium Models", *Journal of International Money and Finance* **14**, 155–177.

Sims, C. A. (1988), "Identifying Policy Effects", *Empirical Macroeconomics for Interdependent Economics, the Brookings Institution, Washington D.C.* pp. 308–321.

Tiao, G. C. and Tsay, R. S. (1989), "Model Specification in Multivariate Time Series (with discussion)", *Journal of the Royal Statistical Society B* **51**, 157–213.

Valentinyi, A. and Herrendorf, B. (2008), 1'Measuring Factor Income Shares at the Sectoral Level", *Review of Economic Dynamics* **11**(4), 820–835.

Yang, S. S.-C. (2005), "Quantifying Tax Effects under Policy Foresight", *Journal of Monetary Economics* **52**(8), 1557–1568.

Supplementary Appendix
(For online publication)

Appendix A. The Structure of the Monetary Search Model

Appendix A.1. Preferences and Technology

Agents' per-period preferences are identically represented by

$$(q^b, q^s, k, X, N, Z) \mapsto u(q^b) - c(q^s/Z, k/Z) + U(X) - h(N),$$

where $u(q^b)$ is the per-period payoff from consuming a special good $q^b \in \mathbb{R}_+$, Z is the aggregate labor-augmenting technology, $c(q^s/Z, k/Z)$ is the utility cost of producing $q^s \neq q^b$ with fixed within-period capital, k. q^s and q^b are the tradable goods in the DM, where s denotes sold good and b denotes bought good.[10] $U(X)$ is utility of consuming the CM general good X, and, $-h(N)$ is the disutility of work effort N in the CM.[11]

Appendix A.2. Stationary Markov Decision Processes

Let the vector of aggregate state variables at the beginning of the DM be $\hat{\mathbf{s}}_t := (M_{t-1},\ K_{t-1},\ Z_t,\ \psi_t,\ \hat{\mu}_t,\ P_t)$, where M_{t-1} is the aggregate money stock; K_{t-1} is the aggregate capital stock; the aggregate labor-augmenting technology Z_t is determined at the beginning of period t; $\psi_t - 1$ is money supply growth rate (determined at the beginning of period t); and $\hat{\mu}_t := \hat{\mu}(\cdot|Z_t, \psi_t) : \mathcal{B}(\mathbb{R}_+) \to [0, 1]$ is a probability measure defined on the measure space of money holdings $(\mathbb{R}_+, \mathcal{B}(\mathbb{R}_+))$. The price level of X, P_t, is included as an auxiliary state variable, since we will focus on stationary Markovian equilibria (see Duffie et al., 1994). Denote m_{t-1} and k_{t-1} as stocks of individual nominal money and capital holdings, determined at the end of period $t - 1$.

Similarly, let $\mathbf{s}_t := (M_{t-1},\ K_{t-1},\ Z_t,\ \psi_t,\ \mu_t,\ P_t)$ denote the aggregate state vector at the beginning of the CM subperiod, in period t. Since money would have changed hands at the end of the DM, the distribution of money holdings would have evolved from $\hat{\mu}_t$ in the DM to μ_t at the start of the CM. At time t, $\mathbf{s}_{t+1}$ is a random vector.

Appendix A.2.1. DM Meeting Process

We assume that there is a probability $\sigma \leq 1/2$ that each agent can access the DM as a buyer of a particular good q^b. With symmetric probability σ, the agent can access the DM to sell his specific q^s. With probability $1 - 2\sigma$, the agent will leave the DM with no exchange. For the sake of simplicity, we assume that "double-coincidence-of-wants" events (where buyers and sellers in the DM are able to barter) and events where the agent can buy q^b and sell q^s simultaneously, both occur with probability zero.

Appendix A.2.2. DM Decision Process

Let $V(m_{t-1}, k_{t-1}, \hat{\mathbf{s}}_t)$ denote the optimal value of an agent at the beginning of the current period in the DM with state $(m_{t-1}, k_{t-1}, \hat{\mathbf{s}}_t)$, where m_{t-1} and k_{t-1} denote the individual money holding

[10] It turns out that in the equilibrium $q^s = q^b = q$ in this model, due to the degeneracy of the distribution of money holding.

[11] Or equivalently, let N_{DM} be the labor effort of an agent expended in a DM. Suppose the production technology, $(N_{DM}, k, Z) \mapsto \tilde{F}(ZN_{DM}, k)$ using capital and labor, is bijective and homogeneous of degree one. Then $Z \cdot N_{DM} = \tilde{F}^{-1}(q^s/k) \cdot k$ and $c(q^s/Z, k/Z) \equiv N_{DM}$.

and individual capital stock, respectively. The Bellman functional characterizing the value function $(m_{t-1}, k_{t-1}, \hat{\mathbf{s}}_t) \mapsto V(m_{t-1}, k_{t-1}, \hat{\mathbf{s}}_t)$ is given by

$$V(m_{t-1}, k_{t-1}, \hat{\mathbf{s}}_t) = \sigma V^b(m_{t-1}, k_{t-1}, \hat{\mathbf{s}}_t) + \sigma V^s(m_{t-1}, k_{t-1}, \hat{\mathbf{s}}_t) \\ + (1-2\sigma) W(m_{t-1}, k_{t-1}, \mathbf{s}_t), \tag{A.1}$$

where the indirect utilities $V^b(m_{t-1}, k_{t-1}, \hat{\mathbf{s}}_t)$ and $V^s(m_{t-1}, k_{t-1}, \hat{\mathbf{s}}_t)$ are determined by a particular pricing protocol in the DM, and $(m_{t-1}, k_{t-1}, \mathbf{s}_t) \mapsto W(m_{t-1}, k_{t-1}, \mathbf{s}_t)$ is the value function for the agent at the start of the CM, to be characterized by the CM decision process in the next section. The assumption in equation (A.1) is that there is no discounting between the DM and CM within the same time period t.

The competitive price-taking assumption for the DM trades implies each *ex post* buyer's problem as:

$$V^b(m_{t-1}, k_{t-1}, \hat{\mathbf{s}}_t) = \max_{q_t^b \in [0, m_{t-1}/\tilde{p}_t]} \left[u(q_t^b) + W\left(m_{t-1} - \tilde{p}_t q_t^b, k_{t-1}, \mathbf{s}_t\right) \right],$$

where $\tilde{p}_t$ is the price of the special good q_t^b and q_t^s, and is taken as given by all buyers and sellers. Each *ex post* seller's problem is:

$$V^s(m_{t-1}, k_{t-1}, \hat{\mathbf{s}}_t) = \max_{q_t^s} \left[-c(q_t^s/Z_t, k_{t-1}/Z_t) + W\left(m_{t-1} + \tilde{p}_t q_t^s, k_{t-1}, \mathbf{s}_t\right) \right].$$

Appendix A.2.3. CM Decision Processes

Let $\delta \in [0,1]$ be the depreciation rate of capital. Denote the competitive rate of return to physical capital by $r_t := r(\mathbf{s}_t)$. Similarly, denote $w_t := w(\mathbf{s}_t)$ as the real wage rate for labor, where each agent's labor supply decision is $N_t := N(m_{t-1}, k_{t-1}, \mathbf{s}_t)$. Denote each individual's CM consumption decision as $X_t := X(m_{t-1}, k_{t-1}, \mathbf{s}_t)$. Let $m_t := m(m_{t-1}, k_{t-1}, \mathbf{s}_t)$ and $k_t := k(m_{t-1}, k_{t-1}, \mathbf{s}_t)$ be, respectively, the money and capital holdings decisions for each individual with the state $(m_{t-1}, k_{t-1}, \mathbf{s}_t)$. Let $P_t := P(\mathbf{s}_t)$ be the competitive price of X_t, and $TR_t := TR(\mathbf{s}_t)$ be the aggregate lump-sum transfer from a monetary authority to the agent.

At the beginning of the CM sub-period, an agent with the state $(m_{t-1}, k_{t-1}, \mathbf{s}_t)$ solves the recursive problem of:

$$W(m_{t-1}, k_{t-1}, \mathbf{s}_t) = \max_{X_t, N_t, m_t, k_t} \left\{ U(X_t) - \phi N_t + \beta \mathbb{E}_\lambda \left[V(m_t, k_t, \hat{\mathbf{s}}_{t+1}) \middle| (Z_t, \psi_t) \right] \right\}, \tag{A.2}$$

subject to

$$\mathbf{s}_{t+1} = \mathcal{G}(\mathbf{s}_t, \mathbf{v}_{t+1}), \qquad \mathbf{v}_t \overset{\text{i.i.d.}}{\sim} \varphi, \tag{A.3}$$

$$X_t + k_t - (1-\delta) k_{t-1} = \frac{m_{t-1} - m_t}{P_t} + w_t N_t + r_t k_{t-1} + TR_t, \tag{A.4}$$

where ϕ is a constant representing the relative importance of CM consumption and leisure in the utility function W; $\lambda(\mathbf{s}_t, \cdot)$ is induced by $\mathcal{G} \circ \varphi$ in equation (A.3) for each given $\mathbf{s}_t$, and defines an equilibrium product probability measure over Borel-subsets containing $\mathbf{s}_{t+1}$. This is a rational expectations constraint that ensures consistency of beliefs in equilibrium. Implicit in constraint (A.3) is the equilibrium transition of the distribution of individual states from the period-t CM, to the period-$(t+1)$ DM, $\hat{\mu}(\hat{\mathbf{s}}_{t+1}, \cdot) =$

$\mathcal{G}_{\hat{\mu}}[\mu(\mathbf{s}_t,\cdot),\mathbf{z}_{t+1}]$, such that the relevant conditional distribution of assets at the beginning of the period-$(t+1)$ CM subperiod is given by $\mu(\mathbf{s}_{t+1},\cdot) = \mathcal{G}_{\mu}[\hat{\mu}(\hat{\mathbf{s}}_{t+1},\cdot),\mathbf{z}_{t+1}] \equiv \mathcal{G}_{\mu} \circ \mathcal{G}_{\nu}(\mathbf{s}_t,\mathbf{z}_{t+1})$, where $\mathcal{G}_{\mu}$ and $\mathcal{G}_{\nu}$ are components of the Markov equilibrium map $\mathcal{G}$. The sequential one-period budget constraint is given by equation (A.4).

Production in the CM is given by the following representative firm's problem:

$$\max_{K_{t-1},N_t^d} \left\{ F(K_{t-1}, Z_t N_t^d) - w_t N_t^d - r_t K_{t-1} \right\},$$

where $F(\cdot,\cdot)$ is a production function, N_t^d is aggregate labor demand by the representative firm in the CM.

Appendix A.2.4. Exogenous Processes

We assume that the money supply growth factor, $\psi_t := M_t/M_{t-1}$, follows an AR(1) process:

$$\ln(\psi_t) = \rho_M \ln(\psi_{t-1}) + \sigma_M \varepsilon_{\psi_t}, \qquad \text{with} \quad \varepsilon_{\psi_t} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1).$$

Stationarity condition requires $0 < \rho_M < 1$.

Following Ireland and Schuh (2008), we specify the technology stochastic process as an AR(1) in its growth factor, $\tilde{Z}_{t+1} := Z_{t+1}/Z_t$:

$$\ln(\tilde{Z}_t) = \rho_Z \ln(\tilde{Z}_{t-1}) + \sigma_Z \varepsilon_{Z_t}, \qquad \text{with} \quad \varepsilon_{Z_t} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1),$$

where $0 < \rho_Z \leq 1$.

Appendix A.2.5. Market Clearing

In the equilibrium, the resource constraint in the CM must hold such that

$$F(K_{t-1}, Z_t N_t) = X_t + K_t - (1-\delta)K_{t-1}. \tag{A.5}$$

Also, the monetary authority's budget constraint must hold,

$$TR_t = \frac{M_t - M_{t-1}}{P_t}. \tag{A.6}$$

These, together with the agent's CM budget constraint (equation (A.4)) in equilibrium, imply that the labor market in the CM must clear as well, *i.e.* $N_t^d = N(\mathbf{s}_t) := \int_{\mathbb{R}_+} N(m_{t-1}, k_{t-1}, \mathbf{s}_t) d\mu_t$.

Appendix A.3. Stationary Monetary Equilibrium

The optimal decision processes and market clearing conditions will give rise to a set of functional equations which characterize the necessary conditions for a stationary monetary equilibrium. We will require more structure on the equilibrium.[12] In particular, we seek a stationary Markov monetary

[12] As a general rule, monetary models such as this can induce many other interesting types of equilibria, including chaotic and sunspot equilibria (see e.g. Lagos and Wright, 2003). However, from an econometric perspective, these equilibria may not be so amenable to econometric analysis.

equilibrium (SME) which is given by allocation and pricing functions that are time-invariant, and depend on past outcomes only through the current state $\mathbf{s}_t$.

We assume the following functional forms:

$$U(X) = \varrho \ln(X), \qquad h(N) = \phi N, \qquad F(K, ZN) = K^{\alpha}(ZN)^{1-\alpha},$$

where ϱ, $\phi > 0$, $\alpha \in (0, 1)$, and

$$u(q) = \ln(q + \underline{q}) - \ln(\underline{q}), \qquad c(q/Z, K/Z) = Z^{-1} q^{\varpi} (K)^{1-\varpi},$$

where $\underline{q} > 0$ is a constant, and $\varpi \geq 1$.

From the first-order conditions of the CM decision problem in equations (A.2)-(A.4) with respect to m_t and k_t, we can deduce that the optimal decision rules for m_t and k_t do not depend on individual states (m_{t-1}, k_{t-1}).[13] Therefore, in equilibrium, all agents exiting from each CM will appear identical in terms of their individual states $(m_{t-1}, k_{t-1}) = (M_{t-1}, K_{t-1})$ for all (m_{t-1}, k_{t-1}). Hence, we can characterize the equilibrium allocations as functions of the aggregate outcomes only — *i.e.*, in terms of "big-M" and "big-K" only — and the labor allocation N_t will be in terms of the aggregate as well.

We transform the original problem into one in terms of stationary variables. Due to the presence of a unit root in the $\{Z_t\}$ process, the real allocations in the equilibrium will inherit the unit root as well. We perform the following transformations: $\tilde{X}_t := X_t/Z_t$, $\tilde{K}_t := K_t/Z_t$, $\tilde{q}_t := q_t/Z_t$, and $\tilde{P}_t := Z_t P_t / M_{t-1}$, and then denote the SME decision and pricing functions as $(\tilde{K}_t, \tilde{q}_t, \tilde{X}_t, \tilde{P}_t, N_t) := (\tilde{K}(\mathbf{s}_t), \tilde{q}(\mathbf{s}_t), \tilde{X}(\mathbf{s}_t), \tilde{P}(\mathbf{s}_t), N(\mathbf{s}_t))$. Given any function $f(x, y, \cdots)$, denote the (partial) derivative of $f(x, y, \cdots)$ with respect to x by $f'_x(x, y, \cdots)$. The SME is defined as follows.

Definition 1 (SME). *Given the exogenous processes $\{\tilde{Z}_t, \psi_t\}_{t \in \mathbb{N}}$, a SME consists of bounded stochastic processes $\{\tilde{K}_t, \tilde{q}_t, \tilde{X}_t, \tilde{P}_t, N_t\}_{t \in \mathbb{N}}$, satisfying the following conditions:*

1. *Optimal investment:*

$$U'_X(\tilde{X}_t) = \beta \mathbb{E}_{\lambda} \left\{ \frac{U'_X(\tilde{X}_{t+1})}{\tilde{Z}_{t+1}} \left[F'_K(\tilde{K}_t/\tilde{Z}_{t+1}, N_{t+1}) - \delta \right] - \sigma \frac{c'_K(\tilde{q}_{t+1}, \tilde{K}_t/\tilde{Z}_{t+1})}{\tilde{Z}_{t+1}} \middle| (\tilde{Z}_t, \psi_t) \right\}, \tag{A.7}$$

2. *Inter-temporal optimal money holdings:*

$$U'_X(\tilde{X}_t) = \beta \mathbb{E}_{\lambda} \left\{ U'_X(\tilde{X}_{t+1}) \frac{\tilde{P}_t}{\psi_t \tilde{P}_{t+1}} \times \left[(1-\sigma) + \sigma \frac{u'_q(\tilde{q}_{t+1})}{c'_q(\tilde{q}_{t+1}, \tilde{K}_t/\tilde{Z}_{t+1})} \right] \middle| (\tilde{Z}_t, \psi_t) \right\}, \tag{A.8}$$

3. *Labor market clearing:*

$$U'_X(\tilde{X}_t) = \frac{\phi}{F'_N(\tilde{K}_{t-1}/\tilde{Z}_t, N_t)}, \tag{A.9}$$

[13] This is a result of the quasi-linearity in the preference functions, *i.e.* there are no wealth effects.

4. DM price-taking solution:

$$\frac{U'_X(\tilde{X}_t)}{\tilde{P}_t}\psi_t = c'_q(\tilde{q}_t, \tilde{K}_{t-1}/\tilde{Z}_t)\tilde{q}_t, \tag{A.10}$$

5. Resource constraint:

$$\tilde{X}_t + \tilde{K}_t + \tilde{G}_t = F(\tilde{K}_{t-1}/\tilde{Z}_t, N_t) + (1-\delta)\tilde{K}_{t-1}/\tilde{Z}_t. \tag{A.11}$$

Appendix A.4. Auxiliary Variable Definitions

This model now has two sectors, the DM and the CM, so we would like to define an aggregate measure of output and employment for the economy as a whole. First, note that the DM price is determined from the DM terms of trade definition $\tilde{p}_t = M_t/q_t$. Therefore, in its stationary form we have $\tilde{\tilde{p}}_t\tilde{q}_t = \psi_t$. The CM total output, in units of the CM final good, is

$$\tilde{X}_{CM,t} = F(\tilde{K}_{t-1}/\tilde{Z}_t, N_t).$$

The DM nominal output, using $\tilde{P}_t$ as the unit of account, is

$$X^{nom}_{DM,t} = \frac{\sigma\tilde{P}_t}{\phi\psi_t}\left[F'_N(\tilde{K}/\tilde{Z}_t, N_t)\right]c'_q(\tilde{q}_t, \tilde{K}_{t-1}/\tilde{Z}_t)\tilde{q}_t,$$

where we make use of the equilibrium DM price taking solution. Define the share of DM output value in the total output value as

$$\chi_t := \frac{X^{nom}_{DM,t}}{X^{nom}_{DM,t} + \tilde{P}_t\tilde{X}_{CM,t}}.$$

Note that this share is time-varying since it is also dependent on the period-t aggregate state $\mathbf{s}_t$. We can now define our measure of aggregate price index as

$$\tilde{P}_{X,t} = \chi_t\tilde{\tilde{p}}_t + (1-\chi_t)\tilde{P}_t.$$

The total real output in this two sector economy is defined as

$$\tilde{X}_{tot,t} = \frac{X^{nom}_{DM,t} + \tilde{P}_t\tilde{X}_{CM,t}}{\tilde{P}_{X,t}}. \tag{A.12}$$

Total labor includes employment in the CM, and also labor effort in DM. In terms of the stationary equilibrium, the total employment is given by

$$N_{tot,t} = \sigma c(\tilde{q}_t, \tilde{K}_{t-1}/\tilde{Z}_t) + N_t. \tag{A.13}$$

Denote the percentage deviations of $N_{tot,t}$ and $\tilde{X}_{tot,t}$ in equations (A.12) and (A.13) as $n_{tot,t}$ and $\tilde{x}_{tot,t}$ respectively. In terms of the corresponding observable variables, employment and the growth of real output, we now have $y_t := (n_{tot,t},\ \Delta\ln X_{tot,t})'$.

Appendix A.5. Calibration

We parameterize the model according to the monetary model literature; see Schlagenhauf and Wrase (1995); Chari et al. (2002); Heathcote and Perri (2002); Ireland and Schuh (2008). First, the discount factor β is set to be 0.99; the capital depreciation rate δ is set to be 2.5%; the share of capital income α is set to be 1/3; and the probability of entering DM as a buyer or seller is $\rho = 0.26$.

As for the parameters in the exogenous shock processes, the steady state values of both technology and the gross money supply are set to be 1, because they do not affect the dynamics. The AR(1) coefficients are $\rho_M = 0.5857$ and $\rho_Z = 0.6$, while the standard deviations are $\sigma_M = 0.00397$ and $\sigma_Z = 0.007$.

We calibrate the remaining parameters (ϕ, ϱ, ϖ) to match the targets of the proportion of total hours worked (DM and CM aggregate), N_{tot}, the velocity of money as defined in the work of Aruoba et al. (2008), and the long run capital-output ratio, K/X_{tot}. The value of $\bar{N}_{tot}$ is 0.33, which is standard. This helps us to pin down the calibration of the disutility of labor in the CM parameter, $\phi = 4.966$. The velocity of money is around 1.3225 per quarter in the data for the M1 definition of monetary aggregate. This is used to pin down the calibration of the utility weight of consuming X_t, which is $\varrho = 0.754$. The target capital-output ratio is 2.23 in annual terms. The calibrated value of $\varpi = 1.289$ implies that the more capital is installed for use in the DM production, the lower the cost of producing a unit of DM output q_t.

Tasmanian School of Business and Economics Working Paper Series

2014-09 VAR Modelling in the Presence of China's Rise: An Application to the Taiwanese Economy. **Mardi Dungey**, Tugrul Vehbi and **Charlton Martin**

2014-08 How Many Stocks are Enough for Diversifying Canadian Institutional Portfolios? **Vitali Alexeev** and Francis Tapon

2014-07 Forecasting with EC-VARMA Models, George Athanasopoulos, Don Poskitt, Farshid Vahid, **Wenying Yao**

2014-06 Canadian Monetary Policy Analysis using a Structural VARMA Model, **Mala Raghavan**, George Athanasopoulos, Param Silvapulle

2014-05 The sectorial impact of commodity price shocks in Australia, S. Knop and **Joaquin Vespignani**

2014-04 Should ASEAN-5 monetary policymakers act pre-emptively against stock market bubbles? **Mala Raghavan** and **Mardi Dungey**

2014-03 Mortgage Choice Determinants: The Role of Risk and Bank Regulation**, Mardi Dungey, Firmin Doko Tchatoka, Graeme Wells, Maria B. Yanotti**

2014-02 Concurrent momentum and contrarian strategies in the Australian stock market, Minh Phuong Doan, **Vitali Alexeev**, Robert Brooks

2014-01 A Review of the Australian Mortgage Market**, Maria B. Yanotti**

2013-20 Towards a Diagnostic Approach to Climate Adaptation for Fisheries, P. Leith, E. Ogier, G. Pecl, **E. Hoshino**, J. Davidson, M. Haward

2013-19 Equity Portfolio Diversification with High Frequency Data, **Vitali Alexeev** and **Mardi Dungey**

2013-18 Measuring the Performance of Hedge Funds Using Two-Stage Peer Group Benchmarks, Marco Wilkens, Juan Yao, **Nagaratnam Jeyasreedharan** and Patrick Oehler

2013-17 What Australian Investors Need to Know to Diversify their Portfolios, **Vitali Alexeev** and Francis Tapon

2013-16 Equity Portfolio Diversification: How Many Stocks are Enough? Evidence from Five Developed Markets**, Vitali Alexeev** and **Francis Tapon**

2013-15 Equity market Contagion during the Global Financial Crisis: Evidence from the World's Eight Largest Economies, **Mardi Dungey** and **Dinesh Gajurel**

2013-14 A Survey of Research into Broker Identity and Limit Order Book, **Thu Phuong Pham** and P Joakim Westerholm

2013-13 Broker ID Transparency and Price Impact of Trades: Evidence from the Korean Exchange, **Thu Phuong Pham**

UNIVERSITY *of* TASMANIA
AUSTRALIA