News, Non-Invertibility, and Structural VARs*

Eric R. Sims
University of Notre Dame and NBER
December 22, 2011

Abstract

An approximate state space representation of a DSGE model implies a VAR in terms of variables observed by an econometrician. The model is said be non-invertible if there exists no mapping from VAR innovations to economic shocks. Non-invertibility arises when the observed variables fail to perfectly reveal the state variables of the model. The imperfect observation of the state drives a wedge between the VAR innovations and the deep shocks, potentially invalidating conclusions drawn from structural impulse response analysis in the VAR. The principle contribution of this paper is to show that non-invertibility should not be thought of as an “either/or” proposition – even when a model has a non-invertibility, the wedge between VAR innovations and economic shocks may be small, and structural VARs may nonetheless perform reliably. As an increasingly popular example, so-called “news shocks” generate foresight about changes in future fundamentals – such as productivity, taxes, or government spending – and lead to an unassailable missing state variable problem and hence non-invertible VAR representations. Simulation evidence from a medium scale DSGE model augmented with news shocks about future productivity reveals that structural VAR methods often perform remarkably well in practice, in spite of a known non-invertibility. Impulse responses obtained from VARs closely correspond to the theoretical responses from the model, and the estimated VAR responses are successful in discriminating between alternative, nested specifications of the underlying DSGE model. Since the non-invertibility problem is, at its core, one of missing information, conditioning on more information, for example through factor augmented VARs, is shown to either ameliorate or eliminate invertibility problems altogether.

* Author contact information: esims1@nd.edu. This is a revised version of a paper that previously circulated under the title “Non-Invertibilities and Structural VARs.” I am grateful to Robert Barisky for helpful comments and advice.
1 Introduction

Structural VARs (SVARs) are frequently used, either formally or informally, as a tool to construct, refine, and parameterize dynamic stochastic general equilibrium (DSGE) models.\(^1\) The validity of this practice hinges upon whether or not SVARs can reliably uncover relevant objects of interest from fully specified DSGE models, such as the impulse responses to structural shocks. There are numerous problems that can arise when analyzing VARs estimated on relatively short data sets; among these are issues due to downward biased autoregressive coefficients in finite samples and the so-called lag truncation bias.\(^2\) Another, potentially more severe, problem is that there may exist no direct mapping between the innovations in the observable variables included in a VAR and the structural shocks of the underlying DSGE model.

While VARs are often touted for their flexibility and lack of imposed structure, and indeed are often pejoratively referred to as “atheoretic,” there nevertheless exists a tight connection between fully specified DSGE models and VARs. The approximate equilibrium of a log-linearized DSGE model can usually be expressed in terms of a state space system. The state space of the model implies a VAR in terms of the variables observed by an econometrician. The model is said to be invertible if there exists, in population, a linear mapping from the VAR innovations to the deep structural shocks of the underlying DSGE model. If no such mapping exists, the model is said to be non-invertible.

The so-called “non-invertibility” (or sometimes “non-fundamental”) problem has been known to exist for some time but has only recently received much attention.\(^3\) At its core, it means that innovations from a VAR on a set of observable variables may not, even in population, be used to exactly uncover the structural shocks from a fully-specified DSGE model that serves as the data generating process for those observables. The non-invertibility problem is fundamentally one of missing information. As shown in Section 2 below, it arises when the observed variables do not span the full state space of the underlying DSGE model. When this happens, the population innovations of a VAR on observed variables are a combination of the true structural shocks from the underlying DSGE model and what essentially amounts to measurement error from forecasting the state space conditional on the observables. The mixing of the structural shocks with this measurement error potentially confounds any analysis based on the rotations of the reduced form VAR innovations, which is the core of the structural VAR methodology.

A principle objective of this paper is to argue and show that non-invertibility should not be thought of as an “either/or” proposition. There may exist situations in which a model has a

---

\(^1\) There is a long list of papers in this literature, and any selection of papers will invariably be incomplete. Nevertheless, a sampling of papers that employ VARs as tools to construct, refine, or estimate the parameters of DSGE models includes Gali (1999); Christiano, Eichenbaum, and Evans (2005); Fisher (2006); Sims and Zha (2006); Altig, Christiano, Eichenbaum, and Linde (2011); Barsky and Sims (2011b); and Sims (2011).

\(^2\) For a discussion of how downward-biased AR coefficients can affect identification, see Faust and Leeper (1997). For a discussion of the lag-truncation bias – which means that most DSGE models imply VAR(∞) models, while researchers in practice estimate finite order models – see Chari, Kehoe, and McGrattan (2008).

\(^3\) For an early treatment, see Lippi and Riechlin (1994). The most recent canonical treatment of this problem is found in Fernandez-Villaverde, Rubio-Ramirez, Sargent, and Watson (2007).
non-invertible VAR representation but where structural VARs nevertheless perform quite reliably.
I make these arguments on the basis of analysis of a conventional, medium scale DSGE model
with a number of real and nominal frictions. A non-invertibility is hard-wired into the model by
augmenting it with a particular kind of shock that generates foresight about changes in future
productivity – a so called “news shock.”

There has recently been renewed interest in the economic effects of “news shocks” about changes
in future fundamentals – although in this paper I consider news about productivity, other papers
have explored news about taxes and government spending changes. Much of this literature is
empirical and makes use of structural VAR techniques. For example, Beaudry and Portier (2006)
and Barsky and Sims (2011a) study the role of news about future productivity, while Mountford
and Uhlig (2009) focus on anticipated changes in government spending, all within the context of
structural VARs. Leeper, Walker, and Yang (2011) show, however, that foresight about changes in
future state variables very likely leads to non-invertible VAR representations. The intuition is fairly
straightforward. News shocks – which are, by construction, unobservable to an econometrician –
are also state variables, as agents in the underlying economy must keep track of lagged values
of these shocks when making current decisions. Hence, foresight leads to an unassailable missing
state variable problem. The non-observation of the state drives a wedge between VAR innovations
and economic shocks, and potentially invalidates any conclusions drawn from structural VARs.
Given the growing popularity of the news literature, an important contribution of this paper is to
investigate the quantitative relevance of non-invertibility for analysis based on SVARs.

The DSGE model laid out in Section 3 nests simpler models that are in common use. The model
features two sources of “real rigidity” – internal habit formation in consumption and investment
adjustment costs – and one source of “nominal rigidity” – price stickiness according to the staggered
contracts in Calvo (1983). Under specific parameter restrictions it nests simpler models – for
example, setting all three parameters governing the degree of frictions to zero yields the canonical
real business cycle (RBC) model, while setting the habit formation and investment adjustment
cost parameters to zero but keeping price stickiness gives rise to a textbook sticky price model with
capital. To keep things simple, the model features only two stochastic disturbances – a conventional
surprise productivity shock and the news shock about future productivity.

Using the “poor man’s invertibility condition” derived in Fernandez-Villaverde, Rubio-Ramirez,
Sargent, and Watson (2007), in Section 4 I analytically show that the presence of news shocks
generates non-invertible VAR representations when TFP growth and any other variable of the model
are observed. I then conduct a battery of Monte Carlo exercises in which I examine the performance
of apparently well-specified SVARs. In particular, I estimate two variable VARs featuring TFP
growth and output on data simulated from the model. I rotate the statistical innovations into
structural shocks using a Choleski decomposition of the innovation variance-covariance matrix with
TFP growth ordered first. This recursive ordering conforms with the theoretical implications of
the model – the innovation in TFP growth is identified with the conventional surprise technology

---

4 As is common in the literature, I restrict attention to the “square case” in which the number of observables equals
the number of shocks.
shock, while the innovation in output orthogonalized with respect to TFP growth is identified with
the news shock. While this is a particularly simple example, it is nevertheless instructive and is
not without precedent in the literature. For example, Beaudry and Portier (2006) identify news
shocks with stock price innovations orthogonalized with respect to TFP growth.\footnote{There is a debate within the literature over how to best identify news shocks within SVAR settings. See, for example, the discussion in Barsky and Sims (2011a). So as to focus on the role of non-invertibility, I restrict attention here to a simple case in which a recursive identification is valid.}

In spite of the presence of a known non-invertibility, SVARs applied to model simulated data
perform remarkably well in recovering the impulse responses to the model’s two shocks. The
estimated responses to both kinds of technology shocks are qualitatively in line with the predictions
of the model. Though there are biases in the estimated responses because of the non-invertibility,
these are typically quantitatively small and are mostly at long forecast horizons. The short horizon
responses, in contrast, are estimated quite precisely. The different nested parameterizations of the
model make very different predictions about the behavior of output in response to both news and
surprise technology shocks. For example, in the RBC model output falls in response to good news
about future productivity and rises by more than productivity after a surprise technology shock.
In contrast, in the fully parameterized model output rises after good news and rises after a surprise
positive technology shock, but by substantially less than productivity. The estimated VARs do a
very good job at picking these features up in the simulations. This means that the VARs can be a
very effective tool at discriminating between different nested versions of a model.

In short, the simulation results show that VARs can be an effective tool for empirical researchers
even if a model is technically non-invertible. While this finding may prove comforting to some, it is
nevertheless not possible to draw sweeping conclusions about the perniciousness of non-invertibility
more generally. To that end, in Section 5 I consider what steps a researcher can take to ameliorate
or eliminate problems stemming from non-invertibility while remaining within the flexible limited
information framework that VARs provide. The second principle contribution of the paper is to
emphasize that non-invertibility is fundamentally a problem of missing information; hence, adding
more information is the most straightforward way to deal with it.

The presence of foresight about future productivity leads to missing state variables – essentially
the agents in the model must keep track of lagged news shocks when making current decisions.
In Section 5 I consider conditioning on additional “information variables” in the VARs applied to
model-generated data. These information variables are noisy signals about future productivity, but
are otherwise not central to the solution of the model. If the signals are precise enough – or if one
conditions on enough information variables – then the missing states are essentially revealed, and
the invertibility problem vanishes. I show that adding these variables to an otherwise standard VAR
reduces the small biases in the estimated VARs. One quickly runs into a sort of “curse of dimension-
ality” problem, however, as conditioning on many information variables quickly consumes degrees
of freedom. I therefore consider compressing the information variables using principal components
and estimate factor-augmented VARs. The impulse responses obtained from the factor augmented
VARs are essentially unbiased at all horizons. These results suggests that factor-augmented VARs,
which are coming into increasing popularity, are an effective tool for researchers interested in using VAR techniques but who are nevertheless concerned about the potential for biases stemming from non-invertibility. Such methods are relatively simple to implement, while maintaining the relative lack of structure that full information techniques require.\footnote{Dupor and Han (2011) develop a four step procedure to partially identify impulse responses when non-invertibility is feared present. Their approach cannot always eliminate non-invertibility, whereas conditioning on a very large information set can.}

The remainder of the paper is organized as follows. Section 2 reviews the mapping between DSGE models and VARs, discusses reasons why invertibility may fail, and derives a simple condition to check whether a system is non-invertible. Section 3 lays out a DSGE model with both real and nominal frictions and a hard-wired non-invertibility because of the presence of foresight about productivity. Section 4 conducts a number of Monte Carlo exercises for different, nested versions of the model to examine the performance of SVARs. Section 5 proposes conditioning on more information as a way to overcome invertibility issues. The final section offers concluding thoughts.

\section{The Mapping Between DSGE Models and VARs}

A log-linearized DSGE models yields a state-space representation of the following form:

\begin{align*}
    s_t &= As_{t-1} + B\varepsilon_t \\
    x_t &= Cs_{t-1} + D\varepsilon_t
\end{align*}

\(s_t\) is a \(k \times 1\) vector of state variables, \(x_t\) is a \(n \times 1\) vector of observed variables, and \(\varepsilon_t\) is a \(m \times 1\) vector of structural shocks. The variance-covariance matrix of these shocks is diagonal and given by \(\Sigma_{\varepsilon}\). \(A\), \(B\), \(C\), and \(D\) are matrixes of conformable size whose elements are functions of the deep parameters of the model. So as to facilitate a comparison with standard assumptions in the structural VAR literature, I restrict attention to the case in which \(n = m\), so that there are the same number of observed variables as shocks. \(D\) is thus square and hence invertible.

One can solve for \(\varepsilon_t\) from (2) as:

\[\varepsilon_t = D^{-1}(x_t - Cs_{t-1})\]

Plugging this into (1) yields:

\[s_t = (A - BD^{-1}C)s_{t-1} + BD^{-1}x_t\]

Solving backwards, one obtains:

\[s_t = (A - BD^{-1}C)^{t-1}s_0 + \sum_{j=0}^{t-1}(A - BD^{-1}C)^{j-1}BD^{-1}x_{t-j}\]
If \( \lim_{t \to \infty} (A - BD^{-1}C)^{t-1} = 0 \), then the history of observables perfectly reveals the current state. This requires that the eigenvalues of \( (A - BD^{-1}C) \) all be strictly less than one in modulus. If this condition is satisfied, (3) can be plugged into (2) to yield a VAR in observables in which the VAR innovations correspond to the structural shocks:

\[
x_t = C \sum_{j=0}^{t-1} (A - BD^{-1}C)^{j-1} BD^{-1}x_{t-1-j} + D\varepsilon_t
\]

The condition that the eigenvalues of \( (A - BD^{-1}C) \) all be strictly less than unity is the “poor man’s invertibility” condition given in Fernandez-Villaverde, Rubio-Ramirez, Sargent, and Watson (2007). It is a sufficient condition for a VAR on observables to have innovations that map directly back into structural shocks in population. When satisfied, a finite order VAR(\( p \)) on \( x_t \) will yield a good approximation to (4), and conventional estimation and identification strategies will allow one to uncover the model’s impulse responses to structural shocks.

When this condition for invertibility is not satisfied the state space system nevertheless yields a VAR representation in the observables, though the VAR innovations no longer correspond to the structural shocks. The crux of the problem when the invertibility condition is not met is that the observables do not perfectly reveal the state vector. To see this, use the Kalman filter to form a forecast of the current state, \( \hat{s}_t \), given observables and a lagged forecast:

\[
\hat{s}_t = (A - KC) \hat{s}_{t-1} + Kx_t
\]

Here \( K \) is the Kalman gain. It is the matrix that minimizes the forecast error variance of the filter, i.e. \( \Sigma_s = E((s_t - \hat{s}_t)(s_t - \hat{s}_t)') \). \( K \) and \( \Sigma_s \) are the joint solutions to the Riccati equations:

\[
\Sigma_s = (A - KC) \Sigma_s (A - KC)' + B\Sigma_\varepsilon B' + K D\Sigma_\varepsilon D' K' - B\Sigma_\varepsilon D'K' - K D\Sigma_\varepsilon B'
\]

\[
K = (A\Sigma_s C' + B\Sigma_\varepsilon D') (C\Sigma_s C' + D\Sigma_\varepsilon D')^{-1}
\]

Given values of \( K \) and \( \Sigma_s \), add and subtract \( C\hat{s}_{t-1} \) from the right hand side of (2) to obtain:

\[
x_t = C\hat{s}_{t-1} + u_t
\]

\[
u_t = C(s_{t-1} - \hat{s}_{t-1}) + D\varepsilon_t
\]

Lagging (5) one period and recursively substituting into (8), one obtains an infinite order VAR representation in the observables:

\[
x_t = (A - KC)^{t-1}\hat{s}_0 + C \sum_{j=0}^{t-1} (A - KC)^j Kx_{t-1-j} + u_t
\]

Under weak conditions, Hansent and Sargent (2007) show that \( (A - KC) \) is a stable matrix, so that the \( (A - KC)^{t-1}\hat{s}_0 \) term disappears in the limit and the infinite sum on the lagged observables...
converges in mean square.

The innovations in this VAR representation are comprised of two orthogonal components: the true structural shocks and the error in forecasting the state. The innovation variance is given by:

$$\Sigma_u = C \Sigma_s C' + D \Sigma_\epsilon D'$$  \hspace{1cm} (11)

Fernandez-Villaverde, Rubio-Ramirez, Sargent and Watson (2007) show that the eigenvalues of $(A - BD^{-1}C)$ being less than unity in modulus implies that $\Sigma_s = 0$. When $\Sigma_s = 0$, then $\Sigma_u = D \Sigma_\epsilon D'$, and it is straightforward to show that (10) reduces to (4). If the “poor man’s invertibility” condition is not satisfied, then $\Sigma_s \neq 0$, and the innovation variance from the VAR is strictly larger than the innovation variance in the structural model. This discussion unveils a critical point – the failure of invertibility is part and parcel a failure of the observables to reveal the state vector. Non-invertibility is fundamentally an issue of missing information.

This discussion also reveals that non-invertibility is not necessarily an “either/or” proposition. (11) makes clear that the extent to which a failure of invertibility might “matter” quantitatively is how large $\Sigma_s$ – i.e. how hidden the state is. This has a number of implications. First, even if the condition for invertibility fails, $\Sigma_s$ may nevertheless be “small,” meaning that $\Sigma_u \approx D \Sigma_\epsilon D'$. Put differently, the VAR innovations may very closely map into the structural shocks even if a given system is technically non-invertible. Second, what observable variables are included in a VAR might matter – some observables may do a better job of forecasting the missing states, hence leading to smaller $\Sigma_s$ and a closer mapping between VAR innovations and structural shocks. Finally, adding more observable variables should always work to lower $\Sigma_s$, and thus ameliorate problems due to non-invertibility. This means that estimating larger dimensional VARs may generally be advantageous relative to the small systems that are frequently estimated in the literature. It also potentially speaks to the benefits of estimating factor augmented models, which can efficiently condition on large information sets. I return to this issue in Section 5 below.

3 A DSGE Model

For the purposes of examining quantitatively how important non-invertibility may be to applied researchers, I consider a standard DSGE model with a particular kind of shock that is known to lead to an invertibility problem. The model is DSGE model with both nominal and real frictions. On the real side, there is habit formation in consumption, convex investment adjustment costs, and imperfect competition. On the nominal side there is price rigidity. A nice feature of the model is that, under certain parameter restrictions, it reverts to a simple neoclassical growth model with variable labor supply. The shock that generates the non-invertibility, for reasons to be discussed below, is a “news shock” about anticipated technological change.

The next subsections describe the decision problems of the various actors in the model as well as results concerning aggregation and the definition of equilibrium.
### 3.1 Households

There is a representative household that consumes a final good, makes decisions to accumulate capital, supplies labor, holds riskless one period nominal bonds issued by a government, and holds nominal money balances. Imposing standard functional forms, its decision problem can be written:

$$\max_{c_t, n_t, I_t, k_{t+1}, M_{t+1}, B_t} \sum_{t=0}^{\infty} \beta^t \left( \ln (c_t - \gamma c_{t-1}) - \frac{\theta n_t^{1+\xi}}{1 + \xi} + \chi \left( \frac{M_{t+1}}{pt} \right)^{1-\nu} \right)$$

subject to:

$$c_t + I_t + \frac{B_{t+1} - B_t}{pt} + \frac{M_{t+1} - M_t}{pt} \leq w_t n_t + R_t k_t + \frac{B_{t-1}}{pt} + \frac{\Pi_t}{pt} + \frac{\Delta I}{pt}$$

$$k_{t+1} = \left( 1 - \frac{\tau}{2} \left( \frac{I_t}{I_{t-1}} - \Delta t \right)^2 \right) I_t + (1 - \delta) k_t$$

$\gamma$ governs the degree of internal habit formation in consumption, $\xi$ is the inverse Frisch labor supply elasticity, and $\nu$ will determine the elasticity of the demand for real balances with respect to the nominal interest rate, $i_t$. $p_t$ is the price of goods in terms of money. $w_t$ and $R_t$ are the real factor prices for labor and capital, respectively. $B_t$ and $M_t$ are the dollar amounts of bonds and money with which the household enters the period. $I_t$ is investment in physical capital, $\Pi_t$ is nominal profits distributed lump sum from firms, and $T_t$ is nominal lump sum tax/transfers from the government. $\tau$ is a parameter governing the cost of adjusting investment, with $\Delta t$ the (gross) balanced growth path growth rate of investment.

The first order conditions for an interior solution to the household problem are:

$$\lambda_t = \frac{1}{c_t - \gamma c_{t-1}} - E_t \frac{\beta \gamma}{c_{t+1} - \gamma c_t}$$  \hspace{1cm} (12)$$

$$\theta n_t^{\xi} = \lambda_t w_t$$  \hspace{1cm} (13)$$

$$\lambda_t = \beta E_t \lambda_{t+1} (1 + i_t) \frac{pt}{pt+1}$$  \hspace{1cm} (14)$$

$$\mu_t = \beta E_t (\lambda_{t+1} R_{t+1} + (1 - \delta) \mu_{t+1})$$  \hspace{1cm} (15)$$

$$\chi \left( \frac{M_{t+1}}{pt} \right)^{-\nu} = \left( \frac{i_t}{1 + i_t} \right) \lambda_t$$  \hspace{1cm} (16)$$

$$\lambda_t = \mu_t \left( 1 - \frac{\tau}{2} \left( \frac{I_t}{I_{t-1}} - \Delta t \right)^2 - \tau \left( \frac{I_t}{I_{t-1}} - \Delta t \right) \left( \frac{I_t}{I_{t-1}} \right) \right) + \beta E_t \mu_{t+1} \tau \left( \frac{I_{t+1}}{I_t} - \Delta t \right) \left( \frac{I_{t+1}}{I_t} \right)^2$$  \hspace{1cm} (17)$$

$\lambda_t$ is the current value Lagrange multiplier on the flow budget constraint and $\mu_t$ is the multiplier on the accumulation equation. (12) defines the marginal utility of consumption, (13) is a labor supply condition, (14) is the Euler equation for bonds, and (15) is the Euler equation for capital. (16) implicitly defines the demand for real balances. (17) is the first order condition with respect
to investment. When there are no adjustment costs, \( \lambda_t = \mu_t \), and (14)-(15) define the usual approximate arbitrage condition between the real interest rate on bonds and the return on capital.

### 3.2 Production

Production is split up into two sub-sectors. The final goods sector is competitive and aggregates a continuum of intermediate goods, \( y_{j,t}, j \in (0,1) \). The production technology for the final good is a CES aggregate of the intermediate goods, with \( \epsilon > 1 \) the elasticity of substitution:

\[
y_t = \left( \int_{0}^{1} \frac{y_{j,t}}{1-\epsilon} \, dj \right)^{\frac{1}{1-\epsilon}}
\]

Profit maximization yields a demand curve for each intermediate and an aggregate price index:

\[
y_{j,t} = \left( \frac{p_{j,t}}{p_t} \right)^{-\epsilon} y_t
\]

\[
p_t = \left( \int_{0}^{1} p_{j,t}^{1-\epsilon} \, dj \right)^{\frac{1}{1-\epsilon}}
\]

Intermediate goods firms are price-takers in factor markets and produce output according to a standard Cobb-Douglas production function:

\[
y_{j,t} = a_t k_{j,t}^{\alpha} n_{j,t}^{1-\alpha}
\]

\( a_t \) is a technology shifter that is common across firms. Because firms have pricing power (as long as \( \epsilon < \infty \)), it is helpful to break the firm problem into two parts. In the first stage firms choose inputs to minimize total cost subject to producing as much as is demanded at a given price:

\[
\min_{n_{j,t}, k_{j,t}} \ w_t n_{j,t} + R_t k_{j,t}
\]

s.t.

\[
a_t k_{j,t}^{\alpha} n_{j,t}^{1-\alpha} \geq \frac{1}{p_t} \left( \frac{p_{j,t}}{p_t} \right)^{-\epsilon} y_t
\]

The first order conditions are:

\[
w_t = mc_{j,t}(1-\alpha)a_t \left( \frac{k_{j,t}}{n_{j,t}} \right)^{\alpha}
\]

\[
R_t = mc_{j,t}\alpha a_t \left( \frac{k_{j,t}}{n_{j,t}} \right)^{\alpha-1}
\]

\( mc_{j,t} \), the multiplier on the production constraint, has the interpretation of real marginal cost. Because all intermediate firms face the same factor prices, it is straightforward to show that real
marginal cost will be the same across firms and that all firms will hire capital and labor in the same ratio.

It is assumed that firms face exogenous price stickiness in setting their prices. This makes the pricing problem dynamic. Following Calvo (1983) and much of the subsequent literature, let $1 - \phi$ be the probability that a firm is allowed to adjust its price in any period. This probability is independent of where the firm’s price is or when it last adjusted. When setting its price, the firm seeks to maximize the expected present discounted value of future profits, where profits are discounted by both the stochastic discount factor, $\Lambda_{t+s} = \beta^s \lambda_{t+s}$, and the probability that a price chosen today is still in effect in the future, $\phi^s$. The problem of a firm with the ability to update in date $t$ is:

$$\max_{p_{j,t}} E_t \sum_{s=0}^{\infty} (\phi \beta)^s \Lambda_{t+s} \left( \frac{p_{j,t}}{p_{t+s}} \right)^{-\epsilon} y_t - mc_t \left( \frac{p_{j,t}}{p_{t+s}} \right)^{-\epsilon} y_t$$

The solution is an optimal reset price, $p^*_t$, satisfying:

$$p^*_t = \frac{\sum_{s=0}^{\infty} (\phi \beta)^s \left( mc_{t+s} p^t_{t+s} y_{t+s} \right)}{\epsilon - 1 \sum_{s=0}^{\infty} (\phi \beta)^s \left( p^t_{t+s} y_{t+s} \right)}$$

Note that $p^*_t$ does not depend on $j$, and is hence the same for all updating price-setters. This follows from the fact that marginal cost, $mc_t$, is the same for all firms. The optimal reset price is essentially a markup over marginal cost. If $\phi = 0$, so that prices are flexible, this formula reduces to the standard fixed markup over marginal cost, with the markup given by $\frac{\epsilon}{\epsilon - 1}$.

### 3.3 Government

There is no government spending in the model. The sole role of the government is to set nominal interest rates according to a Taylor rule. The government then prints a sufficient amount of nominal money, $M_{t+1}$, so that the money market clears at the desired interest rate. Any seignorage revenue is remitted to households lump sum via $T_t$. The Taylor rule is:

$$i_t = \rho i_{t-1} + (1 - \rho) \psi_\pi (\pi_t - \pi^*) + (1 - \rho) \psi_y \left( \frac{y_t}{y_{t-1}} - \Delta_y \right)$$

(23)

$0 \leq \rho \leq 1$ is an interest rate smoothing parameter, $\pi^*$ is an exogenous inflation target, and $\psi_\pi$ and $\psi_y$ are response coefficients to inflation and the output “growth gap,” where $\Delta_y$ is the balanced growth path (gross) growth rate of output. I abstract from a monetary shock and restrict attention to parameter values that yield a determinate rational expectations equilibrium.

Remittance of seignorage revenue requires that:

$$T_t = M_{t+1} - M_t$$

(24)
3.4 Exogenous Process

For simplicity there is only one exogenous stochastic variable in the model – the level of technology, \( a_t \). It is assumed to follow a random walk with drift subject to two stochastic disturbances:

\[
\ln a_t = g_a + \ln a_{t-1} + e_t + u_{t-q}
\]  

(25)

\[
e_t \sim N(0, \sigma_e)
\]

\[
u_t \sim N(0, \sigma_u)
\]

\( e_t \) is a standard technology shock. \( u_t \) is a “news shock” in the sense that agents in the economy see it in period \( t \), but it has no effect on the level of technology until period \( t + q \), where \( q \geq 1 \). For reasons to be spelled out below, the presence of news shocks like this can easily lead to the non-invertibility problem. The two kinds of technology shocks are assumed to be distributed independently. This assumption is far less restrictive than it may seem – it would be straightforward to allow the shocks to be correlated and then partition them into orthogonal components.

3.5 Aggregation and Equilibrium

The notion of equilibrium is standard – it is a set of prices and quantities consistent with the first order conditions of households and firms holding and the budget constraints binding with equality. Market-clearing requires that total capital and labor demand equal that supplied by households:

\[
\int_0^1 n_{j,t} dj = n_t
\]

\[
\int_0^1 k_{j,t} dj = k_t
\]

Aggregate inflation evolves according to:

\[
1 + \pi_t = \left( (1 - \phi) \left( 1 + \pi_t^* \right)^{1-\epsilon} + \phi \right)^{\frac{1}{1-\epsilon}}
\]  

(26)

Here \( 1 + \pi_t^* = \frac{p_{t}^*}{p_{t-1}} \). Aggregation of the intermediate firm production functions yields:

\[
y_t = a_t k_t^\alpha n_t^{1-\alpha} \frac{v_t}{v_t}
\]  

(27)

\( v_t \) is a deadweight loss due to price dispersion:

\[
v_t = \int_0^1 \left( \frac{p_{j,t}}{p_t} \right)^{-\epsilon} dj
\]
It can be written recursively as:

\[ v_t = (1 - \phi) \left( \frac{1 + \pi_t^#}{1 + \pi_t} \right)^{-\epsilon} + \phi(1 + \pi_t)^\epsilon v_{t-1} \] (28)

Aggregate bond market-clearing \((B_t = 0)\) and the combination of the government and household budget constraints yields the standard aggregate accounting identity:

\[ y_t = c_t + I_t \] (29)

After normalizing variables to account for balanced growth owing to the unit root in technology, the model is solved via log-linearizing about the normalized steady state using standard techniques. The normalizations are as follows:

\[
\begin{align*}
\hat{a}_t &= \frac{a_t}{a_{t-1}}, & \hat{c}_t &= \frac{c_t}{a_t^{1/\alpha}}, & \hat{y}_t &= \frac{y_t}{a_t^{1/\alpha}}, & \hat{I}_t &= \frac{I_t}{a_t^{1/\alpha}}, & \hat{k}_t &= \frac{k_t}{a_t^{1/\alpha}}
\end{align*}
\]

Most other variables of the model, including hours, are stationary by construction. In logs, these normalizations imply cointegrating relationships that can be brought to the data. The solution of the transformed model gives rise to a state space system of the form in equations (1)-(2).

Different parameterizations of the model nest popular, simpler models. When \(\tau = 0, \gamma = 0, \phi = 0,\) and \(\epsilon = \infty\) the model reverts to a standard competitive real business cycle model. I will refer to this model as the “RBC model.” \(\epsilon < \infty\) gives rise an RBC model with imperfect competition; this leads to a steady state distortion (due to price being greater than marginal cost), but has no first order effects on the equilibrium dynamics, and is therefore not of much interest in its own right. \(\phi > 0\) but \(\tau = 0\) and \(\gamma = 0\) is a standard “sticky price model,” while \(\phi = 0\) with \(\tau > 0\) and \(\gamma > 0\) is an “RBC model with real frictions.” I will refer to the general specification of the model as the “full model.”

### 3.6 Why Do News Shocks Give Rise to Non-Invertibility?

A non-invertibility means that a VAR estimated on observable variables will fail to perfectly recover a model’s underlying structural shocks, even with an infinite sample size. As noted in Section 2, the problem occurs when the included observable variables fail to perfectly reveal the state vector. When this is the case, the innovations from a VAR in observables are a combination of the structural shocks and errors in forecasting the state.

In many circumstances a researcher concerned about non-invertibility can simply include the relevant state variables in the list of observables and estimate a VAR with those variables included. This is not feasible in a model with news shocks, because the presence of news shocks introduces an unassailable missing state variable problem. When there are anticipation effects, with \(q \geq 1,\) \(u_t\) becomes both a shock and a state variable. This is because agents at time \(t\) must keep track of realizations of \(u_{t-1}, \ldots, u_{t-q-1}\) when making choices at time \(t\); because these shocks have not yet
affected \( a_t \), even if \( a_t \) is observed the full state is hidden. This poses a potentially serious problem, since the entire structural VAR enterprise is about identifying shocks, not conditioning on them.

In order to see this point clearly, it is helpful to introduce additional state variables which keep track of lagged news shocks. Define these as \( z_{i,t} \) for \( i = 1, \ldots q - 1 \). Doing so allows one to write a process for technology that satisfies the Markov property:

\[
\ln a_t = g_a + \ln a_{t-1} + \varepsilon_t + z_{1,t-1} \quad (30)
\]
\[
z_{1,t} = z_{2,t-1} \quad (31)
\]
\[
z_{2,t} = z_{3,t-1} \quad (32)
\]
\[\vdots\]
\[
z_{q-1,t} = u_{t-q} \quad (33)
\]

The agents in the economy must keep track of the \( z \)s; given that these are just equal to the news shock at various lags, the econometrician cannot directly condition on the \( z \)s. Hence, the state vector cannot, in general, be observed based on a history of observables, and the conditions for invertibility are likely to fail.

4 VARs with Non-Invertibility: Monte Carlo Results

In this section I conduct several Monte Carlo experiments. The objective is to examine how well an apparently correctly specified structural VAR performs on model simulated data when there is a known non-invertibility. Can a VAR come close to replicating the structural model’s theoretical impulse responses to shocks when there is a non-invertibility? Can estimated VAR impulse responses be used to differentiate between competing models? This section seeks to provide some answers to these questions.

Consider the model described in the previous section (under any of the nested parameter configurations). The model has two stochastic shocks – the conventional surprise technology shock and the news shock. Suppose that a researcher observes TFP growth, \( \ln \hat{a}_t \), and another variable, say output, \( \ln \hat{y}_t \).\(^7\) Recall that because of the normalization, observed output is: \( \ln y_t = \frac{1}{1-\alpha} \ln a_t \), so this imposes the cointegrating relationships of the structural model.\(^8\) Suppose a researcher estimates a VAR(\( p \)) on the system \( x_t = [\ln \hat{a}_t \ \ln \hat{y}_t] \). A recursive orthogonalization of the innovations, with \( \ln \hat{a}_t \) “ordered” first, is apparently consistent with the implications of the model. Surprise movements in TFP growth would be identified with the surprise technology shock, while surprise movements in output (or any other variable) orthogonalized with respect TFP innovations would be identified with the news shock. This kind of empirical strategy is precisely what is often employed in the

---

\(^7\)I will hereafter focus on the case in which output is observed. The results are essentially the same conditioning on other observed variables.

\(^8\)Alternatives would be to estimate a VAR in levels, \( X_t = [\ln a_t \ \ln y_t] \), or a vector error correction model (VECM). These are all asymptotically equivalent. The Monte Carlo results below turn out to be very similar in all cases.
empirical literature. Beaudry and Portier (2006), for example, estimate a two variable with VAR with TFP and stock prices, identifying stock price movements uncorrelated with TFP innovations as news shocks.

For each of the nested parameter configurations (RBC, sticky price, real frictions, and the full model), I conduct the following Monte Carlo experiments. For the finite sample experiment, I create 500 different data sets with 200 observations each. 200 observations is about the size of most post-war US data sets. On each simulated data set, I estimate a VAR with 8 lags, and orthogonalize the innovations such that TFP growth is ordered first. For each simulation I compute impulse responses to news and surprise technology shocks, and then compare the distribution of estimated response to the true responses from the model. For the large sample experiment, I create one data set with 100,000 observations, estimate a VAR with 8 lags with TFP growth ordered first, and compare the responses to the true model responses.\(^9\)

These experiments require selecting parameter values for the model. Several parameters are fixed at levels across the different nested specifications. The unit of time is taken to be a quarter. The growth rate of TFP, \(g_a\), is set to 0.0025. This means that TFP grows by about one percent at an annual frequency. The Cobb-Douglas share parameter, \(\alpha\), is fixed at 1/3. With average productivity growth of one percent, output and its components will grow at 1.5 percent on average, which is broadly consistent with the post-war US per capita data. The subjective discount factor, \(\beta\), is 0.99, while the inflation target is set to \(\pi^* = 0.005\), or 2 percent at an annual frequency. This implies an average annualized nominal interest rate of \(i = 5.6\%\). The depreciation rate on capital is set to \(\delta = 0.02\). The parameter \(\xi\), which is the inverse Frisch labor supply elasticity, is fixed at one. There is substantial disagreement on the value of this parameter in the literature; many macro models need a high elasticity while most micro studies point to a low elasticity. The central estimate in Kimball and Shapiro (2010) is unity, which strikes a middle ground between the micro and macro literatures. The scaling parameter on the disutility of labor, \(\theta\), is always fixed such that steady state hours are 1/3 of the normalized time endowment of one. Because of the presence of the Taylor rule, the parameters governing the utility from holding real balances, \(\chi\) and \(\nu\), need not be calibrated. The standard deviations of the two shocks are fixed at \(\sigma_\varepsilon = 0.01\) and \(\sigma_u = 0.005\). These shock magnitudes are not chosen with any particular moments in mind, but they do imply that surprise technology shocks drive more of the unconditional variance of TFP growth than do news shocks (80 percent vs. 20 percent). This kind of calibration is necessary to produce data with similar co-movement among output and its components that is observed in actual data.\(^10\)

As we will see below, the surprise technology shocks being more important than news shocks tends to exacerbate any problems due to non-invertibility, and can therefore be considered relatively

---

\(^9\)The choice of 8 lags is somewhat arbitrary. As shown in Section 2, the mapping from model to data yields a VAR(\(\infty\)). Finite data samples require finite lag lengths, with \(p < \infty\). This introduces an additional source of bias, the so-called “lag truncation bias” emphasized, for example, in Chari, Kehoe, and McGrattan (2008). In practice \(p = 8\) lags appears to provide a sufficiently good approximation to the VAR(\(\infty\)), so that estimating with 8 lags isolates the bias due to non-invertibility. The results discussed below are qualitatively the same, though a little worse, with fewer than 8 lags (e.g. the popular \(p = 4\) specification with a year’s worth of lags).

\(^10\)Jaimovich and Rebelo (2009) provides a nice intuitive introduction for why news shocks tend to generate counterfactual co-movement among output and its components in many standard macro models.
conservative. Finally, the time lag between the revelation of news and its effect on productivity, $q$, is set to 4. This means that there is one year of anticipation.

The other parameters of the model govern the degree of frictions, and therefore the magnitudes of the departure from the simple real business cycle framework. When prices are sticky, $\phi$ is set to 0.7. This implies that the average duration between prices changes is between three and four quarters, which is broadly consistent with micro estimates (e.g. Bils and Klenow, 2004) and a number of macro estimates. The parameters of the Taylor rule are set to $\rho = 0.8$, $\psi_\pi = 1.5$, and $\psi_y = 0.5$. These are in line with standard calibrations and estimates within the literature. For real frictions, the habit formation parameter is set to $\gamma = 0.7$ and the investment adjustment cost parameter is $\tau = 2.5$. These are the central estimates in Christiano, Eichenbaum, and Evans (2005). Table 1 summarizes the parameter values for the different cases.

The principle conclusions of the ensuing Monte Carlo exercises can be summarized as follows. First, in spite of the presence of a non-invertibility due to the presence of foresight about productivity, well-specified VAR models can nevertheless do a very good job of recovering the underlying model’s responses to structural shocks. There are some differences in quality of fit across specifications, but in each model, VARs qualitatively capture the dynamics in response to both news and surprise technology shocks. Hence, impulse response analysis from VARs can be effectively used to compare different, nested model specifications. The magnitude of the biases resulting from non-invertibility depend on the relative importance of shocks in the underlying DSGE model – if the shock driving the non-invertibility (the news shock) is very important, then these biases are trivially small.

4.1 Full Model

In the fully parameterized model with both real and nominal frictions, the state vector is:

$$s_t = \begin{bmatrix} \ln \hat{y}_t & \ln \hat{k}_t & i_t & \ln v_t & z_{1,t} & z_{2,t} & z_{3,t} & \ln \hat{I}_t & \ln \hat{c}_t \end{bmatrix}'$$

TFP growth and normalized output are observed. Denote the vectors of observables as $x_t = [\ln \hat{a}_t \ln \hat{y}_t]'$. The shock vector is $\varepsilon_t = [e_t u_t]$. After solving the model at the parameter values given in Table 1, the “poor man’s invertibility condition” of Fernandez-Villaverde, Rubio-Ramirez, Sargent, and Watson (2007), and discussed in Section 2, can easily be checked. In the Appendix section A.1 the numeric values of the $A$, $B$, $C$, and $D$ matrixes are shown, as are the eigenvalues of the matrix $M = (A - BD^{-1}C)$. The maximum modulus of the eigenvalues of $M$ is 1.32. Hence, the invertibility condition fails.

Figure 1 plots impulse response obtained from Monte Carlo simulations of the “full model.” The solid lines are the theoretical responses in the model of the log levels of output and TFP to both news and surprise technology shocks.\footnote{Note that the level of technology does not directly show up in the state space. It is, however, implicitly part of the state in terms of the normalized variables; e.g. $\ln \hat{y}_t = \ln y_t - \frac{1}{1-\alpha} \ln a_t$.\footnote{The log levels are obtained by (i) cumulating the response of the growth rate of TFP to the shock and (ii) adding}} The left panel plots responses from the finite sample
simulations; the right panel shows results for the large sample simulation. In the finite sample panel, the dashed lines are the mean responses to the two shocks averaged over 500 different simulations of the model with 200 observations each. The shaded gray regions represent the middle 68 percent of the distribution of estimated responses across the 500 simulations. In the large sample panel, the dashed lines are the points estimates of the impulse responses obtained from one estimation of the VAR on a single data set with 100,000 observations. All VARs are estimated with \( p = 8 \) lags, and the innovations are orthogonalized with TFP growth ordered first.

A quick visual inspection reveals that the VARs do a good job of capturing the qualitative dynamics of the impulse responses to both kinds of technology shocks. Turning first to the finite sample results, the estimated responses to the news shock are essentially unbiased at horizons up to forty quarters – the average estimated responses roughly lie atop the true model responses. Further, the distributions of estimated responses are fairly tight and are centered around the true responses. The results are somewhat worse for the responses to the surprise technology shock, with some downward bias in the estimated responses, particularly at longer forecast horizons. Nevertheless, the high frequency responses are estimated quite well, and the longer horizon dynamics are qualitatively in line with the true model dynamics.

The right panel plots responses obtained in the “large sample” exercise. The dashed lines are very close to the true model responses at all horizons for both shocks. That they do not lie exactly on top of the true model responses is a direct consequence of the non-invertibility resulting because of the presence of foresight – non-invertibility means that, even in an infinitely large sample, a VAR cannot exactly recover the true impulse responses of the underlying model. What this exercise reveals is that the large sample bias is relatively small, and, for all practice purposes, likely of little importance. In either finite or large samples, the VAR does a very good job at recovering the model’s dynamics in response to both shocks.

Table 2 provides some quantitative evidence on the quality of fit of the estimated VARs. It shows the average absolute deviations of the impulse responses estimated in the simulations relative to the true model responses, at forecast horizons from impact to 40 quarters. For the large sample panel the numbers are simply the average deviation over 40 quarters of the estimated responses on one data set with 100,000 observations relative to the model responses; for the small sample panel the numbers are based on the deviation of the average estimated response across simulations. The numbers in the table are multiplied by 100, so they have the interpretation of average percentage point deviations. The average absolute deviations in the case of the news shock in the large simulation of the full model are 0.026 percentage points and 0.07 percentage points, for the TFP and output responses, respectively. These numbers are very small. For example, the TFP response after 4 quarters is 0.5 percent. On average the estimated response thus lies between 0.47 and 0.53 percent. The large sample biases in response to the surprise technology shock are somewhat larger, but nevertheless small. As one would expect, the finite sample biases tend to be larger, but are again very small and likely of little practical concern.

\[
\frac{1}{1-\alpha} \times \text{cumulated log level TFP response to the response of normalized output.}
\]
4.2 RBC Model

The RBC model is a special case of the full model with the parameter restrictions $\phi = \gamma = \tau = 0$. This turns out to substantially simplify the model, significantly reducing the state space. The state vector for this model is:

$$s_t = \begin{bmatrix} \ln \hat{k}_t & z_{1,t} & z_{2,t} & z_{3,t} \end{bmatrix}^\prime$$

Appendix A.2 shows the values of the parameters of the state space system when TFP growth and normalized output are observed by an econometrician. As in the full model case, the invertibility condition fails – the maximum modulus of the eigenvalues of $M = (A - BD^{-1}C)$ is 1.0572.

Figure 2 shows results for both the finite and large sample Monte Carlo exercises for the RBC model. One interesting thing to note is that the theoretical impulse responses of output to both news and surprise technology shocks differ a great deal relative to the full model with both real and nominal frictions. In particular, in the model output declines on impact in response to good news about productivity growth, while it rises by more than productivity in response to the surprise technology shock. These features are captured very well by the estimated VARs – in both finite and large samples, the impact effects in response to both kinds of technology shocks are estimated quite well. Examining Table 2, the biases in the estimated responses to the news shock are somewhat larger relative to the full model here; the reverse is true in response to the surprise technology shock. In both cases, the large sample biases, while present, are very small. In short, the VAR seems to do a good job in spite of non-invertibility.

In spite of its simplicity, an advantage of the RBC model is that it is very easy to see why the presence of news shocks leads to the invertibility problem. Suppose that there were only one shock in the model, the surprise technology shock, $e_t$. In this case, the state space of the model would simply be the transformed capital stock, $\ln \hat{k}_t$. Hence the matrixes $A$ and $B$ would just be scalars. Referring to Appendix A.2, they would take on values $A = 0.9554$ and $B = -1.4331$. The matrixes of the observer equation, with output observed, would also be scalars, equal to $C = 0.2210$ and $D = -0.3315$. Then $M = (A - BD^{-1}C) = 0.9554 - \frac{-1.4331 \times 0.2210}{-0.3315}$, which is equal to 0. Hence, the invertibility condition is satisfied. This means that a univariate autoregression of $\ln \hat{y}_t$ will correctly (in a large enough sample with enough lags) recover the impulse response to a technology shock as well as the time series of technology shocks. It is the presence of anticipation effects that drives the non-invertibility.\(^\text{13}\)

4.3 Sticky Price Model

The sticky price model imposes that $\gamma = \tau = 0$, so that the only friction relative to a simple RBC model is price stickiness. The state vector for this model is:

\(^\text{13}\)This statement requires some clarification; the impulse response estimated from the autoregression would correctly recover the impulse response of the normalized variable $\ln \hat{y}_t = \ln y_t - \frac{1}{\kappa_{\alpha}} \ln a_t$. The innovations in the autoregression correspond to the technology shocks after normalizing the variance of these shocks to some value (typically 1 in the applied literature).
\[ s_t = \begin{bmatrix} \ln \hat{y}_t & \ln \hat{k}_t & i_t & \ln v_t & z_{1,t} & z_{2,t} & z_{3,t} \end{bmatrix}' \]

The maximum modulus of the eigenvalues of \( M = (A - BD^{-1}C) \) is 1.2397, indicating a failure of the condition for invertibility. Figure 3 shows results from the Monte Carlo simulations. The model impulse responses to the news shock are quite similar to the RBC counterparts, whereas the response to the surprise technology shock are rather different, with output significantly undershooting on impact. In both finite and large samples these effects are captured quite well in the estimated VARs. In finite samples there is some downward bias at longer horizons in response to both shocks, but qualitatively the estimated responses line up well with the model's. The large sample results continue to exhibit some bias but this is again relatively small.

### 4.4 Real Frictions Model

Relative to the full model, the real frictions model imposes the parameter restriction that \( \phi = 0 \), so that prices are flexible. The state vector for the real frictions model is:

\[ s_t = \begin{bmatrix} \ln \hat{y}_t & \ln \hat{k}_t & i_t & \ln v_t & z_{1,t} & z_{2,t} & z_{3,t} & \ln \hat{c}_t & \ln \hat{I}_t \end{bmatrix}' \]

Appendix A.4 shows the numeric values of the \( A, B, C, \) and \( D \) matrixes of the state space. The maximum modulus of the eigenvalues of \( M = (A - BD^{-1}C) \) is 3.84, again indicating that the system is non-invertible.

Figure 4 shows the graphical results from Monte Carlo simulations with the real frictions model as the data generating process. The model responses to both news and surprise technology shocks are quite close to what obtains in the full model. Once again, both qualitatively and quantitatively, the estimated VARs do a very good job at capturing the model-implied dynamic responses to both kinds of shocks. Because of the non-invertibility, small biases in the estimated responses remain even in a very large sample size, but yet again these biases are quantitatively small. For example, looking at Table 2, one sees that the average deviation in the output response to a news shock is 0.09 percentage points in the large sample. Relative to the long horizon response of output of 0.7 percent, this deviation is quite small.

### 4.5 Using SVARs to Conduct Model Comparisons

One of the principle uses of structural VARs is to employ a common restriction that holds across different, potentially non-nested, models and to compare the qualitative pattern of impulse responses in the VAR to predictions from the different models. For example, many models predict that technology shocks should be the sole source of the unit root in labor productivity. But these models differ in terms of their predictions about the high frequency effects of technology shocks – RBC models, for example, predict that hours should increase following technological improvement, whereas some sticky price models, in which output is demand determined, predict the opposite. As a leading example, Gali (1999) identifies technology shocks in a VAR setting and shows that
productivity improvements lead to an immediate hours decline, leading him to conclude that sticky price, New Keynesian models are more promising than RBC models.\textsuperscript{14}

The different nested versions of the model presented above also make very different predictions about the responses of output and hours to both news and surprise technology shocks. In the RBC model, for example, output declines on impact in response to good news (so that hours decline as well), while output rises by more than TFP on impact following a surprise shock (so that hours increase). The full model and the flexible price model with real frictions have the exact opposite predictions: in those models hours and output increase on impact following good news, while hours decline on impact in response to a surprise technology shock. In the sticky price model hours decline on impact in response to both shocks, so that output falls following good news and rises, but by less than TFP, after the surprise technology shock.

Can estimated VARs do a good job of differentiating between these different models? The Monte Carlo results discussed above and graphically depicted in Figures 1 through 4 show a clear answer: yes. The small biases that are present are mostly at longer forecast horizons. In either small or large samples, the impact effects of both news and surprise technology shocks are estimated both accurately and precisely. For example, in the RBC and sticky price model simulations, the estimated output response declines immediately and only rises when TFP improves, just as in the models. In the full and real frictions models, output is estimated to rise on impact following a news shock and is estimated to rise by less than productivity after a surprise shock, again just as in the data. In short, a researcher hoping to use the impulse responses from an estimated VAR to qualitatively differentiate between these different models would likely be quite successful.

4.6 Non-Invertibility and the Relative Importance of Shocks

In the simulation results up to this point, the relative importance of the two stochastic shocks has been held fixed. While this may seem innocuous, the relative importance of the two kinds of shocks turns out to matter somewhat for the empirical performance of the VARs.

Figure 5 plots impulse responses obtained for the large sample Monte Carlo exercise for the full model under two different parameterizations of the standard deviation of the surprise technology shock: $\sigma_e = 0.01$ (which is the benchmark value) and $\sigma_e = 0.001$. The standard deviation of the news shock is fixed at its benchmark value, $\sigma_u = 0.005$. The solid line depicts the true model responses to the news shock, while the dashed and dotted lines, respectively, represent the responses obtained in sample sizes of 100,000 under the two different parameterizations of the magnitude of the surprise shock. In either case the estimation of the responses to the news shock is good, but it is visually apparent that the fit is significantly better when the news shock is relatively more important – i.e. when the standard deviation of the surprise technology shock is small. Quantitatively, the average absolute deviation of the TFP response to the news shock is 0.026 percent points in the

\textsuperscript{14}It should be pointed out that Gali’s (1999) results are disputed within the literature, and the sign of the hours response to a permanent technology shock appears to depend on whether hours enter the VAR in first differences or levels. The purpose of the present paper is not to dissect those results, but rather to use that paper as a motivating example for how VARs are used to make model comparisons.
large shock case and 0.008 percent points in the small shock case. For the response of output, these differences are 0.07 percent points and 0.01, respectively. In short, there is a large improvement in fit for the responses to the news shock when the surprise shock is less important in relative terms.

That the fit of the VARs improves as the news shock becomes relatively more important is easiest to understand by referencing back to (11): \( \Sigma_u = C \Sigma_s C' + D \Sigma_e D' \). The variance of the forecast of the state conditional on observables, \( \Sigma_s \), drives a wedge between the VAR innovations, \( u_t \), and the deep economic shocks, \( \varepsilon_t \). The missing state variables that account for the non-invertibility are lagged values of news shocks: \( u_{t-1}, \ldots, u_{t-q} \). It stands to reason that, as the relative importance of \( \varepsilon_t \) shocks declines, the observed variables will do a better job of revealing lagged values of the news shock.

Figure 6 plots the determinant of \( \Sigma_s \) (a proxy for “size”) as a function of the relative magnitude of the standard deviations of the two shocks. For this exercise, \( \sigma_u \) is held fixed at 0.005 and \( \sigma_e \) is varied. One observes that, as \( \sigma_u / \sigma_e \) gets large, \( \Sigma_s \) goes to zero and the wedge disappears. Feve and Jidoud (2011) make the same point analytically in a simpler environment with news shocks about productivity. Perhaps surprisingly, the fit of the estimated VARs is not much worse than the benchmark when news shocks are relatively unimportant, though the fit does improve fairly significantly as news shocks become relatively more important.

## 5 Adding Information

The Monte Carlo results of the previous section demonstrate that conventional VAR techniques may perform quite well, even in the face of a known non-invertibility. While these results may prove comforting to some, particularly those interested in the economic effects of news shocks about productivity, it is nevertheless not possible to use them to draw sweeping conclusions about the reliability of VAR techniques when the set of observables has a non-invertible VAR representation. Short of imposing the extra structure that full information techniques require, what can a researcher concerned about biases resulting from a potential non-invertibility do?

As the subtitle of this section suggests, adding information to the set of observed variables is the most straightforward route to go. Non-invertible representations arise when the observables fail to perfectly forecast the state vector of the DSGE model serving as the data generating process. Adding additional variables to the set of observables can only improve the forecast of the state, and thus reduce the magnitude of the wedge between VAR innovations and deep shocks.

In the context of the DSGE model considered so far, it is not possible to condition on more observables – one can only condition on as many observables as there are shocks when estimating the VAR. So as to circumvent this stochastic singularity issue, suppose that an econometrician observes a set of “information variables.” These are variables that convey information about the underlying state of the model, but are not otherwise part of the solution of the model. For example, these information variables could be stock prices, survey measures of consumer or business confidence, bond spreads, etc. In particular, I assume that these information variables are noisy signals about the news shock at time \( t \) – in other words, they are potentially useful in forecasting future
productivity conditional on current observed productivity. Let there be $Q$ of these variables, each obeying:

$$s_{i,t} = u_t + v_{i,t}, \quad i = 1, \ldots, Q$$  \hfill (34)

The error term $v_{i,t}$ represents the noise in each signal. These are i.i.d. across $i$, with mean zero and are drawn from a normal distribution with fixed variance. It is clear that, for $Q$ sufficiently large, conditioning on lags of $s_{i,t}$ in a VAR will perfectly reveal the missing states, since the $s_{i,t}$ will average out to the $u_t$ by application of a law of large numbers. This is a particularly simple informational structure, but could easily be extended on a number of different dimensions – for example, there could be persistence in the signals, the signals could respond to other economic shocks, the noise innovations in the signals could be correlated with one another or across time, etc.. The broader point is that conditioning on more information will reduce the variance in the forecast of the state vector, and therefore ought to improve the performance of estimated VARs.

To see this point clearly, Figure 7 shows some Monte Carlo results when incorporating these information series as additional variables in an otherwise standard VAR. Because it arguably has the worst overall fit in the simulations of the previous section, I consider the frictionless RBC model as the data generating process. So as to fix ideas, I focus here on the “large sample” exercise of simulating one data set with 100,000 observations. The standard deviation of the signals is set to 0.005. The solid line shows the true model responses to both shocks, while the dashed lines are the estimated responses from the conventional two variable VAR with $\ln\hat{a}_t$ and $\ln\hat{y}_t$. The thin dotted line shows the estimated responses from that same VAR augmented with four independent signals, while the thick dotted line shows the responses estimated when the VAR includes eight independent signals.\(^{15}\) The identifying restrictions are the same as above – the surprise shock is associated with the innovation in TFP growth ordered first, while the news shock is identified with the innovation in output ordered second. It is visually apparent that the fit improves as more of the information variables are added: the biases are smaller in the VAR with four signals than in the conventional two variable VAR, while the biases in the VAR with eight signals are smaller than the VAR with four signals. Extending this exercise to more information variables continues this pattern: the more information variables on which one conditions, the smaller are the large sample biases in the estimated impulse responses.\(^{16}\)

The above large sample exercise of simply adding more variables to the VAR is informative, but may not be of much practical interest when estimating VAR systems on relatively short sample sizes. Adding more than a couple of additional series in a sample of, say, 200, with more than a year’s worth of lags quickly becomes prohibitive. With different motivations and in different contexts, a number of researchers have made use of factor analytic methods.\(^{17}\) These methods

\(^{15}\)To be clear, the estimated VARs are then: $X_t = [\ln\hat{a}_t \quad \ln\hat{y}_t \quad s_{1,t} \ldots s_{r,t}]'$, for $r = 4$ or $r = 8$.

\(^{16}\)It is worth pointing out that the standard deviation of the noise innovations, $\sigma_v$, does have an effect on these conclusions. If the signals are very precise, then one does not need to add many signals to the VAR for the large sample biases to vanish. In contrast, noisier signals necessitate including more information variables in order to reduce the large sample biases.

\(^{17}\)For applications in macroeconomics, see, for example Bernanke, Boivin, and Eliasz (2005) and Stock and Watson
make use of principal components to compress large sets of data into a small number of common components. This effectively allows one to condition on a large amount of information without consuming too many degrees of freedom.

As an alternative to estimating a VAR system with many additional variables, consider estimating the following factor augmented VAR: $X_t = [\ln \hat{a}_t \ \ln \hat{y}_t \ F_t]'$, where $F_t$ is the first principal component of $Q$ information variables. Isolating just the first principle component in this context makes sense as the information variables only have one common component – the news shock. With enough information variables, conditioning on the common component will be equivalent to conditioning on current and lagged news shocks, and hence the system ought to be invertible.

Figure 8 shows impulse responses obtained from both finite and large sample Monte Carlo exercises making use of the first principle component of $Q$ = 30 factors (the number 30 is arbitrary; the important point is that it be “large”). I take the frictionless RBC model as the data generating process. The standard deviation of the noise innovations in the information variables is again set to 0.005. The estimation procedure takes place in two steps. In the first step, the first principle component of the $Q$ information series is obtained, $F_t$. In the second step, a conventional unrestricted VAR is estimated with $\ln \hat{a}_t$, $\ln \hat{y}_t$, and $F_t$. The identifying restrictions are as above – the surprise technology shock is identified with the innovation in TFP growth, while the news shock is identified as the output innovation ordered second.

Turning first to the large sample results, one observes that the large sample biases have essentially disappeared. The estimated impulse responses virtually lie atop the true model responses at all horizons. That any biases remain is simply a function of $Q$ being small – conditioning on sufficiently more variables in the first stage would cause any remaining biases to vanish entirely. The results are also substantially better in finite samples. Here there remains some downward-bias in the responses at longer horizons, but this is primarily due to finite sample bias in autoregressive coefficients, and is unrelated to non-invertibility. The estimated responses are essentially unbiased on impact and for a number of quarters thereafter. Comparing the finite sample results in Figure 8 with Figure 2, there is a clear improvement. Similar results obtain for simulations from the other version of the model with frictions.

These simulations results suggest that a sensible way of dealing with a potential non-invertibility is to estimate a factor-augmented VAR, with the series used to construct the factor explicitly chosen with the goal in mind of forecasting unobserved states. Indeed, this is consistent with the recommendations in Giannone and Reichlin (2006), Forni, Giannone, Lippi, and Reichlin (2009), and Forni, Gambetti, and Sala (2011). Since non-invertibility is fundamentally a missing information problem, factor methods, which allow an econometrician to condition on a very large data set, are an appropriate, flexible, and simple way to deal with the issue, short of resorting to full information methods.
6 Concluding Thoughts

Fernandez-Villaverde, Rubio-Ramirez, Sargent, and Watson (2007) recommend that researchers estimate the structural parameters of detailed models using full information techniques. If one believes in a particular DSGE specification, then this is sound advice – from an efficiency perspective one would never want to recover impulse response from a VAR if one strongly believed in the underlying DSGE model. The advantage of VARs and similar limited information techniques is that they do not impose as much structure as full information methods, and are therefore less subject to specification bias and can be used to make cross-model comparisons. However, the mapping between DSGE models and VARs is not always clean. The so-called non-invertibility problem arises when the VAR on a set of observables cannot be mapped back into the structural form of an economic model. This means that analysis based on VARs may not prove very useful in building and refining fully specified DSGE models.

This paper has focused on the issue of non-invertibility within the context of a particular shock structure known to create problems – so-called “news shocks” which generate foresight about exogenous changes in future productivity. In so doing, it has made two primary contributions to the literature. First, it has emphasized that invertibility is not an “either/or” proposition – a particular model may be technically non-invertible but the resulting biases in estimated impulse responses may be very small. Second, non-invertibility is best understood as a problem of missing information – therefore, the most straightforward way to deal with it while remaining within the scope of limited information methods is to condition on more information. In particular, estimating VARs which condition on more information – either through adding additional variables informative about the missing states directly or through factor augmented VARs – works to eliminate the biases due to non-invertibility. These methods are relatively straightforward to implement and do not require imposing the structure that full information estimation methods require.
References


<table>
<thead>
<tr>
<th>Parameter</th>
<th>Full Model</th>
<th>RBC Model</th>
<th>Sticky Price Model</th>
<th>Real Frictions Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.99</td>
<td>$\gamma = 0$</td>
<td>$\gamma = 0$</td>
<td>$\phi = 0$</td>
</tr>
<tr>
<td>$\delta$</td>
<td>0.02</td>
<td>$\tau = 0$</td>
<td>$\tau = 0$</td>
<td></td>
</tr>
<tr>
<td>$g_a$</td>
<td>0.0025</td>
<td>$\phi = 0$</td>
<td>$\phi = 0$</td>
<td></td>
</tr>
<tr>
<td>$\pi^*$</td>
<td>0.005</td>
<td></td>
<td>$\phi = 0$</td>
<td></td>
</tr>
<tr>
<td>$\alpha$</td>
<td>$\frac{1}{3}$</td>
<td></td>
<td>$\phi = 0$</td>
<td></td>
</tr>
<tr>
<td>$\xi$</td>
<td>1</td>
<td></td>
<td>$\phi = 0$</td>
<td></td>
</tr>
<tr>
<td>$\sigma_e$</td>
<td>0.01</td>
<td></td>
<td>$\phi = 0$</td>
<td></td>
</tr>
<tr>
<td>$\sigma_u$</td>
<td>0.005</td>
<td></td>
<td>$\phi = 0$</td>
<td></td>
</tr>
<tr>
<td>$\theta \Rightarrow n^* = \frac{1}{3}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\tau$</td>
<td>2.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\phi$</td>
<td>0.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\rho$</td>
<td>0.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi_x$</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\psi_y$</td>
<td>1.5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The left column, labeled “Full Model,” lists the parameter values used in the fully parameterized specification. The remaining columns list the parameter restrictions relative to the full model.
Table 2: Mean Absolute Deviations: VAR Monte Carlo Exercises

<table>
<thead>
<tr>
<th></th>
<th>Tech. to News</th>
<th>Output to News</th>
<th>Tech. to Surprise</th>
<th>Output to Surprise</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Large Sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full</td>
<td>0.026</td>
<td>0.070</td>
<td>0.070</td>
<td>0.087</td>
</tr>
<tr>
<td>RBC</td>
<td>0.068</td>
<td>0.100</td>
<td>0.037</td>
<td>0.060</td>
</tr>
<tr>
<td>Sticky Price</td>
<td>0.106</td>
<td>0.146</td>
<td>0.055</td>
<td>0.081</td>
</tr>
<tr>
<td>Real Fric.</td>
<td>0.048</td>
<td>0.089</td>
<td>0.052</td>
<td>0.071</td>
</tr>
<tr>
<td><strong>Small Sample</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full</td>
<td>0.047</td>
<td>0.052</td>
<td>0.283</td>
<td>0.354</td>
</tr>
<tr>
<td>RBC</td>
<td>0.178</td>
<td>0.244</td>
<td>0.144</td>
<td>0.176</td>
</tr>
<tr>
<td>Sticky Price</td>
<td>0.203</td>
<td>0.267</td>
<td>0.128</td>
<td>0.152</td>
</tr>
<tr>
<td>Real Fric.</td>
<td>0.070</td>
<td>0.076</td>
<td>0.267</td>
<td>0.343</td>
</tr>
</tbody>
</table>

These numbers represent the average absolute deviations of the estimated impulse responses relative to the true responses in the models, times 100, for the various different Monte Carlo exercises. In the “Large Sample” block the numbers are based on one sample with 100,000 observations relative to the true model responses. In the “Small Sample” block the numbers are based on the difference between the average estimated responses across 500 simulations relative to the true model responses.
The solid lines are the theoretical impulse responses to news and surprise technology shocks in the “Full Model” using the parameterization as described in the text. For the Monte Carlo exercises the VARs feature \( \ln \hat{a}_t \) and \( \ln \hat{y}_t \) and are estimated with \( p = 8 \) lags. In the left panel, labeled “Finite Sample,” the dashed lines are the mean responses averaged across 500 simulations of data sets with 200 observations each, while the shaded gray regions depict the middle 68 percent of the distribution of estimated responses across the 500 simulations. In the right panel, labeled “Large Sample,” the dashed lines are the estimated impulse responses from the estimation on one sample with 100,000 observations.
The solid lines are the theoretical impulse responses to news and surprise technology shocks in the “RBC Model” using the parameterization as described in the text. For the Monte Carlo exercises the VARs feature ln $\hat{a}_t$ and ln $\hat{y}_t$ and are estimated with $p = 8$ lags. In the left panel, labeled “Finite Sample,” the dashed lines are the mean responses averaged across 500 simulations of data sets with 200 observations each, while the shaded gray regions depict the middle 68 percent of the distribution of estimated responses across the 500 simulations. In the right panel, labeled “Large Sample,” the dashed lines are the estimated impulse responses from the estimation on one sample with 100,000 observations.
The solid lines are the theoretical impulse responses to news and surprise technology shocks in the “Sticky Price Model” using the parameterization as described in the text. For the Monte Carlo exercises the VARs feature $\ln \hat{a}_t$ and $\ln \hat{y}_t$ and are estimated with $p = 8$ lags. In the left panel, labeled “Finite Sample,” the dashed lines are the mean responses averaged across 500 simulations of data sets with 200 observations each, while the shaded gray regions depict the middle 68 percent of the distribution of estimated responses across the 500 simulations. In the right panel, labeled “Large Sample,” the dashed lines are the estimated impulse responses from the estimation on one sample with 100,000 observations.
The solid lines are the theoretical impulse responses to news and surprise technology shocks in the “Real Frictions Model” using the parameterization as described in the text. For the Monte Carlo exercises the VARs feature $\ln \hat{\alpha}_t$ and $\ln \hat{\gamma}_t$ and are estimated with $p = 8$ lags. In the left panel, labeled “Finite Sample,” the dashed lines are the mean responses averaged across 500 simulations of data sets with 200 observations each, while the shaded gray regions depict the middle 68 percent of the distribution of estimated responses across the 500 simulations. In the right panel, labeled “Large Sample,” the dashed lines are the estimated impulse responses from the estimation on one sample with 100,000 observations.
The solid lines are the theoretical impulse responses to a news shock in the “Full Model” using the parameterization as described in the text. The dashed and dotted lines represent the estimated responses from a large sample Monte Carlo exercise using two different values for the standard deviation of surprise technology shocks: $\sigma_e = 0.01$ (dashed line, which also corresponds to the baseline used in previous figures) and $\sigma_e = 0.001$ (dotted line). 100,000 observations are simulated from the “Full Model” and then a VAR is estimated on a system featuring TFP growth and normalized output, with the news shock identified as the orthogonalized output innovation. The VAR uses $p = 8$ lags.
Figure 6: Theoretical Wedge and Varying Shock Magnitudes

This figure plots the determinant of $\Sigma_s$, the variance-covariance matrix of period $t$ optimal forecast of the state, as a function of the relative shock magnitudes between surprise, $e_t$, and news shocks, $u_t$. This is done in the context of the “Full Model” conditional on observing TFP growth and normalized output. The vertical axis plots the ratio of the the standard deviation of $u_t$ divided by the standard deviation of $e_t$ against the determinant of $\Sigma_s$, which is obtained numerically from solving the Ricatti equations (6)-(7).
The solid lines are the impulse responses to both news and surprise technology shocks in the RBC model. For the Monte Carlo simulations, 100,000 observations are simulated from the model, with additional noisy “information variables” about future productivity included in the model. The dashed, small dotted, and wide dotted lines show the estimated responses with no signals included in the VAR, with four signals, and with eight signals, respectively.
The solid lines are the theoretical responses to news and surprise technology shocks in the RBC model. For the Monte Carlo exercises the estimated VARs include $\ln \tilde{a}_t$, $\ln \tilde{y}_t$, and the first principle component of thirty noisy signals about future productivity. The VARs are estimated with $p = 8$ lags. In the left panel, labeled “Finite Sample,” the dashed lines are the mean responses averaged across 500 simulations of data sets with 200 observations each, while the shaded gray regions depict the middle 68 percent of the distribution of estimated responses across the 500 simulations. In the right panel, labeled “Large Sample,” the dashed lines are the estimated impulse responses from the estimation on one sample with 100,000 observations.
A State Space Representation of the Various Models

This appendix provides details on the parameters of the state space representations of the various nested models.

A.1 The Full Model

The state vector is:

$$s_t = \begin{bmatrix} \ln \hat{y}_t & \ln k_t & i_t & \ln v_t & z_{1,t} & z_{2,t} & z_{3,t} & \ln \hat{I}_t & \ln \hat{c}_t \end{bmatrix}'$$

The vector of observables is:

$$x_t = \begin{bmatrix} \ln \hat{a}_t & \ln \hat{c}_t \end{bmatrix}'$$

These variables represent logarithmic deviations from their normalized steady state values. The shock vector is $$\varepsilon_t = [e_t \ u_t]'$$. The parameters of the state space representation are:

$$A = \begin{bmatrix} 0.0552 & 0.0746 & -0.4402 & -0.0909 & -1.1461 & 0.3006 & 0.2305 & 0.1544 & 0.4798 \\ 0.0034 & 0.9723 & -0.0267 & -0.0049 & -1.4857 & 0.0111 & 0.0070 & 0.0185 & -0.0037 \\ -0.0603 & -0.0278 & 0.4806 & 0.0424 & -0.0452 & 0.0096 & 0.0407 & 0.0307 & 0.0875 \\ 0.0175 & -0.0179 & -0.1398 & 0.7656 & -0.0410 & -0.0105 & 0.0089 & 0.0077 & 0.0200 \\ 0.0000 & 0.0000 & -0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 0.1417 & -0.1717 & -1.1291 & -0.2051 & -0.8958 & 0.4691 & 0.2960 & 0.7831 & -0.1558 \\ 0.0319 & 0.1411 & -0.2543 & -0.0600 & -1.2136 & 0.2552 & 0.2129 & -0.0153 & 0.6513 \end{bmatrix}$$

$$B = \begin{bmatrix} -1.1461 & 0.1603 \\ -1.4857 & 0.0029 \\ -0.0452 & 0.0551 \\ -0.0410 & 0.0198 \\ 0.0000 & 0.0000 \\ 0.0000 & 0.0000 \\ 0.0000 & 1.0000 \\ -0.8958 & 0.1206 \\ -1.2136 & 0.1711 \end{bmatrix}$$

$$C = \begin{bmatrix} 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0552 & 0.0746 & -0.4402 & -0.0909 & -1.1461 & 0.3006 & 0.2305 & 0.1544 & 0.4798 \end{bmatrix}$$

$$D = \begin{bmatrix} 1.0000 & 0.0000 \\ -1.1461 & 0.1603 \end{bmatrix}$$

The modulus of the eigenvalues of the matrix $$M = (A - BD^{-1}C)$$ are:
There are two eigenvalues with modulus outside the unit circle. Hence, the “poor man’s invertibility condition” is not satisfied.

A.2 The RBC Model

The state vector is:

\[ s_t = \begin{bmatrix} \ln \hat{k}_t & z_{1,t} & z_{2,t} & z_{3,t} \end{bmatrix}' \]

The vector of observables is the same as above.

The parameters of the state space representation are:

\[ A = \begin{bmatrix} 0.9554 & -1.4331 & -0.0674 & -0.0704 \\ 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix} \]

\[ B = \begin{bmatrix} -1.4331 & -0.0666 \\ 0.0000 & 0.0000 \\ 0.0000 & 0.0000 \\ 0.0000 & 1.0000 \end{bmatrix} \]

\[ C = \begin{bmatrix} 0.0000 & 1.0000 & 0.0000 & 0.0000 \\ 0.2210 & -0.3315 & -0.2903 & -0.2745 \end{bmatrix} \]

\[ D = \begin{bmatrix} 1.0000 & 0.0000 \\ -0.3315 & -0.2597 \end{bmatrix} \]

The modulus of the eigenvalues of the matrixes \( M = (A - BD^{-1}C) \) are:

\[ \Lambda = \begin{bmatrix} 1.3208 \\ 1.3208 \\ 0.5781 \\ 0.9471 \\ 0.7239 \\ 0.8057 \\ 0.0000 \\ 0.0000 \end{bmatrix} \]

Again, the conditions required for invertibility are not met.
A.3 The Sticky Price Model

The state vector is:

\[ s_t = \begin{bmatrix} \ln \tilde{y}_t & \ln \tilde{k}_t & i_t & v_{t,t} & z_{2,t} & z_{3,t} \end{bmatrix}' \]

The vector of observables is as above. The parameters of the state space representation are:

\[
A = \begin{bmatrix} 1.1043 & -0.4590 & -8.8015 & -0.2184 & -0.9680 & -0.4384 & -0.3754 \\ 0.1192 & 0.8796 & -0.9498 & -0.0143 & -1.4982 & -0.1024 & -0.0921 \\ 0.0996 & -0.1080 & -0.7942 & 0.0036 & 0.0125 & -0.0737 & -0.0590 \\ 0.0453 & -0.0314 & -0.3611 & 0.7524 & -0.0208 & -0.0151 & -0.0108 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}
\]

\[
B = \begin{bmatrix} -0.9680 & -0.3095 \\ -1.4982 & -0.0818 \\ 0.0125 & -0.0455 \\ -0.0208 & -0.0073 \\ 0.0000 & 0.0000 \\ 0.0000 & 0.0000 \\ 0.0000 & 1.0000 \end{bmatrix}
\]

\[
C = \begin{bmatrix} 0.0000 & 0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 \\ 1.1043 & -0.4590 & -8.8015 & -0.2184 & -0.9680 & -0.4384 & -0.3754 \end{bmatrix}
\]

\[
D = \begin{bmatrix} 1.0000 & 0.0000 \\ -0.9680 & -0.3095 \end{bmatrix}
\]

The modulus of the eigenvalues of the matrixes \( M = (A - BD^{-1}C) \) are:

\[
\Lambda = \begin{bmatrix} 1.2397 \\ 1.2397 \\ 0.8087 \\ 0.8087 \\ 0.7638 \\ 0.0000 \\ 0 \end{bmatrix}
\]

The conditions for invertibility are not satisfied.

A.4 The Real Frictions Model

The state vector is:

\[ s_t = \begin{bmatrix} \ln \tilde{y}_t & \ln \tilde{k}_t & z_{1,t} & z_{2,t} & z_{3,t} & \ln \tilde{c}_t & \ln \tilde{l}_t \end{bmatrix}' \]
The vector of observables is as above. The parameters of the state space representation are:

\[
A = \begin{bmatrix}
0.9757 & -1.4747 & 0.0092 & 0.0002 & 0.0166 & -0.0092 \\
0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 \\
0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
-0.0281 & -0.4298 & 0.3895 & 0.0087 & 0.7030 & -0.3884 \\
0.1743 & -1.0942 & 0.2098 & 0.1182 & -0.0353 & 0.5905
\end{bmatrix}
\]

\[
B = \begin{bmatrix}
-1.4747 & -0.0049 \\
0.0000 & 0.0000 \\
0.0000 & 0.0000 \\
0.0000 & 1.0000 \\
-0.4298 & -0.2091 \\
-1.0942 & 0.0773
\end{bmatrix}
\]

\[
C = \begin{bmatrix}
0.0000 & 1.0000 & 0.0000 & 0.0000 & 0.0000 & 0.0000 \\
0.1313 & -0.9531 & 0.2480 & 0.0949 & 0.1215 & 0.3826
\end{bmatrix}
\]

\[
D = \begin{bmatrix}
1.0000 & 0.0000 \\
-0.9531 & 0.0165
\end{bmatrix}
\]

The modulus of the eigenvalues of the matrixes \( M = (A - BD^{-1}C) \) are:

\[
\Lambda = \begin{bmatrix}
3.8376 \\
3.8376 \\
0.9620 \\
0.8083 \\
0.0000 \\
0.0000
\end{bmatrix}
\]

Since the maximum modulus of the eigenvalues lies outside of the unit circle, the conditions for invertibility are not satisfied.