On Weak Identification in Structural VAR(MA) Models

Wenying Yao* Timothy Kam† Farshid Vahid‡

Abstract

We simulate synthetic data from known data generating processes (DGPs) that arise from economic theory, and compare the performance of VAR and VARMA models in fitting the dynamics of our DGPs. We show that in the case of small samples typical of macroeconomic data, the moving average (MA) component of 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 in approximately characterizing the known DGPs. We provide a new multivariate insight into why this is so. Our extended findings suggest that there are pitfalls in relying on VAR/VARMA representations for identifying policy and market-relevant shifts underlying observed macroeconomic data.

Keywords: VARMA; VAR; DSGE; impulse response analysis
JEL Classification: C15; C52; C32

1 Introduction

Macroeconomists have long recognized that approximate solutions to theoretical dynamic stochastic general equilibrium (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). 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 data generating process (DGP). We, as do others, show that the correct VARMA structure is not identified in realistic sample

*Department of Economics, Faculty of Business and Law, Deakin University, Burwood VIC 3125, Australia; Tasmanian School of Business and Economics, Private Bag 84, University of Tasmania, Hobart, TAS 7001, Australia. E-mail address: wenying.yao@deakin.edu.au.

†Research School of Economics, The Australian National University, Canberra, ACT 2601, Australia. E-mail address: tcy.kam@gmail.com.

‡Department of Econometrics and Business Statistics, Monash University, Clayton VIC 3800, Australia. E-mail address: farshid.vahid@monash.edu.

We thank Heather Anderson, Mardi Dungey, Christian Kascha, Eric Leeper, Don Poskitt, and Shaun Vahey for their valuable comments. This research was supported by the Australian Research Council Grant DP0984399. This verion: August 22, 2016. An Online Appendix containing more details and results can be accessed from https://github.com/phantomachine/_varmita.

sizes, which in turn leads to unreliable inference of the structure of the true DGP. Kascha and Mertens (2009) attribute it to the fact that the DGPs used in the previous literature (in particular Christiano et al., 2006; Chari et al., 2007) are:

> "[n]*early non-stationary, nearly non-invertible and the correct VARMA representation is close to being not identified.*"

In contrast to the claim of Kascha and Mertens (2009), we show that the identification problem also arises in a counterexample where the true VARMA DGP is strictly invertible and far away from being non-stationary. We provide some more general analytical understanding of this problem. 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, and this is true regardless of whether the system is close to being non-stationary and non-invertible, or not. 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. Here, we provide a multivariate generalization of this insight.

Our experimental design follows and extends the work of Kascha and Mertens (2009), Erceg et al. (2005), and Ravenna (2007). We consider as a benchmark, the stylized real business cycle (RBC) model by Hansen (1985) and derive its equilibrium-restricted VARMA representation (i.e., the true DGP taken by us to be known with certainty). Within our experiments, a hypothetical econometrician fits reduced-form VAR and VARMA models to simulated data from the DGP. As a further contribution, we conduct such a thought experiment across a wide class of DGPs, each with increasing layers of dynamic sophistication.

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. Section 3 discusses the near cancellation in the AR and MA dynamics in the resulting VARMA DGP from the RBC model. Section 4 experiments with several alternative DGPs that come from more complex DSGE models to examine the weak identification of the correct VARMA structure. Section 5 concludes.

2 DGP 1: Hansen's RBC and VARMA Identification

We begin with the most used DGP assumption in the literature: The indivisible labor RBC model of Hansen (1985). There are two exogenous structural shocks in the model—i.e., a non-stationary technology shock Z_t, and a stationary labor supply shock D_t. 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 $\mathbb{E}_0 \sum_{t=1}^{\infty} \beta^t [\ln C_t + \phi D_t (1 - N_t)]$, subject to capital accumulation and production technologies, respectively, $K_t = X_t - C_t + (1-\delta) K_{t-1}$, and, $X_t = K_{t-1}^{\alpha} (Z_t N_t)^{1-\alpha}$, where $\alpha, \beta \in (0, 1)$, and $\delta \in (0, 1]$. The labor-augmenting technology level Z_t and the labor supply shifter D_t follow exogenous stochastic processes, respectively defined by, $\ln Z_t = \ln Z_{t-1} + \mu_z + \varepsilon_{z_t}$, where $\varepsilon_{z_t} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_z^2)$, and, $\ln D_t = (1 - \rho_d) \ln \bar{D} + \rho_d \ln D_{t-1} + \varepsilon_{d_t}$, where $\rho_d \in (0, 1)$ and $\varepsilon_{d_t} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma_d^2)$. 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 DGP and experimental design

The RBC model is parameterized following Erceg et al. (2005) and Ravenna (2007). This is reported in the first column of Table 1. This RBC instance implies an equilibrium-restricted VARMA(1,1) representation in terms of y_t:

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

where u_t is the reduced form disturbance with zero mean and non-singular covariance matrix Σ_u. The VARMA process (1) is strictly stationary and invertible. The eigenvalues of the AR and MA coefficient matrices are given in Table 1.

We will now fix (1) as the true DGP in the following Monte Carlo experiments: Consider econometricians within the experiments who only observe data generated by (1). They then fit atheoretical VAR and VARMA structures to each simulated sample path, and consider them as competing models. We then examine whether the reduced-form VAR or VARMA statisticians are able to match the impulse dynamics of the true VARMA DGP (1).

We consider two cases in every experiment: From the DGP (1), we simulate large-sample-path ($T = 20,000$), and, small-sample-path ($T = 200$) scenarios. Each of such paths is replicated or sampled 1,000 times. The setting of $T = 200$ corresponds to 50 years of quarterly data, which is similar to the sample sizes examined in previous studies (see, e.g., Kascha and Mertens, 2009), and represents a typical sample size for macroeconomic data. This experimental design is repeated in all remaining DSGE models used as "true DGPs" in alternative experimental cases.

The specification of the lag order in VAR models allows for the estimation of all parameters. However, VARMA models are not always fully identified in terms of their specification of AR and MA orders. We need to specify the largest lag orders in each equation in order to estimate a VARMA model, because the lag orders in one equation can have implication for the identification of parameters in other equations. Consider the VARMA(1,1) example below:

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix} + \begin{pmatrix} \theta_{11} & \theta_{12} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} u_{1,t-1} \\ u_{2,t-1} \end{pmatrix}.$$

The fact that the longest lag in the second equation is zero implies that ϕ_{12} and θ_{12} are not separately identifiable.

Therefore, we need to determine the orders of each equation and impose additional zero

Table 1: Properties of the DSGE models, their implied VARMA DGPs and estimation results on simulated data

		(1) RBC	(2) HI[a]	(3) NEWS	(4) MS
Key parameter values of the models					
	α	0.35	0.36	0.36	1/3
	β	$1.03^{-0.25}$	0.99	0.99	0.99
	δ	2%	$\delta_0 = 2.5\%$	2%	2.5%
	$\bar{N}$	1/3	1/3	1/3	1/3
	μ_z	0.0037	0.0037	0.0037	1
	σ_z	0.0148	0.0148	0.009	0.007
	ρ_d	0.8	0.8	0.8	
	σ_d	0.009	0.009	0.005	
			$\bar{U} = 1$	$\rho_G = 0.99$, $\sigma_G = 0.018$	$\rho = 0.26$
			$\theta_c = 0.7$	$\rho_K = 0.97$, $\sigma_K = 0.025$	$\rho_m = 0.5857$
			$\theta_l = 0.5$	$\rho_L = 0.99$, $\sigma_L = 0.020$	$\rho_z = 0.6$
			$\delta_2 = 0.11$	$\tau_K = 0.39$, $\tau_L = 0.21$	$\sigma_m = 0.00397$
			$\varsigma = 4$	$r_{K,L} = 0.261$	
Properties of the DGPs:					
eig(AR)[b]		0.9459, 0.8	0.99, 0.79, 0.0002, 0	0.98, 0.90, 0.8, 0	0.9386, 0.6, 0.5857, 0
eig(MA)		0.9563, 0	0.91, 0.20	0.92, 0.92	0.9425, 0.0087
True VARMA orders		(1,1)	(2,1)	(2,1)	(2,1)
Correct SCM structure		(1,1)∼(1,0)	(2,1)∼(1,1)	(2,1)∼(1,1)	(2,1)∼(1,1)
VAR(MA) estimation results on simulated data from DGPs:					
$T = 20,000$	% of corr.id. SCM[c]	93.1%	85.9%	97.0%	93.5%
	Median VAR lag[d]	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[e]	(1,0)∼(0,0)	(1,1)∼(1,0)	(1,0)∼(0,0)	(1,1)∼(1,0)
	% of most.id. SCM	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 4.1, NEWS stands for the model with news shocks from Section 4.2, and MS denotes for monetary search model discussed in Section 4.3.

[b] eig(AR) and eig(MA) denote roots of the AR and MA characteristic polynomials of the true (RBC and MS) and estimated (HI and NEWS) VARMA processes.The MA roots for NEWS model are calculated from the corresponding fundamental representation of the VARMA process.

[c] Represents the percentage of correct identification of the SCM structure when the sample size is $T = 20,000$.

[d] Represents the median of the lag lengths chosen by AIC for VAR using $T = 20,000$.

[e] Represents the mostly identified SCM structure when the sample size is $T = 200$. The percentage of identifying this SCM is shown in the row below.

and normalization restrictions before estimating the VARMA model. This can be done using 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.

2.2 Large-sample simulation

For each simulated large sample path, the structure of the VARMA models used by the econometrician within each experiment is specified using the scalar components model (SCM) methodology. For example, most sample paths simulated from process (1) are identified as a VARMA(1,1) model with SCM(1,1) ~ SCM(1,0). The canonical form for this SCM VARMA structure is

$$\begin{pmatrix} 1 & 0 \\ a_0 & 1 \end{pmatrix} y_t = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} y_{t-1} + u_t + \begin{pmatrix} \theta_{11} & \theta_{12} \\ 0 & 0 \end{pmatrix} u_{t-1}, \tag{2}$$

where the SCM(1,0) (second row) does not have an MA lag. There are also some zero and normalization restrictions imposed on the left-hand side transformation matrix, in order to ensure a unique identification of the unknown parameters.

Out of the total number of simulated sample paths, the percentage instances of identifying the correct VARMA(1,1) model with SCM(1,1) ~ SCM(1,0) is shown in Table 1 as 95.6%. This is the simplest structure of the underlying true DGP for y_t.[1] 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. Most of the chosen lags are higher than 15.

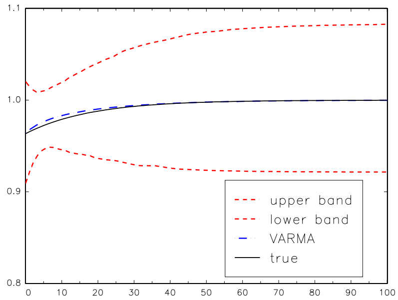
The impulse responses of $\ln X_t$ and n_t to the technology shock ε_{z_t} are plotted in Figure 1 up to 100 periods after a shock occurs. It shows the mean as well as the 2.5$^{\text{th}}$ and 97.5$^{\text{th}}$ percentiles of the impulse responses estimated from fitted VARs and VARMA models. In each 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$^{\text{th}}$ and 97.5$^{\text{th}}$ percentiles is referred to as the "97.5/2.5 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 97.5/2.5 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 97.5/2.5 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

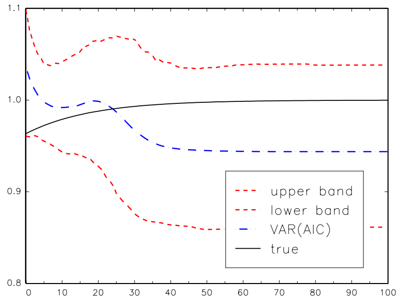
[1] In the other 4.4% of the time, the SCM methodology always finds a structure with higher orders that nests the SCM(1,1) ~ SCM(1,0).

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

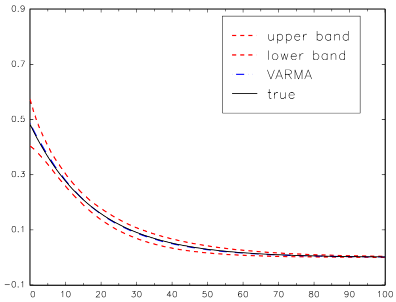
(a) $\ln X_t$ to ε_{z_t} (VARMA)



(b) $\ln X_t$ to ε_{z_t} (VAR)



(c) n_t to ε_{z_t} (VARMA)



(d) n_t to ε_{z_t} (VAR)

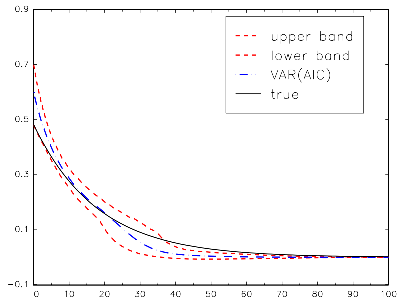
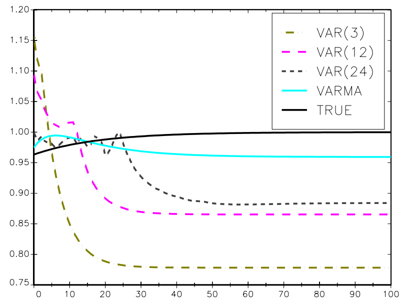
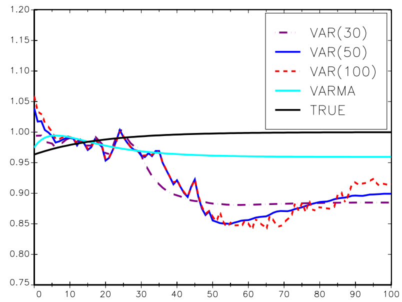


Figure 2: Estimated impulse responses from one particular sample path (large sample)

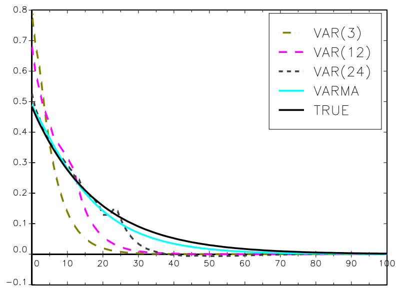
(a) $\ln X_t$ to ε_{z_t} (short VARs)



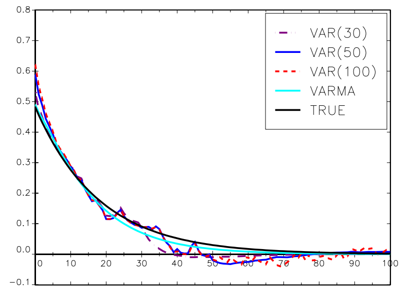
(b) $\ln X_t$ to ε_{z_t} (long VARs)



(c) n_t to ε_{z_t} (short VARs)



(d) n_t to ε_{z_t} (long VARs)



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(1,1) $\sim$ SCM(1,0). 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.[2] 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 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.

2.3 Small-sample simulation

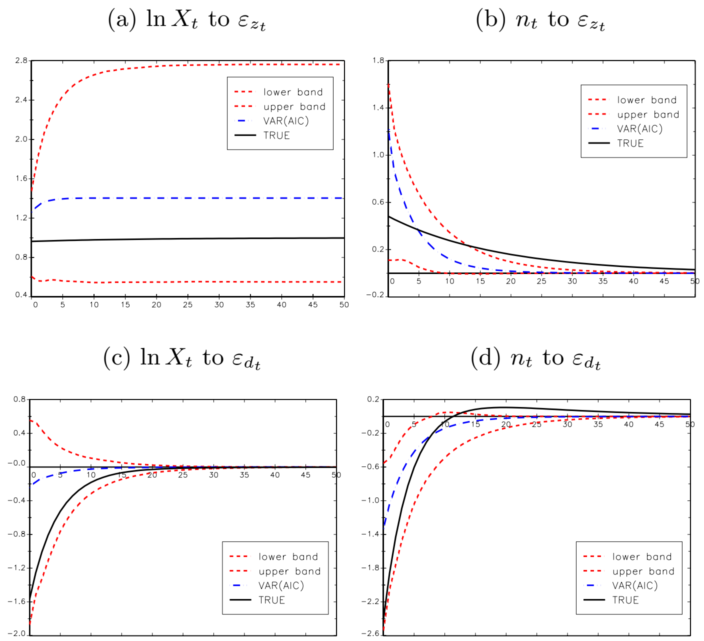
When working with empirical macroeconomic data, usually there are only a limited number of observations available. Hence, the comparison of VAR and VARMA models based on small samples is crucial for practitioners. Unfortunately, the SCM identification procedure for VARMA models always fails to detect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

Figure 3 depicts the mean and 97.5/2.5 interval 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

[2] 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.

will appear to be less persistent. Kilian (2011) suggests 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 and Mertens (2009).[3]

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}

3 Examining the AR and MA roots

We would like to undestand why there is, with small sample sizes, non-identification of the AR and MA components of the true DGP in (2). Here we decompose analytically and show, by simulating over many parametric instances of the DGP, why that is the case. To do so, we can rotate each parametric instance of the DGP VARMA system in such a way that we can focus on the two-dimensional space containing an equivalent scalar AR and a scalar MA parameter. Note that this rotation still preserves the original SCM(1,1) $\sim$ SCM(1,0) structure. This makes our analysis algebraically and graphically more instructive.

Specifically, the algebraic transforms are as follows: In the instance of the RBC DGP, the resulting version of (2) as DGP is one where $\phi_{22} = 0$:

$$\begin{pmatrix} 1 & 0 \\ a_0 & 1 \end{pmatrix} y_t = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & 0 \end{pmatrix} y_{t-1} + u_t + \begin{pmatrix} \theta_{11} & \theta_{12} \\ 0 & 0 \end{pmatrix} u_{t-1}. \tag{2'}$$

[3]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.

Recall that the coefficients a_0, ϕ_{ij} and θ_{ij}, i, $j = 1, 2$, are functions of the deep parameters of the RBC model. Note that the MA component only appears in the first row of the system of equations.

As any non-singular linear transformation of y_t does not affect the overall orders of the VARMA process, we consider pre-multiplying equation (2') by the matrix $\begin{pmatrix} 1 & \kappa \\ 0 & 1 \end{pmatrix}$, where κ is some scalar to be determined. This preserves the original SCM(1,0) form for the second-row equation of (2'), whereas the first row is rotated as:

$$\Big[(1+\kappa a_0)-(\phi_{11}+\kappa\phi_{21})\,L\Big]n_t+(\kappa-\phi_{12}\,L)\Delta\ln X_t = u_{1t}+\kappa\,u_{2t}+\theta_{11}u_{1,t-1}+\theta_{12}u_{2,t-1}. \quad (3)$$

Equation (3) in principle will have a SCM(1,1) structure. However, order reduction can occur if all of the AR and MA roots in (3) turn out to be equal to the same value.

From equation (3), the two AR(1) coefficients for n_t and $\Delta\ln X_t$ are $AR_n^{(1)} = \frac{\phi_{11}+\kappa\phi_{21}}{1+\kappa a_0}$, and $AR_x^{(1)} = \frac{\phi_{12}}{\kappa}$, respectively. The terms on the left-hand side of equation (3) 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. Any value of κ corresponds to a point in the three-dimensional space $(AR_n^{(1)}(\kappa),\ AR_x^{(1)}(\kappa),\ \gamma(\kappa))$. In order to reduce the problem to two dimensions with only one AR coefficient and one MA coefficient, we find the value of κ that makes the two AR(1) coefficients the same, and then use this value of κ to calculate the MA(1) coefficient γ.[4] This is done for each experimental or parametric instance of the RBC DGP.

In the exact instance where $AR_n^{(1)} = AR_x^{(1)} = \gamma$, the first row equation (3) degenerates to a static equation with no lagged variables involved. As a result, the MA component is not detectable in the system (2'), whose overall orders will be dominated by the second row equation as a VARMA(1,0) process. More generally, even when each estimated instance of the triplets are not equal to each other but they are close, it still poses a challenge to the identification of the correct VARMA structure, particularly in small samples. Cogley and Nason (1993) came across the same situation in the setting of a univariate ARMA process, where the AR and MA lag polynomials have roughly the same factors that almost cancel each other out. What we have here is a multivariate generalization of this insight.

Given what we can learn from the algebraic transforms just described, we next experimentally evaluate the possibility of near cancellation by examining whether the AR(1) and MA(1) coefficients of the first row equation (3) always stay close to each other. The simulation procedure to examine the near cancellation of the AR and MA roots is as follows. We vary the true DGP by simulating the values of the deep parameters in the RBC model randomly from the distributions tabulated in Table 2, and then compute their implied VARMA(1,1) coefficients in equation (2'). The value of κ is obtained by equating the two AR(1) coefficients $\frac{\phi_{11}+\kappa\phi_{21}}{1+\kappa a_0} = \frac{\phi_{12}}{\kappa}$, which in turn allows us to record their transforms as equivalent scalar AR(1) and MA(1) coefficients. We should emphasize again that the computed κ for each

[4] In general this will yield two different values of κ. We choose the one that generates the smaller distance between the AR(1) and MA(1) coefficients.

DGP instance may not be the one that minimizes the differences among the set of triplets $(AR_n^{(1)}(\kappa),\ AR_x^{(1)}(\kappa),\ \gamma(\kappa))$, hence any result here is a lower bound on the actual severity of the near-cancellation problem.

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

parameter	distribution	mean	parameter	distribution	mean	variance
β	$\boldsymbol{U}[0.98, 1]$	0.99	μ_z	$\boldsymbol{U}[0, 0.01]$	0.005	
$\bar{N}$	$\boldsymbol{U}[0, 2/3]$	1/3	ρ_d	Beta	0.8	0.022
δ	$\boldsymbol{U}[0, 5\%]$	2.5%	σ_z	Inverse Gamma	0.0148	1
α	$\boldsymbol{U}[20\%, 60\%]$	40%	σ_d	Inverse Gamma	0.009	1

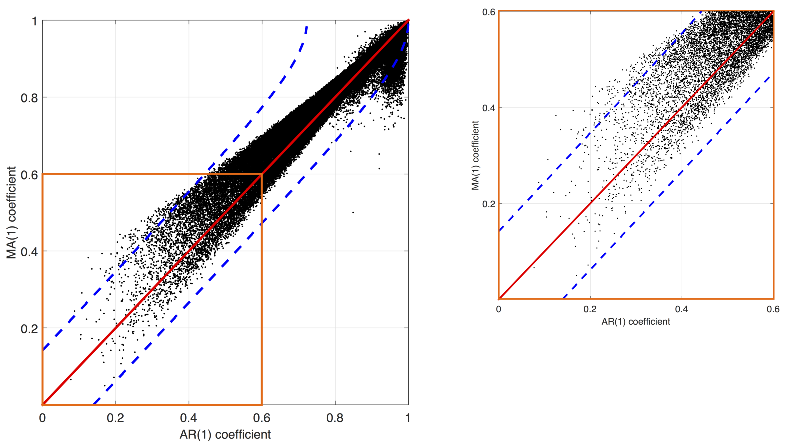
Uniform distributions are used for parameters on which we do not have strong priors, an idea borrowed from the Bayesian DSGE literature. The range of the steady state level of employment is set to be $[0, 2/3]$ so that its mean is the same as is in the previous benchmark parameterization. The capital 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.

Figure 4 plots corresponding AR(1) and MA(1) coefficient pairs, together with the 45-degree line. The latter is the set of all instances in which the DGP is exactly unidentifiable. The two dashed bands in Figure 4 outlines the region in 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, with a sample size $T = 200$, given combinations of AR(1) and MA(1) coefficients within these bands, the true first order autocorrelation will not be recognized as statistically significantly different from zero. Consequently, the first row equation (3) is likely to be identified as a SCM(0,0) structure, and hence the second row equation of SCM(1,0) dictates the overall orders of the system (2') as VARMA(1,0).

Figure 4: Instances of the RBC-DGPs projected as AR(1)-MA(1) coefficient pairs

(a) Entire feasible space

(b) Not "near-unit root" feasible subspace



We can see that from the experimental DGPs, 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: 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.

We ensured the invertibility of the VARMA processes in the simulation by setting $|\gamma| < 1$ in the rotated representation. Further, we can also check that the identification problem is not because of the roots of the RBC model's VARMA DGP being close to unit roots. To illustrate this point, we used $(0.6, 0.6)$ as an arbitrary cutoff point to demarcate roots that are sufficiently far away from being near unit root. From the experiment, out of 1 million simulated DGP instances, about 7% have both roots fall in the set $(0, 0.6]^2$, whereas about 33.5% have both roots fall within $(0, 0.8]^2$. Thus, even conditioning on DGP cases where the true processes are nowhere near unit root, the near cancellation problem is still quite prevalent: from Figures 4a and 4b we can see that a majority of these DGP instances are still inside the non-identification region, irrespective of whether the DGPs have roots that are near unity or not.

This reveals an important challenge in real-world applications: Even when one has a known DGP that implies a VARMA process, from a econometrician's point of view, it is difficult to distinguish this VARMA process from a finite order VAR statistically in small samples. In contrast to the claim of Kascha and Mertens (2009), our simulation experiments demonstrate that even when the VARMA process under investigation is far from being non-stationary and non-invertible, the identification difficulty still persists due to the closeness of the AR and MA roots.

4 Near-cancellation in other DGPs

In the basic RBC model above, we show that the roots of the AR and MA lag polynomials are almost 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, and, when we consider DGPs that are very far away from implying near-unit root stochastic processes. Therefore, we conclude that the similarity in the AR and MA dynamics is an inherent feature of that RBC model itself.

A natural question that arises is whether all VARMA DGPs implied by other types of DSGE models have near cancellation in the AR and MA dynamics. Although it is not possible to analyze the entire universe of alternative DGPs, we select a few simple but representative DSGE models which have richer and more complex internal 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, so we will not repeat them again. This exercise will shed some light on the ubiquity of the near cancellation problem.

4.1 DGP 2: Habit formation and investment adjustment cost

Consider a variant of the RBC model with habit formation in consumption C_t and leisure $L_t = 1 - N_t$, combined with costly investment adjustment. We provide the model description and justify the choice of this dynamically more interesting RBC variant in our Online Appendices C.1 and D. Parameter values in this model are given in Table 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 the true process underlying the observables, $y_t = (n_t,\ \Delta \ln X_t)'$, 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(2,1) $\sim$ SCM(1,1), 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.

4.2 DGP 3: RBC model with news shocks

We also consider a DGP variant in terms of a RBC model with news shock.[5] Since this version of the model is quite well known, we relegate its description to our Online Appendices C.2 and E. The third column of Table 1 presents the values of the structural parameters. Most

[5]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).

of them are chosen following Yang (2005).

Taking $y_t = (n_t,\ \Delta \ln X_t)'$ again as observable, 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(2,1)∼SCM(1,1). 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.[6] 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.

4.3 DGP 4: A monetary 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).[7] We relegate the description of this model to our Online Appendix C.3 and F.

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). Some key parameter values are shown in the last column of Table 1. Other calibrated parameters are discussed in our Online Appendix F. The minimal VARMA representation of the log-linearized solution is a VARMA(2,1) process with the structure SCM(2,1) ∼ SCM(1,1)

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

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(1,1) ∼ SCM(1,0) 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.

[6]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 obtained using Blaschke matrices as introduced by Lippi and Reichlin (1994).

[7]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.

5 Conclusion

In empirical work attempting to identify unobserved policy, demand-, and/or supply-side shocks from observed macroeconomic time-series data, a practitioner would often condition their insights on VAR or VARMA representations of the data. She would also impose minimal long-run and sign-restriction insights from some implicit theory—usually some form of a DSGE model—on the impulse dynamics of their statistical model-windows to the data.

We have shown that even if econometricians were endowed with the correct long-run and sign-restriction insights, using a VAR still ends up with a misleading conclusion about the true DGPs' dynamics; and this problem is ever-present even if the VAR econometrician had the luxury of having observed an arbitrarily long sample of data. Further, across many and varied forms of structural DGPs, in small samples the VARMA econometrician will almost always fail to identify the correct VARMA structure coming from the DGP. However, this problem goes away asymptotically—i.e., if the econometrician had the luxury of an arbitrarily long sample of time-series data.

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, whilst being prepared to believe that the data would have come from some unspecified DGP that shares some behavior with well-accepted theories of business cycles. The earlier conclusions in the literature (using the prototype RBC-as-DGP example) and our more extensive findings here, suggest that using VARs or VARMAs as identification devices to uncover policy and market-relevant shifts in the data is quite a problematic endeavor—the researcher may end up with quite misleading conclusions.

References

Aruoba, S. B. (2011), "Money, Search, And Business Cycles", *International Economic Review* **52**(3), 935–959. Cited on page(s): [13], [SA.3 — E]

Aruoba, S. B., Waller, C. J. and Wright, R. (2008), "Money and Capital", *Journal of Monetary Economics* **58**(2), 98–116. Cited on page(s): [13], [SA.3 — E], [SA.9 — E]

Athanasopoulos, G. and Vahid, F. (2008), "A Complete VARMA Modelling Methodology Based on Scalar Components", *Journal of Time Series Analysis* **29**(3), 533–554. Cited on page(s): [5]

Bagliano, F. C. and Favero, C. A. (1998), "Measuring Monetary Policy with VAR Models: An Evaluation", *European Economic Review* **42**(6), 1069–1112. Cited on page(s): [1]

Blanchard, O. J. and Quah, D. (1989), "The Dynamic Effects of Aggregate Demand and Supply Disturbances", *American Economic Review* **79**(4), 655–73. Cited on page(s): [3]

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. Cited on page(s): [13], [SA.9 — E]

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. Cited on page(s): [1], [2]

Christiano, L. J., Eichenbaum, M. and Vigfusson, R. (2006), "Assessing Structural VARs", NBER Working Papers 12353, National Bureau of Economic Research, Inc. Cited on page(s): [1], [2]

Cogley, T. and Nason, J. M. (1993), "Impulse Dynamics and Propagation Mechanisms in a Real Business Cycle Model", *Economics Letters* **43**(1), 77–81. Cited on page(s): [2], [9] Cooley, T. and Dwyer, M. (1998), "Business Cycle Analysis without Much Theory. A Look at Structural VARs", *Journal of Econometrics* **83**, 57–88. Cited on page(s): [1]

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. Cited on page(s): [12], [SA.2 — D]

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. Cited on page(s): [1], [2], [3]

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. Cited on page(s): [1]

Gomis-Porqueras, P., Kam, T. and Lee, J. (2013), "Money, Capital, and Exchange Rate Fluctuations", *International Economic Review* **54**(1), 329–353. Cited on page(s): [13]

Hansen, G. D. (1985), "Indivisible Labor and the Business Cycle", *Journal of Monetary Economics* **16**(3), 309–327. Cited on page(s): [2]

Hansen, L. P. and Sargent, T. J. (1991), "Two Difficulties in Interpreting Vector Autoregressions", *Rational Expectations Econometrics, Westview Press, Boulder and London*, 77–119. Cited on page(s): [12], [SA.2 — D]

Heathcote, J. and Perri, F. (2002), "Financial Autarky and International Business Cycles", *Journal of Monetary Economics* **49**(3), 601–627. Cited on page(s): [13], [SA.9 — E]

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. Cited on page(s): [13], [SA.3 — E], [SA.6 — E], [SA.9 — E]

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. Cited on page(s): [7]

Kascha, C. and Mertens, K. (2009), "Business Cycle Analysis and VARMA Models", *Journal of Economic Dynamics and Control* **33**(2), 267–282. Cited on page(s): [2], [3], [8], [11]

Kilian, L. (2011), "Structural Vector Autoregressions", CEPR Discussion Papers 8515, C.E.P.R. Discussion Papers. Cited on page(s): [8]

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. Cited on page(s): [1]

Lagos, R. and Wright, R. (2005), "A Unified Framework for Monetary Theory and Policy Analysis", *Journal of Political Economy* **113**(3), 463–484. Cited on page(s): [13]

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. Cited on page(s): [12], [SA.2 — D]

Lippi, M. and Reichlin, L. (1994), "VAR Analysis, Nonfundamental Representations, Blaschke Matrices", *Journal of Econometrics* **63**(1), 307–325. Cited on page(s): [13]

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. Cited on page(s): [1]

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. Forthcoming in the *Journal of Business and Economics Statistics*. Cited on page(s): [7]

Ravenna, F. (2007), "Vector Autoregressions and Reduced Form Representations of DSGE Models", *Journal of Monetary Economics* **54**(7), 2048–2064. Cited on page(s): [1], [2], [3] 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. Cited on page(s): [13], [SA.9 — E]

Sims, C. A. (1988), "Identifying Policy Effects", *Empirical Macroeconomics for Interdependent Economics, the Brookings Institution, Washington D.C.* pp. 308–321. Cited on page(s): [12], [SA.2 — D]

Valentinyi, A. and Herrendorf, B. (2008), "Measuring Factor Income Shares at the Sectoral Level", *Review of Economic Dynamics* **11**(4), 820–835. Cited on page(s): [10]

Yang, S. S.-C. (2005), "Quantifying Tax Effects under Policy Foresight", *Journal of Monetary Economics* **52**(8), 1557–1568. Cited on page(s): [13], [SA.2 — D]

NOT FOR PUBLICATION
ONLINE APPENDIX

Available publicly from: https://github.com/phantomachine/_varmita

A DGP 1: The Hansen RBC Model

Here we further describe the "correct" long-run restriction information that we have available to the statisticians within our Monte Carlo experiments in Section 2.1 of the paper.

In the RBC model only the technology shock ε_{z_t} has a long-run effect on total production. Denote the VMA representation of y_t as $y_t = \Upsilon(L)u_t = \Upsilon(L)A_0\varepsilon_t$, where the matrix A_0 transform the structural shocks ε_t into u_t. The impulse response functions are derived from the coefficients in the matrix polynomial $\Upsilon(L)A_0$. The response of $\Delta \ln X_t$ to a labor supply shock ε_{d_t} is associated with the term $[\Upsilon(L)A_0]_{22}$. Thus, the long-run identification constraint implies that $[\Upsilon(L)A_0]_{22} = 0$. The other restrictions needed for identification come from the relationship between the variance-covariance matrices of u_t and ε_t that $\Sigma_u = A_0\Sigma_\varepsilon A_0'$, where $\Sigma_\varepsilon = \text{diag}\{\sigma_z^2, \sigma_d^2\}$. The last step is to determine the sign of the impulse responses 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. 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.

B Examining the AR and MA Roots

This section details the distributions of the DGP parameters used in the simulation exercises in Section 3 in the paper.

C DGP 2: RBC with Habits and Capital Adjustment Cost

The agent's criterion function is now: $\mathbb{E}_0\left\{\sum_{t=1}^{\infty}\beta^t[\ln(C_t - \theta_c C_{t-1}) + D_t(L_t - \theta_l L_{t-1})]\right\}$, where $\theta_c, \theta_l > 0$. The agent's utility depends on the current consumption relative to a fraction of the past consumption and leisure.[8]

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 a quadratic loss function $S_t = \frac{\varsigma}{2}\left(\frac{I_t}{I_{t-1}} - 1\right)^2$, $\varsigma > 0$ as

[8] 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. 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.

the adjustment cost. The total capital stock in the economy accumulates according to $K_t = (1-\delta_t) K_{t-1} + (1-S_t) I_t$, where the capital depreciation rate δ_t is the time-varying as a quadratic function of the capital utilization rate U_t, $\delta_t = \delta_0 + \delta_1 (U_t - 1) + \frac{\delta_2}{2} (U_t - 1)^2$, $\delta_0, \delta_1, \delta_2 > 0$. Here U_t will also affect the production function via $X_t = (U_t K_{t-1})^\alpha (Z_t N_t)^{1-\alpha}$.

D DGP 3: The RBC Model with News Shock and Fiscal Policy

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 faces the same maximization problem given in our first RBC model, but now subject to the per-period budget constraint $C_t + K_t - (1-\delta) K_{t-1} + T_t = (1-\tau_t^L) W_t N_t + (1-\tau_t^K) R_t K_{t-1}$. 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 government's per-period budget constraint requires $G_t = T_t + \tau_t^L W_t N_t + \tau_t^K R_t K_{t-1}$. We assume that fiscal policy is exogenously specified. In particular, the tax rate on labor obeys

$$\ln \tau_t^L = \rho_L \ln \tau_{t-1}^L + \mu_L \ln(X_t/Z_t) + \varepsilon_{L_{t-1}} + r_{K,L}\, \varepsilon_{K_{t-1}}. \tag{4a}$$

The tax rate on capital follows

$$\ln \tau_t^K = \rho_K \ln \tau_{t-1}^K + \mu_K \ln(X_t/Z_t) + \varepsilon_{K_{t-1}} + r_{K,L}\, \varepsilon_{L_{t-1}}, \tag{4b}$$

and, government spending in efficiency units follows

$$\ln(G_t/Z_t) = \rho_G \ln(G_{t-1}/Z_{t-1}) + \varepsilon_{G_t}. \tag{4c}$$

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 (4a) and (4b), tax shocks occurring at period-t will change the tax rates at period-$(t+1)$. Hence, the agents have foresight of, or "news" about, tax policy one period ahead.

E DGP 4: Monetary Search Model

Since this monetary model is not so standard in terms of it being used as a data-generating process for experiments in the VAR versus VARMA literature, we describe it here. The model is a version of Aruoba (2011) or Aruoba et al. (2008), which was shown to match U.S. monetary business cycle facts quite well.

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.[9]

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

$$\ln(\psi_t) = \rho_m \ln(\psi_{t-1}) + \sigma_m \varepsilon_{\psi_t}, \quad 0 < \rho_m < 1, \text{ and } \quad \varepsilon_{\psi_t} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1). \tag{5}$$

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

$$\ln(\tilde{Z}_t) = \rho_z \ln(\tilde{Z}_{t-1}) + \sigma_z \varepsilon_{z_t}, \quad 0 < \rho_z \leq 1, \text{ and } \quad \varepsilon_{z_t} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1). \tag{6}$$

[9] 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*).

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 Appendix E. In terms of the observable variables—percentage deviations of the aggregate employment and the growth of real output—we have $y_t := (n_{tot,t},\ \Delta \ln X_{tot,t})'$.

E.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]

E.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.

E.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.

[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 $c(q^s/Z, k/Z) \equiv N_{DM}$.

E.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 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

$$V(m_{t-1}, k_{t-1}, \hat{\mathbf{s}}_t) = \sigma V^b(m_{t-1}, k_{t-1}, \hat{\mathbf{s}}_t) + \sigma V^s(m_{t-1}, k_{t-1}, \hat{\mathbf{s}}_t) + (1-2\sigma) W(m_{t-1}, k_{t-1}, \mathbf{s}_t), \tag{7}$$

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 (7) 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 the *ex post* buyer's problem:

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

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:

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

E.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:

$$W(m_{t-1}, k_{t-1}, \mathbf{s}_t) = \max_{X_t, N_t, m_t, k_t} \left\{ U(X_t) - \phi N_t + \beta \mathbb{E}_\lambda \left[V(m_t, k_t, \hat{\mathbf{s}}_{t+1}) \middle| (Z_t, \psi_t) \right] \right\}, \tag{8}$$

subject to

$$\mathbf{s}_{t+1} = \mathcal{G}(\mathbf{s}_t, \mathbf{v}_{t+1}), \qquad \mathbf{v}_t \overset{\text{i.i.d.}}{\sim} \varphi, \tag{9}$$

$$X_t + k_t - (1-\delta)k_{t-1} = \frac{m_{t-1} - m_t}{P_t} + w_t N_t + r_t k_{t-1} + TR_t, \tag{10}$$

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 (9) 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 (9) 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 (10).

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.

E.2.4 Exogenous Processes

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

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

where $0 < \rho_M < 1$.

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

$$\ln(\tilde{Z}_t) = \rho_z \ln(\tilde{Z}_{t-1}) + \sigma_z \varepsilon_{z_t}, \qquad 0 < \rho_z \leq 1 \;\text{ and }\; \varepsilon_{z_t} \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1).$$

E.2.5 Market Clearing

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

$$F(K_{t-1}, Z_t N_t) = X_t + K_t - (1 - \delta) K_{t-1}. \tag{11}$$

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

$$TR_t = \frac{M_t - M_{t-1}}{P_t}. \tag{12}$$

These, together with the agent's CM budget constraint (equation (10)) in equilibrium, imply that the labor market in the CM must clear as well, *i.e.* $N_t^d = N(\mathbf{s}_t) := \int_{\mathbb{R}_+} N(m_{t-1}, k_{t-1}, \mathbf{s}_t) d\mu_t$.

E.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 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 (8)-(10) 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:*

$$U'_X(\tilde{X}_t) = \beta \mathbb{E}_{\lambda} \left\{ \frac{U'_X(\tilde{X}_{t+1})}{\tilde{Z}_{t+1}} \left[F'_K(\tilde{K}_t/\tilde{Z}_{t+1}, N_{t+1}) - \delta \right] - \sigma \frac{c'_K(\tilde{q}_{t+1}, \tilde{K}_t/\tilde{Z}_{t+1})}{\tilde{Z}_{t+1}} \middle| (\tilde{Z}_t, \psi_t) \right\}, \tag{13}$$

[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.

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

2. *Inter-temporal optimal money holdings:*

$$U'_X(\tilde{X}_t) = \beta \mathbb{E}_\lambda \Bigg\{ U'_X(\tilde{X}_{t+1}) \frac{\tilde{P}_t}{\psi_t \tilde{P}_{t+1}} \times \left[(1-\sigma) + \sigma \frac{u'_q(\tilde{q}_{t+1})}{c'_q(\tilde{q}_{t+1}, \tilde{K}_t/\tilde{Z}_{t+1})} \right] \Bigg| (\tilde{Z}_t, \psi_t) \Bigg\}, \tag{14}$$

3. *Labor market clearing:*

$$U'_X(\tilde{X}_t) = \frac{\phi}{F'_N(\tilde{K}_{t-1}/\tilde{Z}_t, N_t)}, \tag{15}$$

4. *DM price-taking solution:*

$$\frac{U'_X(\tilde{X}_t)}{\tilde{P}_t} \psi_t = c'_q(\tilde{q}_t, \tilde{K}_{t-1}/\tilde{Z}_t) \tilde{q}_t, \tag{16}$$

5. *Resource constraint:*

$$\tilde{X}_t + \tilde{K}_t + \tilde{G}_t = F(\tilde{K}_{t-1}/\tilde{Z}_t, N_t) + (1-\delta)\tilde{K}_{t-1}/\tilde{Z}_t. \tag{17}$$

E.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

$$\tilde{X}_{tot,t} = \frac{X^{nom}_{DM,t} + \tilde{P}_t \tilde{X}_{CM,t}}{\tilde{P}_{X,t}}. \tag{18}$$

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

$$N_{tot,t} = \sigma c(\tilde{q}_t, \tilde{K}_{t-1}/\tilde{Z}_t) + N_t. \tag{19}$$

Denote the percentage deviations of $N_{tot,t}$ and $\tilde{X}_{tot,t}$ in equations (18) and (19) 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 $y_t := (n_{tot,t}, \Delta \ln X_{tot,t})'$.

E.5 Parametrization and 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. 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.

References

Aliprantis, C. D., Camera, G. and Puzzello, D. (2007*a*), "Anonymous Markets and Monetary Trading", *Journal of Monetary Economics* **54**(7), 1905–1928. Cited on page(s): [SA.3 — E]

Aliprantis, C. D., Camera, G. and Puzzello, D. (2007*b*), "Contagion Equilibria in a Monetary Model", *Econometrica* **75**(1), 277–282. Cited on page(s): [SA.3 — E]

Aruoba, S. B. (2011), "Money, Search, And Business Cycles", *International Economic Review* **52**(3), 935–959. Cited on page(s): [13], [SA.3 — E]

Aruoba, S. B., Waller, C. J. and Wright, R. (2008), "Money and Capital", *Journal of Monetary Economics* **58**(2), 98–116. Cited on page(s): [13], [SA.3 — E], [SA.9 — E]

Athanasopoulos, G. and Vahid, F. (2008), "A Complete VARMA Modelling Methodology Based on Scalar Components", *Journal of Time Series Analysis* **29**(3), 533–554. Cited on page(s): [5]

Bagliano, F. C. and Favero, C. A. (1998), "Measuring Monetary Policy with VAR Models: An Evaluation", *European Economic Review* **42**(6), 1069–1112. Cited on page(s): [1]

Blanchard, O. J. and Quah, D. (1989), "The Dynamic Effects of Aggregate Demand and Supply Disturbances", *American Economic Review* **79**(4), 655–73. Cited on page(s): [3]

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. Cited on page(s): [SA.1 — C]

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. Cited on page(s): [SA.1 — C]

Carroll, C. D., Overland, J. and Weil, D. N. (2000), "Saving and Growth with Habit Formation", *The American Economic Review* **90**(3), 341–355. Cited on page(s): [SA.1 — C]

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. Cited on page(s): [13], [SA.9 — E]

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. Cited on page(s): [1], [2]

Christiano, L. J., Eichenbaum, M. and Vigfusson, R. (2006), "Assessing Structural VARs", NBER Working Papers 12353, National Bureau of Economic Research, Inc. Cited on page(s): [1], [2]

Cogley, T. and Nason, J. M. (1993), "Impulse Dynamics and Propagation Mechanisms in a Real Business Cycle Model", *Economics Letters* **43**(1), 77–81. Cited on page(s): [2], [9]

Cooley, T. and Dwyer, M. (1998), "Business Cycle Analysis without Much Theory. A Look at Structural VARs", *Journal of Econometrics* **83**, 57–88. Cited on page(s): [1]

Duffie, D., Geanakoplos, J., Mas-Colell, A. and McLennan, A. (1994), "Stationary Markov Equilibria", *Econometrica* **62**(4), 745–781. Cited on page(s): [SA.4 — E]

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. Cited on page(s): [12], [SA.2 — D]

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. Cited on page(s): [1], [2], [3]

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. Cited on page(s): [1]

Fuhrer, J. C. (2000), "Habit Formation in Consumption and Its Implications for Monetary Policy Models", *The American Economic Review* **90**(3), 367–390. Cited on page(s): [SA.1 — C]

Gomis-Porqueras, P., Kam, T. and Lee, J. (2013), "Money, Capital, and Exchange Rate Fluctuations", *International Economic Review* **54**(1), 329–353. Cited on page(s): [13]

Hansen, G. D. (1985), "Indivisible Labor and the Business Cycle", *Journal of Monetary Economics* **16**(3), 309–327. Cited on page(s): [2]

Hansen, L. P. and Sargent, T. J. (1991), "Two Difficulties in Interpreting Vector Autoregressions", *Rational Expectations Econometrics, Westview Press, Boulder and London*, 77–119. Cited on page(s): [12], [SA.2 — D]

Heathcote, J. and Perri, F. (2002), "Financial Autarky and International Business Cycles", *Journal of Monetary Economics* **49**(3), 601–627. Cited on page(s): [13], [SA.9 — E]

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. Cited on page(s): [13], [SA.3 — E], [SA.6 — E], [SA.9 — E]

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. Cited on page(s): [7]

Kascha, C. and Mertens, K. (2009), "Business Cycle Analysis and VARMA Models", *Journal of Economic Dynamics and Control* **33**(2), 267–282. Cited on page(s): [2], [3], [8], [11]

Kilian, L. (2011), "Structural Vector Autoregressions", CEPR Discussion Papers 8515, C.E.P.R. Discussion Papers. Cited on page(s): [8]

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. Cited on page(s): [1]

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

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. Cited on page(s): [SA.7 — E]

Lagos, R. and Wright, R. (2005), "A Unified Framework for Monetary Theory and Policy Analysis", *Journal of Political Economy* **113**(3), 463–484. Cited on page(s): [13]

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. Cited on page(s): [12], [SA.2 — D]

Lippi, M. and Reichlin, L. (1994), "VAR Analysis, Nonfundamental Representations, Blaschke Matrices", *Journal of Econometrics* **63**(1), 307–325. Cited on page(s): [13]

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. Cited on page(s): [1]

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. Forthcoming in the *Journal of Business and Economics Statistics*. Cited on page(s): [7]

Ravenna, F. (2007), "Vector Autoregressions and Reduced Form Representations of DSGE Models", *Journal of Monetary Economics* **54**(7), 2048–2064. Cited on page(s): [1], [2], [3]

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. Cited on page(s): [13], [SA.9 — E]

Sims, C. A. (1988), "Identifying Policy Effects", *Empirical Macroeconomics for Interdependent Economics, the Brookings Institution, Washington D.C.* pp. 308–321. Cited on page(s): [12], [SA.2 — D]

Valentinyi, A. and Herrendorf, B. (2008), "Measuring Factor Income Shares at the Sectoral Level", *Review of Economic Dynamics* **11**(4), 820–835. Cited on page(s): [10]

Yang, S. S.-C. (2005), "Quantifying Tax Effects under Policy Foresight", *Journal of Monetary Economics* **52**(8), 1557–1568. Cited on page(s): [13], [SA.2 — D]