A, B, C'S (AND D)'S FOR UNDERSTANDING VARS

JESÚS FERNÁNDEZ-VILLAVERDE
UNIVERSITY OF PENNSYLVANIA

JUAN F. RUBIO-RAMÍREZ
FEDERAL RESERVE BANK OF ATLANTA

THOMAS J. SARGENT
NEW YORK UNIVERSITY AND HOOVER INSTITUTION

Preliminary and Incomplete

"(Likelihood Principle) The information brought by an observation x about [a parameter] θ is entirely contained in the likelihood function." *The Bayesian Choice*, by Christian P. Robert, p. 15.[1]

"... with a specific parameterization of preferences the theory would place many restrictions on the behavior of endogenous variables. But these predictions do not take the form of locating blocks of zeros in a VAR description of these variables." *Money and Interest in a Cash-in-Advance Economy*, Robert E. Lucas, Jr., and Nancy L. Stokey, p. 512.[2]

I. INTRODUCTION

These epigraphs frame our topic. Unrestricted Vector Autoregressions (VARs) have become a crucial component of the economists' toolbox. They are used, among other purposes, to answer questions like: What is the effect of a technology shock on hours worked? How does output respond to monetary perturbations? What happens after a fiscal shock? Knowing the answers to these questions will guide us in building better theoretical models.

To provide answers to these questions the researcher needs to recover economic shocks from the unrestricted VAR innovations. This is done through identification restrictions on the unrestricted VAR. This practice is attractive because, at least up to a linear approximation, many dynamic equilibrium models have a representation as a restricted VAR. This

[1] See Robert, (2001).

[2] See Lucas and Stokey (1987).

restricted VAR is a recursive expression of the model's likelihood function, which according to the likelihood principle contains all that the data have to say about the model's parameters. Consequently, if our theoretical model has a restricted VAR representation and we correctly place our identification restrictions in a unrestricted VAR, we can learn everything is to be known about the theoretical model's behavior without the need to fully specify and estimate a dynamic equilibrium model.

However, the promise of unrestricted VARs of providing a coherent framework for structural inference is not without difficulties. As we will describe below in more detail, the travel between a unrestricted VAR and a dynamic equilibrium model is full of dangers. When can we recover economic shocks from unrestricted VAR innovations? Which are the restrictions that a theoretical model's stationary distribution places on a unrestricted VAR? Can the impulse response function of the identified VAR match the impulse response function of a theoretical model? What are the consequences of estimating the unrestricted VAR with a finite set of data?.

This paper addresses these questions in six steps. First, it collects a set of convenient formulae that summarize how we can move between a unrestricted VAR and a dynamic equilibrium model.

Second, we make precise how, in general, unrestricted VAR innovations amalgamize current and past realizations of the economic shocks hitting preferences, technologies, and information sets. The formulae impose a taxonomy of potential challenges in interpreting unrestricted VAR innovations and the impulse response function of observables to them in terms of the economic shocks and their impulse response function. This taxonomy structures the rest of the paper.

Third, we study when the mapping between unrestricted VAR innovations and the history of economic shocks is invertible, i.e., when we can write the economic shocks as a function of linear combinations of present and past unrestricted VAR innovations. For some theories, the number of economic shocks differ from the number of observables and therefore the number of shocks in the unrestricted VAR. Even for the lucky situation in which the economic shocks equal the number of variables, it is possible that the history of economic shocks spans a bigger space than the history of the observables. We show a necessary and sufficient condition to test for invertibility for the 'square case', i.e., when we have the same number of economic shocks and observables.

Fourth, the paper analyzes the implications of invertibility (or lack of) for the impulse response function of an identified VAR. Even when the theory is such that there are equal numbers of shocks and no invertibility problem, there remains the challenge of partitioning the contemporaneous covariation among unrestricted VAR innovations in a way that captures the contemporaneous covariance of economic shocks. In practice, researchers accomplish this by imposing long run theoretical restrictions from theory on the VAR or by placing zeroes. For the 'square case', we show that when the model is invertibility there is an identification scheme such that the impulse response function of a identified VAR matches those of the economic theory.

Fifth, we also explore the implications of employing small samples to estimate the unrestricted VAR. Even when all of the previous problems can be resolved, because the theoretical model implies in general an infinite order VAR (technically, it is a finite order VARMA system), one must either fit moving average terms or use an information-theoretic (e.g., a Bayesian information criterion) to select a suitable lag length for the unrestricted VAR. Because the unrestricted VAR is profligately parameterized relative to the economic theory, specifying a good lag length for small samples means that the researcher must call for help from different tools.

Finally, we illustrate our results with some examples, including two interesting application, one by Fisher (**?**) and one by Erceg, Henderson and Levin (**?**).

The daunting hierarchy of problems described above tempts the researcher to decline the challenge to match his theory to a unrestricted VAR. But because the VAR is simply an expression of the likelihood principle, a researcher who claims to take his theory seriously, in the sense of analyzing the data as if his theory is true, simply cannot turn his back on the implications of his theory for a unrestricted VAR.[3]

The paper is organized as follows. In section 2 we ...

II. THE QUESTION

The main objective of the paper is to set up necessary and sufficient conditions under which the economic shocks can be recovered from the linear combinations of history of innovations generated by a reduced form VAR. We first define economic shocks, w_t's, and innovations, a_t's. Economic shocks are defined in a economic model equilibrium, while innovations are defined in an unrestricted VAR. Second, we show how unrestricted VAR innovations are function of linear combinations of the history economic shocks. Finally, we ask whether economic shocks are function of linear combinations of the history of unrestricted VAR innovations? The conditions under which the answer is "yes" will be given in section III.

II.1. **Economic Model Equilibrium.** We start with the equilibrium laws of motion of an economic model or a linear approximation to it. These laws of motion have a representation

[3] A theorist who wants to analyze the data as if his theory *is not* true would not turn his back on the likelihood function either. However a complete study of the robustness problems raised by the explicit use of a false theory will divert us too much from the main focus of our paper.

in the state-space form

$$x_{t+1} = Ax_t + Bw_t \tag{1}$$

$$y_t = Cx_t + Dw_t \tag{2}$$

where w_t is a Gaussian vector satisfying $Ew_t = 0$, $Ew_t w_t' = I$, $Ew_t w_{t-j} = 0$ for $j \neq 0$. Here x_t is an $n \times 1$ vector of possibly unobserved state variables, y_t is a $k \times 1$ vector of variables observed by an economist or econometrician, and w_t is an $m \times 1$ vector of economic shocks impinging on the states and observables. Typically, the observation vector includes some prices, quantities, and capital stocks. With m shocks in the economic model, n states, and k observables, A is $n \times n$, B is $n \times m$, C is $k \times n$, and D is $k \times m$. In general, $k \neq m$, although we shall soon devote some special attention to an interesting 'square case' in which the number of economic shocks equals the number of observables ($k = m$.)

There are two main ways of obtaining equilibrium representations of the form (1)-(2). The first is to compute a linear or loglinear approximation of a nonlinear model about a non-stochastic steady state, as exposited for example, in Christiano (XXXX), Uhlig (XXXX), King and Watson (XXXX), or the `dynare` manual.[4] It is straightforward to collect the linear or log linear approximations to the equilibrium decision rules and to arrange them into the state-space form (1)-(2). We provide an extended examples in section EXAMPLES. A second way that (1)-(2) emerges is directly as a representation of a member of a class of dynamic stochastic general equilibrium models with linear transition laws and quadratic preferences. Ryoo and Rosen (XXXX), Rosen and Topel (XXXX), Rosen, Murphy, and Scheinkman (XXXX), and Hansen and Sargent (2005) provide many examples.

II.2. **Unrestricted VAR.** On the other hand, a mainstay of empirical macroeconomics is the reduced form vector autoregression on observables y_t

$$y_t = \alpha + \sum_{j=1}^{\infty} A_j y_{t-j} + a_t \tag{3}$$

where $a_t = y_t - E[y_t|y^{t-1}]$, $\sum_{j=1}^{\infty} \text{trace}(A_j A_j') < +\infty$, $Ea_t a_t' = \Omega$, and the α and the A_js satisfy the least squares orthogonality conditions

$$Ea_t = 0$$

$$Ea_t y_{t-j}' = 0, \quad j \geq 1$$

We call a_t innovations. Let ε_t be an stochastic process that satisfies $E\varepsilon_t = 0$, $E\varepsilon_t\varepsilon_t' = I$, and $E\varepsilon_t\varepsilon_{t-j}' = 0$ for $j \neq 0$, then exists an unique G such that $GQ\varepsilon_t = a_t$ for any orthonormal matrix Q. To economize on notation, from now on, we shall assume that $Q = I$.

[4]Dynare is a suite of Matlab programs that computes linear approximations of a big class of dynamic stochastic general equilibrium models.

II.3. **The Innovations Representation.** To build a map from linear combinations of the history of economic shocks to innovations, we define what is called an innovations representation: Associated with any state space system (A,B,C,D) of the form (1)-(5) is the following *innovations representation*:

$$\hat{x}_{t+1} = A\hat{x}_t + KG\varepsilon_t \tag{4}$$

$$y_t = C\hat{x}_t + G\varepsilon_t, \tag{5}$$

where $\hat{x}_t = E[x_t|y^{t-1}]$, $G\varepsilon_t \equiv a_t = y_t - E[y_t|y^{t-1}]$, and K is the Kalman gain from the steady state Kalman filter equations:

$$\begin{aligned} \Sigma &= A\Sigma A' + BB' - (A\Sigma C' + BD') \\ &\quad (C\Sigma C' + DD')^{-1}(A\Sigma C' + BD')' \qquad (6) \\ K &= (A\Sigma C' + BD')(C\Sigma C' + DD')^{-1} \qquad (7) \end{aligned}$$

where $\Sigma = E(x_t - \hat{x}_t)(x_t - \hat{x}_t)'$. The covariance matrix of the innovations $a_t = G\varepsilon_t$ equals

$$Ea_t a_t' = GG' = C\Sigma C' + DD'. \tag{8}$$

With m shocks in the economic model, n states, and k observables, K is $n \times k$ and G is $k \times k$. The vector processes a_t and ε_t are each of dimension $k \times 1$, as is the y_t process, and the matrix G is $k \times k$. Let $H(z^t)$ be the Hilbert space consisting of all square summable linear combinations of the one-sided infinite history of random vectors z^t. The Kalman filter takes the history y^t and by using a Gram-Schmidt procedure constructs a history a^t with orthogonal increments that spans the same linear space, i.e., is such that $H(y^t) = H(a^t)$ and for which $Ea_t a_s' = 0$ for $t \neq s$.

The innovations representation (4)-(5) is also an equilibrium representation for the y process, one that resembles the original equilibrium representation (1)-(2). It differs from it in that (a) the $n \times k$ matrix KG replaces the $n \times m$ matrix B; (b) the $k \times k$ matrix G replaces the $k \times m$ matrix D; and (c) the $k \times 1$ process ε_t replaces the $m \times 1$ process w_t.

We can combine and rearrange the two representations to obtain the following system that describes the mapping from the economic shocks w_t to the innovations $G\varepsilon_t$:

$$\begin{bmatrix} x_{t+1} \\ \hat{x}_{t+1} \end{bmatrix} = \begin{bmatrix} A & 0 \\ KC & A-KC \end{bmatrix} \begin{bmatrix} x_t \\ \hat{x}_t \end{bmatrix} + \begin{bmatrix} B \\ KD \end{bmatrix} w_t \tag{9}$$

$$G\varepsilon_t = \begin{bmatrix} C & -C \end{bmatrix} \begin{bmatrix} x_t \\ \hat{x}_t \end{bmatrix} + Dw_t. \tag{10}$$

Define

$$A^* \equiv \begin{bmatrix} A & 0 \\ KC & A-KC \end{bmatrix}$$

if the eigenvalues of $A-KC$ are less than unity in modulus, we can write (9)-(10) as

$$G\varepsilon_t = \left\{ D + \begin{bmatrix} C & -C \end{bmatrix} \left[I - A^* L \right]^{-1} \begin{bmatrix} B \\ KD \end{bmatrix} L \right\} w_t. \tag{11}$$

Equation (11) can be writen as:

$$G\varepsilon_t = \sum_{j=0}^{\infty} h_j w_{t-j} \tag{12}$$

with

$$h_0 = I$$
$$h_j = \begin{bmatrix} C & -C \end{bmatrix} (A^*)^{j-1} \begin{bmatrix} B \\ KD \end{bmatrix}, \quad j \geq 1.$$

Equation (12) verifies that by construction $H(a^t) \subset H(w^t)$, i.e., equation (11) defines a mapping from linear combinations of the history of economy shochs, w^t, to innovations, $G\varepsilon_t$. We want to know whether $H(a^t) = H(w^t)$. If it is, we say that the mapping (11) is invertible.

In section III, we give a neat condition for checking invertibility in the 'square' case that there are as many observables as economic shocks.

II.4. **The Conditional Likelihood.** Before showing the condition to check invertibility, let us spend some time on the implications of the innovations representation (4)–(5) for estimation.

The innovations representation (4)–(5) can be rearranged to assume the form of a Wold moving average representation

$$y_t = [G + C(I - AL)^{-1} KGL] \varepsilon_t \tag{13}$$

By applying a partitioned inverse formula to invert the operator $[G + C(I - AL)^{-1} KGL]$ in (13), Hansen and Sargent (2005) show that when the eigenvalues of $A - KC$ are less than unity in modulus, y_t has an autoregressive representation given by

$$y_t = C[I - (A - KC)L]^{-1} K y_{t-1} + G\varepsilon_t \tag{14}$$

which is of the form (3) with

$$A_j = C(A - KC)^{j-1} K, \quad j \geq 0. \tag{15}$$

Formulas (14) and (15) make explicit the restrictions that the dynamic model defined by (A,B,C,D) imposes on a VAR. These restrictions depend on A,B,C, and D, but also on K, the Kalman gain from the steady state Kalman filter equations (6)-(7). In section IV, we show that, when the model is invertibility there exists another version of (15) that only depends on A,B,C, and D, therefore, in that case the restrictions that the dynamic model described imposes on a VAR only depend on A,B,C, and D.

The innovations representation, (4)-(5) , and the restricted VAR, (14), implied by (A,B,C,D) exhausts the implications of the economic model for all first and second moments of the process $\{y_t\}$. When w_t is Gaussian, it exhausts the implications of the model for the joint distribution of any sequence of y's. This claim follows from the fact that all of the information that a time series of observations $\{y_t\}_{t=1}^T$ contains about the economic parameters underlying (A,B,C,D) is contained in the model's likelihood function. The innovations representation (4)-(5) or the theoretical infinite order VAR (14) contains all of the information needed to construct a Gaussian likelihood function conditional on an initial infinite history of observations.[5]

Denote the likelihood function of a sample of data $\{y_t\}_{t=1}^T$ conditional on the infinite history y^0 by $f(y_T, y_{T-1}, \ldots, y_1|y^0)$. Factor this likelihood as

$$L = f(y_T, y_{T-1}, \ldots, y_1|y^0) = f_T(y_T|y^{T-1}) f_{T-1}(y_{T-1}|y^{T-2}) \cdots f_1(y_1|y^0). \tag{16}$$

Under the assumption that w_t is a Gaussian process, the conditional density $f_t(y_t|y^{t-1})$ is $\mathcal{N}(C\hat{x}_t, GG')$. Recalling that $a_t = y_t - C\hat{x}_t$ from (5), it follows that $\log f(y_T, y_{T-1}, \ldots, y_1|y^0)$, the log of the conditional likelihood (16), equals

$$\log L = -.5 \sum_{t=1}^{T} \left\{ k \log 2\pi + \ln|GG'| + a_t'(GG')^{-1} a_t \right\}. \tag{17}$$

III. THE ANSWER IN THE "SQUARE CASE"

In section II.3, we show how the innovation representation defines a mapping from linear combinations of the history of economy shochs, w^t, to innovations, $G\varepsilon_t$. In this section we investigate when that mapping is invertible for the 'square' case, i.e., when there are as many observables as economic shocks. This is the case that is 'least likely' to have an invertibility problem. In such an framework, we can get a nice necessary and sufficient condition for invertibility in terms of the fundamentals of the problem.

We shall make the following assumptions:

ASSUMPTION 1. The state space system (1), (2) is stable: all of the eigenvalues of A are less than one in modulus, except possibly one associated with a constant.

ASSUMPTION 2: D is square and invertible.

Assumption 2 often applies to systems with equal numbers of economic shocks and observables (i.e., variables in the pertinent VAR). Under Assumption 2, (11) can be represented

[5] See Hansen and Sargent (2005), chapter 9) for how the Kalman filter can also be used to construct an unconditional likelihood function.

as:

$$G\varepsilon_t = \left\{ I + \begin{bmatrix} C & -C \end{bmatrix} \begin{bmatrix} I - A^* L \end{bmatrix}^{-1} \begin{bmatrix} BD^{-1} \\ K \end{bmatrix} L \right\} D w_t \tag{18}$$

To check whether $H(a^t) = H(w^t)$, we must inquire whether the polynomial in L on the right side of (18) has a square-summable inverse in nonnegative powers of L. Such an inverse exists if and only if the zeros of $\det\left\{ I + \begin{bmatrix} C & -C \end{bmatrix} \begin{bmatrix} zI - A^* \end{bmatrix}^{-1} \begin{bmatrix} BD^{-1} \\ K \end{bmatrix} \right\}$ are all less than unity in modulus.[6] The following theorem gives an easy way to check this condition

Theorem III.1. *When D^{-1} exists, the zeros of*

$$\det\left\{ I + \begin{bmatrix} C & -C \end{bmatrix} \begin{bmatrix} zI - A^* \end{bmatrix}^{-1} \begin{bmatrix} BD^{-1} \\ K \end{bmatrix} \right\}$$

equal the eigenvalues of $A - BD^{-1}C$ and the eigenvalues of A.

Proof. Write

$$I + C^*(zI - A^*)^{-1} B^* = \left\{ I + \begin{bmatrix} C & -C \end{bmatrix} \begin{bmatrix} zI - A^* \end{bmatrix}^{-1} \begin{bmatrix} BD^{-1} \\ K \end{bmatrix} \right\}$$

where

$$\begin{aligned} C^* &= \begin{bmatrix} C & -C \end{bmatrix} \\ B^* &= \begin{bmatrix} BD^{-1} \\ K \end{bmatrix}. \end{aligned}$$

Now apply the partitioned inverse formula

$$\det(a)\det(d + ca^{-1}b) = \det(d)\det(a + bd^{-1}c)$$

with $a = I, b = C^*, c = B^*, d = (zI - A^*)$ to get

$$\det(I + C^*(zI - A^*)^{-1} B^*) = \frac{\det(zI - A^* + B^* C^*)}{\det(zI - A^*)}. \tag{19}$$

Compute

$$zI - A^* + B^* C^* = zI - \begin{bmatrix} A - BD^{-1}C & BD^{-1}C \\ 0 & A \end{bmatrix},$$

an equation that shows that the zeros of $\det zI - A^* + B^* C^*$ equal the eigenvalues of $A - BD^{-1}C$ and the eigenvalues of A. Using this result in (19) shows that the zeros of $\det(I + C^*(zI - A^*)^{-1} B^*)$ equal the eigenvalues of $A - BD^{-1}C$ and the eigenvalues of A. □

[6] If one or more zeros equal unity in modulus, an autoregressive representation fails to exist, but nevertheless, it is true that $H(a^t) = H(w^t)$. See Whittle (XXXX) and Hansen and Sargent (2005), chapter 2).

Under assumptions 1 and 2, theorem III.1 implies that in order to check whether $H(a^t) = H(z^t)$, we can simply check the eigenvalues of $A - BD^{-1}C$. The theorem allows us to check invertibility in terms of the fundamental objects A, B, C, D without actually computing the innovation representation and K, Σ via the Kalman filter.

Another way to express this point is to note that we can compute K and Σ directly without having to solve the Riccati equation (6), as we show in the following theorem.

Theorem III.2. *Suppose that D^{-1} exists and that $A - BD^{-1}C$ is a stable matrix. Then in the steady state Kalman filter, $K = BD^{-1}$ and $\Sigma = 0$.*

Proof. Notice that $\Sigma = 0$ solves the steady state Riccati equation (6). Notice also that with $\Sigma = 0$, equation (7) implies that $K = BD^{-1}$. Furthermore, the Riccati difference equation corresponding to the steady state equation (6) can be represented as

$$\begin{aligned} \Sigma_{t+1} &= (A - K_t C)\Sigma_t (A - K_t C)' + BB' \\ &+ KDD'K' - BD'K' - KDB' \end{aligned} \tag{20}$$

where

$$K_t = (A\Sigma_t C' + BD')(C\Sigma_t C' + DD')^{-1}.$$

Under the conditions of the theorem, $A - K_t C$ converges to a stable matrix $A - BD^{-1}C$ and successive iterates Σ_t converge to zero starting from any positive semidefinite initial Σ_0. □

Corollary III.3. *Under the conditions of theorem III.1, $Dw_t = G\varepsilon_t$ and the innovations covariance matrix $GG' = DD'$. Thus, we are free to set $G = D$. Of course, the choice of G is unique only up to postmultiplication by an orthogonal matrix.*

Proof. It can be verified directly from (11) that when the conditions of theorem III.1 hold and, therefore, $K = BD^{-1}$, it follows that $G\varepsilon_t = Dw_t$. □

The assertions in theorems III.1 and III.2 can be viewed as extensions to the case of a vector process of well known results for a first order pure moving average process that are contained in the following example:

Example III.4. *Take the scalar pure m.a. process*

$$y_{t+1} = w_{t+1} + \alpha w_t.$$

Let the state be $x_t = w_t$ so that we have a state space representation with $A = 0, B = 1, C = \alpha, D = 1$. Evidently,

$$A - BD^{-1}C = -\alpha,$$

which is a stable matrix if and only if $|\alpha| < 1$, in which case $K = B$.

The consequences of theorems III.1, III.2, and corollary III.3 for the impulse response function are studied in the following section.

IV. Impulse Response Functions

In section III we provide an easy way to check invertibility (theorem III.1) and a mapping from contemporaneous economic shocks to contemporaneous innovations for the invertible case (corollary III.3). In this section, we investigate the implications of these results for the impulse responses function from innovations to observables and its relationship with the impulse responses function from economic shocks to observables. Finally, we will also comment on the implications of invertibility for We assume that conditions 1 and 2 hold.

IV.1. **From Economic Shocks w to Observables y.** We are often interested in an impulse response function from the of economic shocks, w_t's, to the observables, y_t's,

$$y_t = \mu_y + d(L)w_t \tag{21}$$

where L is the lag operator, $d(L) = \sum_{j=0}^{\infty} d_j L^j$, $\sum_{j=0}^{\infty} \text{trace}(d_j d_j') < +\infty$, and μ_y is the mean of y. If all eigenvalues of A are less than unity in modulus, except for a single unit eigenvalue associated with a constant state variable, elementary calculations with system (1)-(2) deliver

$$y_t = \mu_y + [C(I - AL)^{-1}BL + D]w_t, \tag{22}$$

so that evidently

$$d_0 = D \tag{23}$$

$$d_j = CA^{j-1}B \quad j \geq 1. \tag{24}$$

Formula (22) tells us how to compute the impulse response function directly from the state space representation (A,B,C,D) of the economic model.[7]

Although the impulse response function is unique, the matrices (A,B,C,D) used to represent an impulse response function in (22) are not. Different (A,B,C,D)'s can deliver the same μ_y, $d(L)$. For convenience, one often selects a particular member of this class by choosing a minimum state realization of μ_y, $d(L)$.[8]

---

[7]The Matlab control toolkit program `impulse.m` calculates $d(L)$ from (A,B,C,D).

[8]The Matlab control toolkit command `sys=ss(sys,'min')` replaces an (A,B,C,D) with an equivalent minimal state realization.

IV.2. **From Innovations ε to Observables y.** For some G, let us now to calculate the impulse response functions from innovations, ε_t's, to observables , y_t's,

$$y_t = \mu_y + c(L)G\varepsilon_t \tag{25}$$

where $c(L) = \sum_{j=0}^{\infty} c_j L^j = (I - \sum_{j=1}^{\infty} A_j L^j)^{-1}$. This is a Wold moving average representation.[9]

The shock process ε_t is said to be 'fundamental for y' because it is by construction in the space spanned by square summable linear combinations of current and past values of the y process. Representation (25) is a population version of the impulse response function reported by a typical VAR researcher.

IV.3. **Invertibility and Impulse Responses Functions.** As corollary III.3 states, if the model is invertible, then we can choose $G = D$. If this is the case, we can rewrite (25) as

$$y_t = \mu_y + c(L)Dw_t, \tag{26}$$

but, since the impulse response function is unique, (23), (24), and (26) imply that

$$c_0 = I \tag{27}$$

$$c_j = CA^{j-1}BD^{-1} \quad j \geq 1. \tag{28}$$

Hence

$$c_j D = d_j \quad j \geq 0. \tag{29}$$

Equation (29) shows that, if the model is invertible, there exists and identification scheme, $G = D$, such that the impulse response function asociated with the VAR matches the one of the economic model.

IV.4. **Invertibility and Restricted VAR.** In addition, (27) and (28) imply that we can rewrite (26) as

$$y_t = [D + C(I - AL)^{-1}BD^{-1}DL]w_t \tag{30}$$

Since the eigenvalues of $A - BD^{-1}C$ are less than unity in modulus, the partitioned inverse formula allow us to invert the operator $[D + C(I - AL)^{-1}BD^{-1}DL]$ in (30) and get an autoregressive representation given by

$$y_t = C[I - (A - BD^{-1}C)L]^{-1}BD^{-1}y_{t-1} + Dw_t, \tag{31}$$

implying that

$$A_j = C(A - BD^{-1}C)^{j-1}BD^{-1}, \quad j \geq 0.$$

in (3).

[9] A VAR representation does not exist when $\det c(z)$ has zeros on the unit circle. See Whittle (XXX) and Hansen and Sargent (1991). **Tom: add statement of finite approximations that still exist.**

V. A BADLY BEHAVED EXAMPLE: A PERMANENT INCOME MODEL

A consumer maximizes

$$-.5\sum_{t=0}^{\infty}\beta^t[(c_t-b)^2+\varepsilon(k_t-k_{t-1})^2]$$

subject to the asset accumulation equations

$$k_t+c_t\le Rk_{t-1}+d_t$$

with k_{-1} as an initial condition, and where the pure endowment process d_t follows the two-component process

$$d_t=\mu_d+\frac{1}{1-\rho_1 L}\sigma_1 w_{1t}+\frac{1}{1-\lambda_1 L}\sigma_2 w_{2t}$$

where $|\rho_1|<1$ and $|\lambda_1|<1$. Where ε is a very small positive number, the term $\varepsilon(k_t-k_{t-1})^2$ is a small adjustment cost that we include to select an interesting solution.[10] We follow Hall (1978) and set $R\beta=1$ in order to deliver the outcome that k_t and c_t are cointegrated.[11] Set $R=1.05, \rho_1=.9, \lambda_1=.6, \mu_d=5, b=30$. Let $d_{1t}=\frac{1}{1-\rho_1 L}\sigma_1 w_{1t}$, $d_{2t}=\frac{1}{1-\lambda_1 L}\sigma_2 w_{2t}$, and $d_t=\mu_d+d_{1t}+d_{2t}$. Define the state vector as $x_t=\begin{bmatrix}k_{t-1} & 1 & d_{1t} & d_{2t}\end{bmatrix}'$ and let the observable variables be $y_{t+1}=\begin{bmatrix}c_{t+1} & d_{t+1}\end{bmatrix}'$. Then the solution of the consumer's problem can be expressed as[12]

$$\begin{aligned}x_{t+1} &= Ax_t+Bw_{t+1}\\ y_{t+1} &= Cx_t+Dw_{t+1}\end{aligned}$$

where $w_t=\begin{bmatrix}w_{1t} & w_{2t}\end{bmatrix}'$ is a vector white noise with mean zero and identity contemporaneous covariance matrix; and[13]

$$A=\begin{bmatrix}1.0000 & 0.0000 & 0.6667 & 0.8889\\ 0 & 1.0000 & 0 & 0\\ 0 & 0 & 0.9000 & 0\\ 0 & 0 & 0 & 0.6000\end{bmatrix}$$

[10] If we set $\varepsilon =$, the solution of the problem is $c_t=b$.

[11] This outcome occurs in the limit as $\varepsilon \searrow 0$.

[12] We used Hansen and Sargent's (2005) Matlab program `solvea.m`. To produce (A,B,C,D) from the solution produced by `solvea.m`, set `S=[sc;sd(1,:)], A=ao, B=c, C=S*ao, D=S*c`.

[13] We can read the random walk property of the optimal k_t directly from A and B.

$$B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0.5000 & 0 \\ 0 & 0.8000 \end{bmatrix}$$

$$C = \begin{bmatrix} 0.0500 & 5.0000 & 0.3333 & 0.1111 \\ 0 & 5.0000 & 0.9000 & 0.6000 \end{bmatrix}$$

$$D = \begin{bmatrix} 0.1667 & 0.0889 \\ 0.5000 & 0.8000 \end{bmatrix}.$$

It follows that

$$A - BD^{-1}C = \begin{bmatrix} 1.0000 & 0.0000 & 0.6667 & 0.8889 \\ 0 & 1.0000 & 0 & 0 \\ -0.2250 & -20.0000 & -0.1500 & -0.2000 \\ 0.2250 & 15.0000 & 0.1500 & 0.2000 \end{bmatrix}.$$

This matrix has an eigenvalue of 1.05 (which equals R). Therefore, the mapping (18) is not invertible. It follows that the Hilbert space $H(\varepsilon^t)$ spanned by the history of VAR shocks is smaller than the space $H(w^t)$ spanned by the space of economic shocks. Furthermore, in general the shapes of the impulse responses to ε_t and w_t differ.

Invertibility of the mapping (18) for the permanent income model is obtained for the square cases in which the observation vector is either $\begin{bmatrix} c_t & k_t \end{bmatrix}'$ or $\begin{bmatrix} c_t & k_t - k_{t-1} \end{bmatrix}'$. With these observation vectors, the offending zero at R 'flips' to become a zero at R^{-1}. The reader can verify this claim numerically by recomputing $A - BD^{-1}C$ and applying Theorem III.1.

V.1. **Historical note.** Sargent (1987a), Hansen, Roberds, and Sargent (1991), and Roberds (1991) studied a version of this example in response to a question asked by Robert E. Lucas, Jr., at a 1985 Minneapolis Fed conference: with a constant interest rate, what restrictions are placed on a vector autoregression for government expenditures and tax receipts by the hypothesis of present value budget balance? The permanent income model is isomorphic to a stochastic version of a tax smoothing model in the style of Barro (1979) with total tax collections τ_t replacing consumption c_t and government expenditures g_t replacing the endowment d_t. This model imposes two restrictions on the c_t, d_t process: (1) present value budget balance, and (2) c_t must be a martingale. Because it implies equal present values of the moving average coefficients of d_t and c_t to either economic shock w_{it}, present value budget balance puts a zero of R into the operator on the right side of (18) and is therefore the source of noninvertibility.

Hansen, Roberds, and Sargent (1991) went on to answer Lucas's question by showing that present value budget balance by itself puts no testable restrictions on the infinite order VAR of $\begin{bmatrix} c_t & d_t \end{bmatrix}'$.

The permanent income example with c_t, d_t as the observables is one in which the invertibility condition is bound to fail. That stands as a counterexample to a presumption that VAR shocks always readily match up with the economic shocks w_t. It is thus one important example of things that can go wrong. However, there are other examples in which things go right. In the next section, we turn to such an example.

VI. A Better Behaved Example: Jonas Fisher's Two Technology Shock Model

Jonas Fisher (2003) assesses the impact of technology shocks on business cycles by imposing what he motivates as long-run restrictions on an estimated non-structural VAR. Fisher explicitly acknowledges that a necessary condition for his procedure to be compelling is that the mapping (11) be invertible, assumes that his model is such that this condition is satisfied, and imposes a long-run restriction on G that is suggested by an analysis of an exogenous growth model with two orthogonal unit-root technology processes. In this section, we use our theorem III.1 to verify that that invertibility assumption is indeed valid at calibrated values for the parameters in Fisher's model. Section VII extends the example by replacing calibrated parameter values with their Bayesian prior from Fisher's data, obtained using an MCMC algorithm.

There is a representative household whose preferences over stochastic sequences of consumption C_t and leisure $1-L_t$ are representable by the utility function

$$E_0 \sum_{t=0}^{\infty} \beta^t \left(\log C_t + \psi \log(1-L_t)\right) \tag{32}$$

where $\beta \in (0,1)$ is the discount factor and E_0 is the conditional expectation operator. The resource constraint is given by

$$C_t + X_t = A_t K_t^{\alpha} L_t^{1-\alpha}, \tag{33}$$

and the law of motion for capital is:

$$K_{t+1} = (1-\delta) K_t + V_t X_t, \tag{34}$$

and:

$$\begin{aligned} A_t &= e^{\gamma + C_a(L) w_{at}} A_{t-1},\ \gamma \geq 0 && (35) \\ V_t &= e^{\upsilon + C_\upsilon(L) w_{\upsilon t}} V_{t-1},\ \upsilon \geq 0 && (36) \\ \left[w_{at}, w_{\upsilon t}\right]' &\sim \mathcal{N}(0, D),\ D \text{ diagonal} && (37) \end{aligned}$$

where $C_a(L)$ and $C_\upsilon(L)$ are square summable polynomials in the lag operator L. We assume that C_a and C_υ are both the identity operator.

VI.1. **Transforming Variables.** Define the scaling variable $Z_t = A_{t-1}^{\frac{1}{1-\alpha}} V_{t-1}^{\frac{\alpha}{1-\alpha}} = \left(A_{t-1}V_{t-1}^{\alpha}\right)^{\frac{1}{1-\alpha}}$ and the transformations $\widetilde{C}_t = \frac{C_t}{Z_t}$ and $\widetilde{K}_t = \frac{K_t}{Z_t V_{t-1}}$. Using methods described, for example, by Uhlig (1999), first compute a steady state transformed capital stock $\widetilde{K}_{ss}$, then obtain log-linearized decision rules for capital

$$\log\widetilde{K}_{t+1} - \log\widetilde{K}_{ss} = a_1\left(\log\widetilde{K}_t - \log\widetilde{K}_{ss}\right) + a_2 w_{a,t} + a_3 w_{\upsilon,t},$$

for hours worked

$$\log L_t - \log L_{ss} = b_1\left(\log\widetilde{K}_t - \log\widetilde{K}_{ss}\right) + b_2 w_{a,t} + b_3 w_{\upsilon,t},$$

and for consumption

$$\log\widetilde{C}_t - \log\widetilde{C}_{ss} = c_1\left(\log\widetilde{K}_t - \log\widetilde{K}_{ss}\right) + c_2 w_{a,t} + c_3 w_{\upsilon,t},$$

where $a_1, a_2, a_3, b_1, b_2, b_3, c_1, c_2$ and c_3 are constants that depend on the structural parameters of the economy. For this model, it turns out that $a_2 = a_3$, $a_2 = -\frac{a_1}{1-\alpha}$, and $b_2 = -\frac{b_1}{1-\alpha}$. Then use these loglinear decision rules and the definitions of the transformed variables to obtain the following state-space system in logarithms of our original (untransformed) variables:

$$\begin{pmatrix} 1 \\ \Delta\log K_{t+1} \\ \log L_t \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1-a_1}{1-\alpha}(\gamma+\upsilon) & a_1 & 0 \\ -b_1\frac{\gamma+\upsilon}{1-\alpha} & b_1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ \Delta\log K_t \\ \log L_{t-1} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ \frac{1}{1-\alpha}+a_2 & \frac{1}{1-\alpha}+a_3 \\ b_2 & b_3 \end{pmatrix} \begin{pmatrix} w_{a,t} \\ w_{\upsilon,t} \end{pmatrix} \tag{38}$$

$$\begin{pmatrix} \Delta\log\frac{Y_t}{L_t} \\ \log L_t \end{pmatrix} = \begin{pmatrix} \gamma+\alpha b_1\frac{\gamma+\upsilon}{1-\alpha} & \alpha(1-b_1) & 0 \\ -b_1\frac{\gamma+\upsilon}{1-\alpha} & b_1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \Delta\log K_t \\ \log L_{t-1} \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 1-\alpha b_2 & -\alpha b_3 \\ b_2 & b_3 \end{pmatrix} \begin{pmatrix} w_{a,t} \\ w_{\upsilon,t} \end{pmatrix} \tag{39}$$

Equations (38), (39) form a state space system of the form A,B,C, and D. For the rest of the argument in this section, we use the following parameter values $\beta = 1.00, \psi = 5.49, \alpha = 0.452, \delta = 0.0006, \gamma = 0.002, \upsilon = 0.011, \sigma_a = 0.012, \sigma_\upsilon = 0.008$. Later we will describe how we estimated these parameter values.

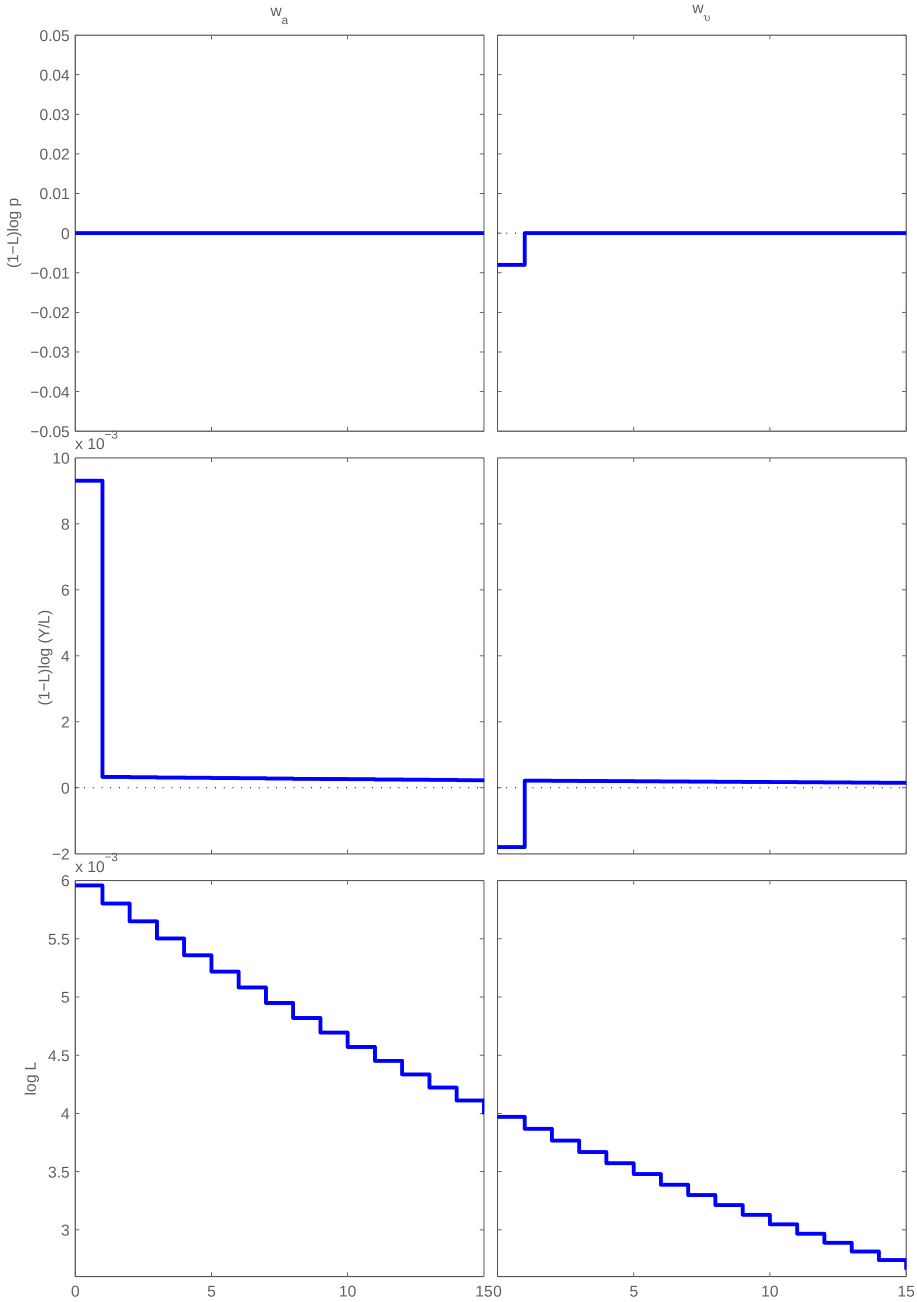

FIGURE 1. Common impulse response functions for VAR and economic structure for Fisher's model. For the two-observed-variable model, only the bottom two panels are pertinent.

VI.2. **Using Theorem III.1.** We turn first to the bivariate system formed by (38) and (39). This is a 'square system' with two shocks and two observables. The eigenvalues of $A - B * D^{-1}C$ are zero, which means that (11) is invertible, and also that by setting $G = D$ and $K = BD^{-1}$, we attain an innovations representation with $G\varepsilon_t = Dw_t$. We also know from theorem III.2 that the history y^t provides a perfect estimate of the state $\Delta \log K_t$ and $\log L_{t-1}$. By setting $G = D$, the impulse response function to ε_t associated with a identified VAR perfectly matches the impulse response function to the theoretical shocks w_t. This impulse response function is reported in the bottom two rows of panels of Figure 2. For this example, invertibility prevails so that we are assured that there exists some G satisfying $GG' = DD'$ that makes the impulse response for the identified VAR match the theoretical impulse response to the w's. However, the example also confirms the doubt expressed in the epigraph from Lucas and Stokey. The required G must be equal to $D = \begin{bmatrix} 0.4370 & -0.0252 \\ 0.1908 & 0.0763 \end{bmatrix}$, which lacks zeros, as Lucas and Stokey feared. To make the impulse response function from an unrestricted VAR match requires a way of discovering G while being initially ignorant of D.

VI.3. **Fisher's Identification Procedure.** Fisher fits an unrestricted VAR with $\Delta \log p_t$ as an observable, therefore in order to explain his procedure, we need to deifne the system formed by (38) and the observer equation

$$\begin{pmatrix} \Delta \log p_t \\ \Delta \log \frac{Y_t}{L_t} \\ \log L_t \end{pmatrix} = \begin{pmatrix} -\nu & 0 & 0 \\ \gamma + \alpha b_1 \frac{\gamma+\upsilon}{1-\alpha} & \alpha(1-b_1) & 0 \\ -b_1 \frac{\gamma+\upsilon}{1-\alpha} & b_1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ \Delta \log K_t \\ \log L_{t-1} \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 1-\alpha b_2 & -\alpha b_3 \\ b_2 & b_3 \end{pmatrix} \begin{pmatrix} w_{a,t} \\ w_{\upsilon,t} \end{pmatrix} \tag{40}$$

Before we describe Fisher's bit of magic, have to work around a technical difficulty. When using the three variable observation vector (40), we have to confront the fact that we have a stochastically singular system because two shocks are driving three observables. Two eliminate that stochastic singularity while staying as close as possible to Fisher's model, we add a very small measurement error to $\log L_t$.

To identify G from a three variable system, Fisher notes that $\Delta \log p_t = -\Delta \log V_t$ is an exogenous white noise that equals $w_{\upsilon,t}$. Therefore, any scheme for factoring GG' that identifies the row of $G\varepsilon_t$ associated with $\Delta \log p_t$ with $w_{\upsilon,t}$ should work. Fisher used the following scheme.

Let $\Omega = C\Sigma C' + DD'$ be the covariance matrix of $G\varepsilon_t$ from the infinite order VAR (see equation (8). Fisher (2003), footnote 5) used the following version of a procedure of Blanchard and Quah to identify G. First, he formed $\hat{c}(1) = (I - \sum_{j=1}^{\infty} A_j)^{-1}$. Second, he computed a lower triangular Cholesky factor x of $\hat{c}(1)\Omega\hat{c}(1)'$, so that $xx' = \hat{c}(1)\Omega\hat{c}(1)'$. Third, after noting that $\hat{c}(1)G$ is a factor of $\hat{c}(1)\Omega\hat{c}(1)'$, he computes $G = \hat{c}(1)^{-1}x$.

We have verified that applying this scheme produces a $G = D$. The impulse response associated both with the infinite order VAR with $G = D$ conforms with the impulse response to the economic shocks. It is reported in Figure 2. The only change from the VAR computed for our two variable system is the addition of the top panel in Figure 2 and the addition of equation (**??**) to our ARMA representation for the VAR and the theoretical economic model.[14]

While these calculations confirm the validity of Fisher's identification procedure for his theoretical model, they don't really contradict the skepticism about zero restrictions on G or $\sum_{j=1}^{\infty} A_j L^j$ expressed in Lucas and Stokey's epigraph. The phrase '*endogenous variables*' in the epigraph bears remembering. Fisher's zero restriction that $\Delta \log p_t$ is never influenced by $w_{a,t}$ comes from having specified the model so that $\Delta \log p_t$ is econometrically exogenous.

VI.4. **Interpretation.** While these calculations confirm the validity of Fisher's identification procedure for his theoretical model, they don't really contradict the skepticism about zero restrictions on G or $\sum_{j=1}^{\infty} A_j L^j$ expressed in Lucas and Stokey's epigraph. The phrase '*endogenous* variables' in the epigraph bears remembering. Fisher's zero restriction that $\Delta \log p_t$ is never influenced by $w_{a,t}$ comes from having specified the model so that $\Delta \log p_t$ is econometrically exogenous.[15]

VI.5. **Finite order VARs.** Using the projection formulas in Appendix A, we computed population versions of finite order vector autoregressions for both the two and three variable VARs.[16] We computed these for VARs of length 1 and 4. Both gave such close approximations to the impulse response functions reported in Figure 2 that it was impossible to detect any difference when we plotted them on along side those in Figure 2.

VII. BAYESIAN ESTIMATION OF FISHER'S MODEL

In this section we describe how we estimate Fisher's model.

VIII. AN ALMOST WELL BEHAVED EXAMPLE: THE EHL MODEL

Many sticky price models imply a reduction in hours worked after a productivity shock hits the economy. This theoretical finding has been the motivation of some empirical work (see Gali, XXXX) trying to identify a productivity shock and its consequences for hours

[14] A simple alternative to Fisher's scheme would also work, namely, choosing G as a triangular Cholesky factor of the innovation covariance matrix Ω that sets $G_{12} w_{\upsilon,t} = \Delta \log p_t$.

[15] Fisher has a nice discussion of this point in his paper and describes how the particular zero restriction that we have imposed wouldn't prevail with a modified technology for producing investment goods.

[16] Chari, Kehoe, and McGrattan use Monte Carlo methods to approximate these VARs. It is easier and more accurate to use the analytic formulas in the appendix.

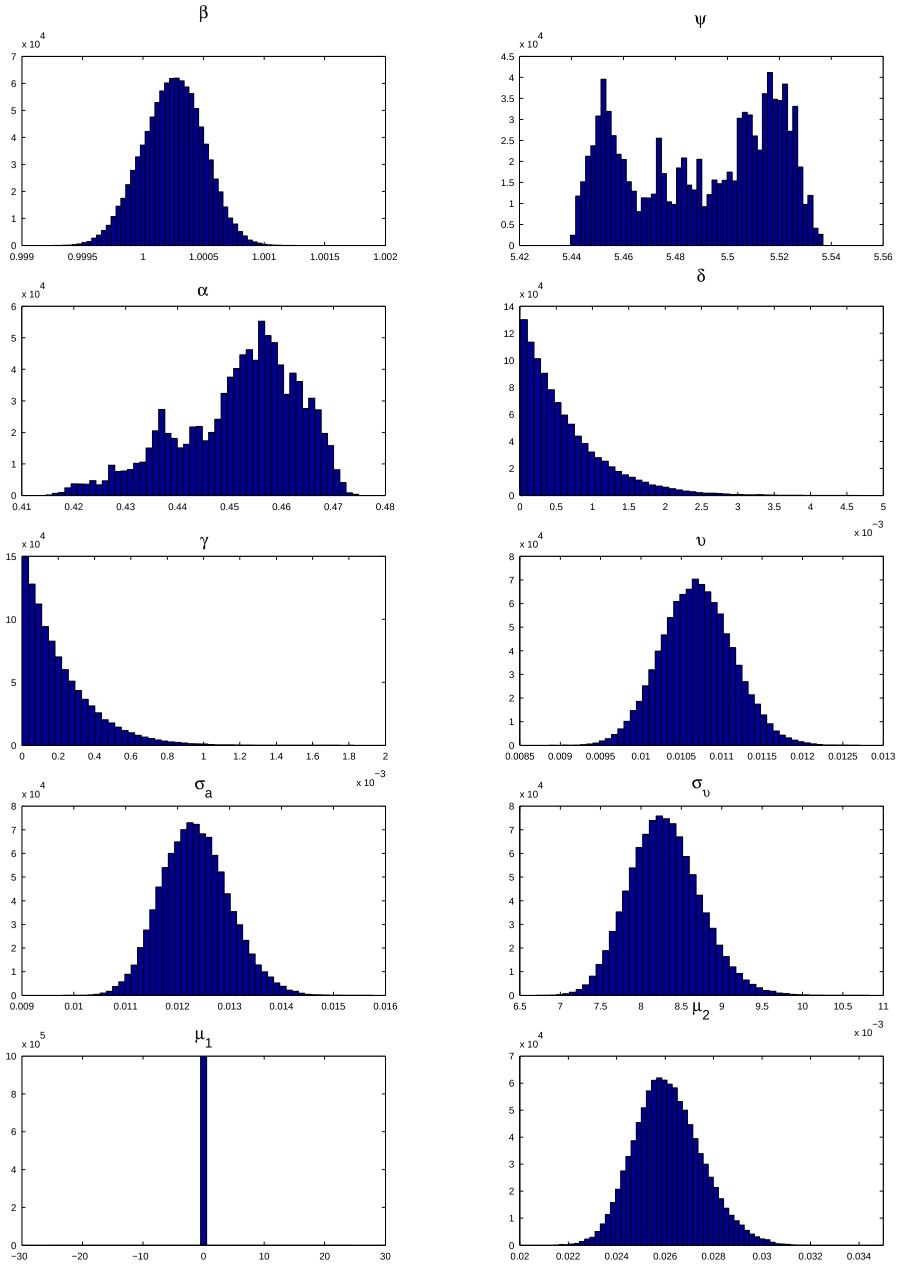

FIGURE 2. Posterior Distribution for Fisher's model.

worked. As mentioned in the introduction, if we want to recover the theoretical impulse response function with a unrestricted VAR we need to check for invertibility of our theory.

For that porpouse we analyze Erceg, Henderson and Levin's (2003) model with sticky prices and sticky wages. In this section, we use our theorem III.1 to verify that that invertibility assumption is indeed valid at calibrated values for the parameters in Erceg, Henderson and Levin's model. Section IX extends the example by replacing calibrated parameter values with their Bayesian posterior from Erceg, Henderson and Levin's model, obtained using an MCMC algorithm. Since this model is well known in the literature we will only present the equations describing the log deviation from steady-state values of the variables.

VIII.1. **The Model Equations.** First, we have the Euler equation which relates output growth with the real rate of interest

$$y_t = E_t y_{t+1} - \sigma(r_t - E_t \Delta p_{t+1} + E_t g_{t+1} - g_t) \quad (41)$$

where y_t denotes output, r_t is the nominal interest rate, g_t is the preference shifter shock, p_t is the price level, and σ is the elasticity of intertemporal substitution.

The production function and the real marginal cost of production are:

$$y_t = a_t + (1-\delta)n_t \quad (42)$$

$$mc_t = w_t - p_t + n_t - y_t \quad (43)$$

where a_t is a technology shock, n_t is the amount of hours worked, mc_t is the real marginal cost, w_t is the nominal wage, and δ is the capital share of output.

The marginal rate of substitution (mrs_t) between consumption and hours is:

$$mrs_t = g_t + \frac{1}{\sigma} y_t + \gamma n_t \quad (44)$$

where γ is the inverse elasticity of labor supply with respect to real wages. Hence, the preference shifter shock affects both the consumption Euler equation and the marginal rate of substitution.

The pricing decision of the firm under the Calvo-type restriction delivers the following forward looking equation for price inflation (Δp_t):

$$\Delta p_t = \beta E_t \Delta p_{t+1} + \kappa_p(mc_t + \lambda_t) \quad (45)$$

where $\kappa_p = \frac{(1-\delta)(1-\theta_p\beta)(1-\theta_p)}{\theta_p(1+\delta(\bar{\varepsilon}-1))}$ and $\bar{\varepsilon} = \frac{\bar{\lambda}}{\lambda - 1}$ is the steady state value of ε, the elasticity of substitution between types of goods. λ_t is the price markup shock, θ_p is the probability of keeping prices fixed during the period, and β is the discount factor.

Since the model presents sticky wages, the following equation relates wages and the mrs:

$$\Delta w_t = \beta E_t \Delta w_{t+1} + \kappa_w(mrs_t - (w_t - p_t)) \quad (46)$$

where $\kappa_w = \frac{(1-\theta_w)(1-\beta\theta_w)}{\theta_w(1+\phi\gamma)}$, θ_w is the probability of keeping wages fixed in a given period, and ϕ is the elasticity of substitution between different types of labor in the production function. With staggered wage setting, it is no longer true that workers remain on their labor supply schedule. Hence, the driving force of current nominal wage growth is expected nominal wage growth, as well as the distance between the marginal rate of substitution and the real wage.

We use the following specification for the Taylor rule:

$$r_t = \rho_r r_{t-1} + (1-\rho_r)\left[\gamma_\pi \Delta p_t + \gamma_y y_t\right] + ms_t \tag{47}$$

where γ_π and γ_y are the long run responses of the monetary authority to deviations of inflation and output from their steady state values, and ms_t is the monetary shock. We include an interest rate smoothing parameter, ρ_r, following recent empirical work (see Clarida, Galí and Gertler, 2000).

In order to close the model, we need the identity that links real wage growth, nominal wage growth and price inflation:

$$w_t - p_t = w_{t-1} - p_{t-1} + \Delta w_t - \Delta p_t \tag{48}$$

We specify the shocks to follow the following processes:

$$\begin{aligned} a_t &= \rho_a a_{t-1} + w_t^a \\ g_t &= \rho_g g_{t-1} + w_t^g \\ ms_t &= w_t^{ms} \\ \lambda_t &= w_t^\lambda \end{aligned}$$

where each innovation w_t^i follows a $\mathcal{N}\ (0, \sigma_i^2)$ distribution, for $i = a, g, m, \lambda$. The innovations are uncorrelated with each other.

VIII.2. **The A, B, C, and D matrices.** With the system in this form, we can apply Uhlig's (1999) formulae to find the coefficients of the policy function are of the form:

$$K_t = P * K_{t-1} + Qz_t, \tag{49}$$

and

$$L_t = R * K_{t-1} + Sz_t, \tag{50}$$

where $K_t = \left[\begin{array}{ccccc} w_t - p_t & r_t & \Delta p_t & \Delta w_t & y_t \end{array}\right]'$, $L_t = \left[\begin{array}{cccc} n_t & mc_t & mrs_t & c_t \end{array}\right]'$, and $z_t = \left[\begin{array}{cccc} a_t & g_t & ms_t & \lambda_t \end{array}\right]'$.

A more convinient way of writing (49) and (50) is

$$\left[\begin{array}{cc} K_t' & z_t' \end{array}\right]' = \left[\begin{array}{cc} P & Q*N \\ N & 0 \end{array}\right] \left[\begin{array}{cc} K_{t-1}' & z_{t-1}' \end{array}\right]' + \left[\begin{array}{c} Q \\ I \end{array}\right] w_t, \tag{51}$$

and

$$L_t = \left[\begin{array}{cc} R & S*N \end{array}\right] \left[\begin{array}{cc} K_{t-1}' & z_{t-1}' \end{array}\right]' + Sw_t, \tag{52}$$

where $w_t = \left[\begin{array}{cccc} w_t^a & w_t^{ms} & w_t^\lambda & w_t^g \end{array}\right]'$.

Let us assume that the observables we consider are $Y_t = [\ \Delta p_t \quad \Delta n_t \quad y_t \quad w_t - p_t\ ]'$, then we obtain the following state-space system in log deviations from steady state:

$$\left[\ K_t' \quad z_t' \quad n_t\ \right]' = A \left[\ K_{t-1}' \quad z_{t-1}' \quad n_{t-1}\ \right]' + Bw_t, \tag{53}$$

$$Y_t = C \left[\ K_{t-1}' \quad z_{t-1}' \quad n_{t-1}\ \right]' + Dw_t, \tag{54}$$

where

$$A = \begin{bmatrix} P & Q*N & 0 \\ N & 0 & 0 \\ R_{1,\cdot} & (S*N)_{1,\cdot} & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} QQ \\ R_{1,\cdot} \end{bmatrix}$$

$$C = \begin{bmatrix} A_{3,\cdot} \\ A_{6,\cdot} - [\ 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 1\ ]' \\ A_{5,\cdot} \\ A_{1,\cdot} \end{bmatrix},$$

and

$$D = \begin{bmatrix} B_{3,\cdot} \\ B_{6,\cdot} \\ B_{5,\cdot} \\ B_{1,\cdot} \end{bmatrix}$$

where $A_{j,\cdot}$ stands for the jth row of matrix A.

Equations (53), (54) form a state space system of the form A,B,C, and D. We use the following calibrated parameter values $\beta = 0.9, \bar{\varepsilon} = 6, \delta = 0.4, \sigma = 0.5, \theta_p = \theta_w = 0.9, \gamma = 2, \phi = 6, \gamma_y = 0.125, \gamma_\pi = 1.5, \rho_r = \rho_a = \rho_g = 0.9$, and $\sigma_i = 0.05$ for $i = a, m, \lambda, g$.

VIII.3. **Using Theorem III.1.** We turn first to the system formed by (53) and (54). This is a 'square system' with four shocks and four observables. The eigenvalues of $A - B*D^{-1}C$ are all zero but one that is exactly equal to one. Since the model is specified in deviations from the steady state, the mean is zero, hence (11) is not invertible. The good news is that since the eigenvalues are not strictly bigger than one, we still have that $H(a^t) = H(w^t)$ and by setting $G = D$ and $K = BD^{-1}$, we attain an innovations representation with $G\varepsilon_t = Dw_t$. We also know from theorem III.2 that the history y^t provides a perfect estimate of the state $\left[\ K_{t-1}' \quad z_{t-1}' \quad n_{t-1}\ \right]'$. By setting $G = D$, the impulse response function to ε_t associated with a identified VAR perfectly matches the impulse response function to the theoretical shocks w_t.

IX. Bayesian estimation of EHL's Model

In this section we describe how we estimate the EHL's model.

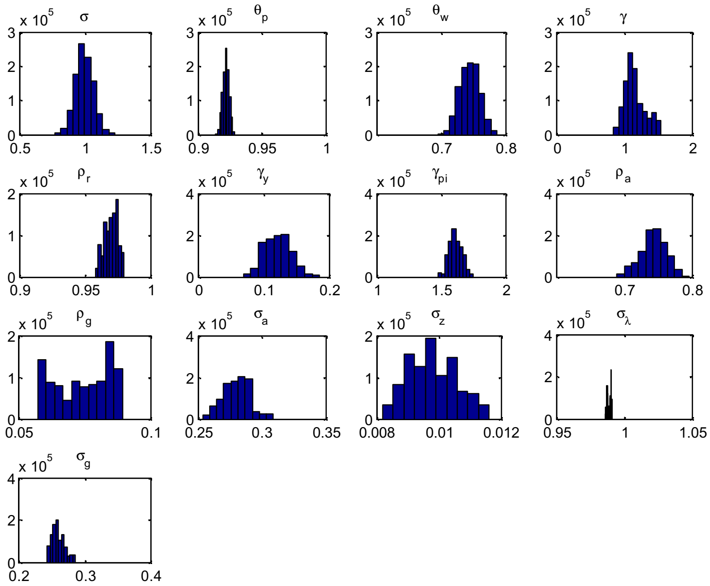

FIGURE 3. Posterior Distribution for EHL's model.

X. Concluding Remarks

Appendix A. Finite order autoregressions

This appendix describes formulas for taking an A,B,C,D and forming the associated nth order vector autoregression.[17]

A.1. Moment formulas.

Take an economic model in the state-space form (1)-(2). Assume that all of the eigenvalues of A are less than unity in modulus, except possibly for a unit eigenvalue that is affiliated with the constant. If present, the unit eigenvalue determines the unconditional mean vector μ_x of x via

$$(I-A)\mu_x = 0. \tag{A1}$$

The stationary covariance matrix of x is $c_x(0) = E(x-\mu_x)(x-\mu_x)'$ and can be computed by solving the discrete Sylvester equation

$$c_x(0) = Ac_x(0)A' + BB', \tag{A2}$$

which can be solved by Hansen and Sargent's matlab program `doublej`. (The indigenous matlab program `dlyap.m` works only when there are no unit eigenvalues of A.) The autocovariance of x is $c_x(j) = E(x_t-\mu_x)(x_{t-j}-\mu_x)'$ and can be computed from

$$c_x(j) = A^j c_x(0), \quad j \geq 1. \tag{A3}$$

Let $\mu_y = C\mu_x$ be the mean of y and $c_y(j) = E(y_t-\mu_y)(y_{t-j}-\mu_y)'$. Elementary calculations establish:

$$\begin{aligned}
c_y(0) &= Cc_x(0)C' + DD' && \text{(A4)}\\
c_y(j) &= CA^j c_x(0)C' + CA^{j-1}BD', \quad j \geq 1 && \text{(A5)}\\
c_y(-j) &= c_y(j)', \quad j \geq 1. && \text{(A6)}
\end{aligned}$$

A.2. Projection formulas.

We want to calculate the nth order vector autoregressions

$$y_t - \mu_y = \sum_{j=1}^{n} A_j^{(n)}(y_{t-j}-\mu_y) + \varepsilon_t^{(n)} \tag{A7}$$

where $\varepsilon_t^{(n)}$ satisfies the orthogonality conditions

$$E[\varepsilon_t^{(n)}(y_{t-j}-\mu_y)'] = 0, \quad j = 1,\ldots,n. \tag{A8}$$

---

[17]Riccardo Colacito has written a Matlab programs `ssvar.m` that by implementing these formulas accepts an (A,B,C,D) and a positive integer n and yields all of the objects defining an nth order VAR. His program `varss.m` takes an nth order VAR and forms a state space system A,B,C,D, a useful tool for using Matlab to compute impulse response functions for estimated VAR's.

The orthogonality conditions, also known as the normal equations, can be written

$$c_y(k) = \sum_{j=1}^{n} A_j^{(n)} c_y(k-j)', \quad k = 1, \ldots, n. \tag{A9}$$

Writing out (A9) and solving for $\begin{bmatrix} A_1^{(n)} & A_2^{(n)} & \ldots & A_n^{(n)} \end{bmatrix}$ gives:

$$\begin{bmatrix} c_y(1)' \\ c_y(2)' \\ \vdots \\ c_y(n)' \end{bmatrix}' \begin{bmatrix} c_y(0) & c_y(1) & \cdots & c_y(n-1) \\ c_y(-1) & c_y(0) & \cdots & c_y(n-2) \\ \vdots & \vdots & & \vdots \\ c_y(1-n) & c_y(2-n) & \cdots & c_y(0) \end{bmatrix}^{-1} = \begin{bmatrix} A_1^{(n)} & A_2^{(n)} & \ldots & A_n^{(n)} \end{bmatrix}. \tag{A10}$$

The covariance matrix $\Sigma^{(n)} = E\varepsilon_t^{(n)}\varepsilon_t^{(n)\prime}$ of the innovations is

$$\Sigma^{(n)} = c_y(0) - \begin{bmatrix} A_1^{(n)\prime} \\ A_2^{(n)\prime} \\ \vdots \\ A_n^{(n)\prime} \end{bmatrix}' \begin{bmatrix} c_y(0) & c_y(1) & \cdots & c_y(n-1) \\ c_y(-1) & c_y(0) & \cdots & c_y(n-2) \\ \vdots & \vdots & & \vdots \\ c_y(1-n) & c_y(2-n) & \cdots & c_y(0) \end{bmatrix} \begin{bmatrix} A_1^{(n)\prime} \\ A_2^{(n)\prime} \\ \vdots \\ A_n^{(n)\prime} \end{bmatrix}. \tag{A11}$$

Please note that the nth order autoregression can also be expressed as

$$y_t = \alpha^{(n)} + \sum_{j=1}^{n} A_j^{(n)} y_{t-j} + \varepsilon_t^{(n)} \tag{A12}$$

where

$$\alpha^{(n)} = (I - \sum_{j=1}^{n} A_j^{(n)})\mu_y. \tag{A13}$$

Here μ_y is a properly scaled eigenvector of $\sum_{j=1}^{(n)}$ associated with the unit eigenvalue, where the proper scaling assures that the mean of the constant 1 is 1. Our Matlab program `ssvar` takes an (A,B,C,D), with the understanding that the constant 1 is the first state variable, and computes an nth order VAR. Our program `varss` takes an nth order VAR and forms the pertinent (A,B,C,D).

APPENDIX B. APPROXIMATION FORMULAS

Hansen and Sargent (1993) develop the following frequency domain approximation criterion. The truth is a stationary stochastic process for observables y with unconditional mean ν and spectral density matrix $F(\omega)$. An incorrect approximating model with parameter vector δ implies that the mean of y is $\mu(\delta)$ and that the spectral density $G(\omega,\delta)$ is defined as

$$E[y_t - \mu(\delta)][y_{t-j} - \mu(\delta)]' = \frac{1}{2\pi}\int_{-\pi}^{\pi} \exp(i\omega j)G(\omega,\delta)d\omega. \tag{B1}$$

Hansen and Sargent show that maximum likelihood estimates of δ asymptotically minimize the approximation criterion

$$\mathscr{A}(\delta) = \mathscr{A}_1(\delta) + \mathscr{A}_2(\delta) + \mathscr{A}_3(\delta) \tag{B2}$$

where

$$\mathscr{A}_1(\delta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \log\det G(\omega,\delta)d\omega \tag{B3}$$

$$\mathscr{A}_2(\delta) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \operatorname{trace}[G(\omega,\delta)^{-1}F(\omega)]d\omega \tag{B4}$$

$$\mathscr{A}_3(\delta) = [\nu - \mu(\delta)]'G(0,\delta)^{-1}[\nu - \mu(\delta)]. \tag{B5}$$

To compute (B2), we can use Riemann sums to approximate the integrals. Let N be even and take as our approximant

$$\hat{\mathscr{A}} = \frac{1}{N}\sum_{j=0}^{N-1}\left\{\log\det G(\omega_j,\delta) + \operatorname{trace}[G(\omega,\delta)^{-1}F(\omega_j)]\right\} + [\nu-\mu(\delta)]'G(0,\delta)^{-1}[\nu-\mu(\delta)] \tag{B6}$$

where $\omega_j = \frac{2\pi j}{N}$, $j = 0,1,\ldots,N-1$.

B.1. **Calculating theoretical spectral density matrices.** For a state space system A,B,C,D the spectral density matrix of the observables is

$$S_y(\omega) = C(I - A\exp(-i\omega))^{-1}BB'(I - A'\exp(i\omega))^{-1}C' + DD'. \tag{B7}$$

To compute $F(\omega)$ and $G(\omega,\delta)$ in (B6), we will apply this formula twice, once for the A,B,C,D matrices that define the truth, and again with the A,B,C,D matrices associated with the approximating nth order VAR. The A,B,C,D associated with the approximating nth order VAR are defined implicitly by the following state space representation of the VAR:

$$\begin{bmatrix} 1 \\ y_{t+1} \\ y_t \\ y_{t-1} \vdots \\ y_{t-n} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \alpha & A_1^{(n)} & A_2^{(n)} & \cdots & A_n^{(n)} \\ 0 & I & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} 1 \\ y_t \\ y_{t-1} \\ y_t \vdots \\ y_{t-n+1} \end{bmatrix} + \begin{bmatrix} 0 \\ S \\ 0 \\ \vdots \\ 0 \end{bmatrix} w_{t+1}^{(n)} \tag{B8}$$

$$y_{t+1} = \begin{bmatrix} \alpha & A_1^{(n)} & A_2^{(n)} & \cdots & A_n^{(n)} \end{bmatrix} \begin{bmatrix} 1 \\ y_t \\ y_{t-1} \\ y_t \vdots \\ y_{t-n+1} \end{bmatrix} + Sw_{t+1}^{(n)} \tag{B9}$$

where S is a Cholesky factor of $\Sigma^{(n)} = E\varepsilon_t^{(n)}\varepsilon_t^{(n)\prime}$ computed from (A11):

$$\Sigma^{(n)} = SS'.$$

The misspecification consists of the assertion that $w_{t+1}^{(n)}$ is orthogonal to the history of y for dates before $t-n$.

REFERENCES

FISHER, J. D. M. (2003): "Technology Shocks Matter," Working paper, Federal Reserve Bank of Chicago.

HANSEN, L. P., AND T. J. SARGENT (1991): "Lecture Notes on Least Squares Prediction Theory," in *Rational expectations econometrics*, ed. by L. P. Hansen, and T. J. Sargent, pp. 13–44. Westview Press, Boulder.

——— (2005): *Recursive Linear Models of Dynamic Econmies*. Princeton, Princeton, New Jersey.

LARS PETER HANSEN, W. R., AND T. J. SARGENT (1991): "Time Series Implications of Present Value Budget Balance and of Martingale Models of Consumption and Taxes," in *Rational expectations econometrics*, ed. by L. P. Hansen, and T. J. Sargent, pp. 121–162. Westview Press, Boulder.

ROBERDS, W. (1991): "Implications of Expected Present Value Budge Balance: Application to Postwar U.S. Data," in *Rational expectations econometrics*, ed. by L. P. Hansen, and T. J. Sargent, pp. 163–174. Westview Press, Boulder.

ROBERT, C. C. (2001): *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation, Second Edition*. Springer-Verlag.

ROBERT E. LUCAS, J., AND N. L. STOKEY (1987): "Money and Interest in a Cash-in-Advance Economy," *Econometrica*, 55(1), 491–513.