News, Non-Invertibility, and Structural VARs*

Eric R. Sims
University of Notre Dame, NBER, and ifo Institute

June 14, 2012

Abstract

A state space representation of a linearized DSGE model implies a VAR in terms of observable variables. The model is said be non-invertible if there exists no linear rotation of the VAR innovations which can recover the 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 principal 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 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 Barsky as well as two anonymous referees for helpful comments.

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 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 observed variables. The model is said to be invertible if there exists, in population, a linear rotation of the VAR innovations which recovers 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 of a fully-specified DSGE model. The non-invertibility problem is fundamentally one of missing information. As shown in Section 2, 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 amounts to measurement error from forecasting the state 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 principal 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 non-invertible VAR representation but where structural VARs nevertheless perform reliably. I make these arguments on the basis of analysis of a conventional, medium scale DSGE model

[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 (2012); 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 early treatments of this issue, see Quah (1990), Hansen and Sargent (1991), Hansen, Roberds, and Sargent (1991), and Lippi and Riechlin (1994). A well-known recent treatment of this problem is found in Fernandez-Villaverde, Rubio-Ramirez, Sargent, and Watson (2007).

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 (2011) study the role of news about future productivity, while Mountford and Uhlig (2009) and Forni and Gambetti (2011) 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 may drive 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.

Section 3 lays out a DSGE model that serves as a laboratory for investigating the significance of non-invertibility for VARs. 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 baseline 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, as long as there is more than one period of anticipation, generates non-invertible VAR representations when TFP growth and any other variable of the model are observed.[4] I then conduct a battery of Monte Carlo exercises in which I examine the performance of apparently well-specified SVARs. Although the model has a number of frictions, the relatively simple shock structure lends itself to estimating a small VAR system and using a conventional recursive identifying assumption. 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

[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.

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 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.[5]

In spite of the presence of a known non-invertibility, SVARs applied to model simulated data perform 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, 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 up these features in the simulations. This means that the VARs can be an effective tool at discriminating between different nested versions of a model.

The simplicity of the assumed shock structure in the DSGE model lends itself to estimating two variable VARs with a recursive restriction. This is nice in that it allows one to cleanly focus in on the role of non-invertibility, but it is rather unrealistic and therefore may lack wider appeal. In Section 5 I consider a number of robustness checks in richer model environments. In particular, I consider situations in which TFP is not observed by the econometrician, in which there are additional "demand shocks," and in which an econometrician mistakenly introduces other sources of bias into the estimated VAR. Among other things, these modifications of the model structure invalidate the simple recursive identifying assumption that is used throughout Section 4. Using long run restrictions, shape/sign restrictions, and combinations of short run, long run, and shape restrictions I conduct additional Monte Carlo exercises on VAR systems featuring between two and four variables. In all cases the identified VAR impulse responses are good approximations to the true model impulse responses to both news and surprise technology shocks.

In summary, the simulation results of Sections 4 and 5 suggest 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. If one wants to use a VAR in a situation in which non-invertibility may arise, the onus is on that researcher to convince his or her readers that invert-

[5] 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 (2011). 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.

ibility is not a major problem in that particular context. I therefore consider in Section 6 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. Non-invertibility is fundamentally a problem of missing information; hence, adding more information is the most straightforward way to deal with it.

In Section 6 I consider conditioning on additional "information variables." 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 the VAR systems considered in Section 4 works to reduce the small biases in the estimated impulse responses to both news and surprise technology shocks. One quickly runs into a sort of "curse of dimensionality" 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. Recent papers such as Giannone and Reichlin (2006), Forni, Giannone, Lippi, and Reichlin (2009), and Forni, Gambetti, and Sala (2011) have also advocated for the use of factor-augmented VARs to overcome problems due to non-invertibility, the last paper specifically in the context of news shocks. Factor-augmented VAR methods have appeal in that they are straightforward to implement, while maintaining the relative lack of structure that full information techniques require.[6]

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 conducts additional Monte Carlo exercises in richer environments in which a simple recursive identification is not available. Section 6 considers conditioning on more information as a way to overcome invertibility issues. The final section offers concluding thoughts.

2 The Mapping Between DSGE Models and VARs

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

[6]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 a non-invertibility problem, whereas conditioning on a very large information set can.

$$s_t = As_{t-1} + B\epsilon_t \tag{1}$$

$$x_t = Cs_{t-1} + D\epsilon_t \tag{2}$$

s_t is $k \times 1$ vector of state variables, x_t is a $n \times 1$ vector of observed variables, and ϵ_t is a $m \times 1$ vector of structural shocks. The variance-covariance matrix of these shocks is diagonal and given by Σ_ϵ. 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 ϵ_t from (2) as:

$$\epsilon_t = D^{-1}\left(x_t - Cs_{t-1}\right)$$

Plugging this into (1) yields:

$$s_t = \left(A - BD^{-1}C\right)s_{t-1} + BD^{-1}x_t$$

Solving backwards, one obtains:

$$s_t = \left(A - BD^{-1}C\right)^{t-1} s_0 + \sum_{j=0}^{t-1} \left(A - BD^{-1}C\right)^{j-1} BD^{-1}x_{t-j} \tag{3}$$

If $\lim_{t\to\infty} \left(A - BD^{-1}C\right)^{t-1} = 0$, then the history of observables perfectly reveals the current state. This requires that the eigenvalues of $\left(A - BD^{-1}C\right)$ 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:

$$x_t = C\sum_{j=0}^{t-1} \left(A - BD^{-1}C\right)^{j-1} BD^{-1}x_{t-1-j} + D\epsilon_t \tag{4}$$

The condition that the eigenvalues of $\left(A - BD^{-1}C\right)$ 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, $\widehat{s}_t$, given observables and a lagged forecast:

$$\widehat{s}_t = (A - KC)\,\widehat{s}_{t-1} + Kx_t \tag{5}$$

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

$$\Sigma_s = (A - KC)\,\Sigma_s\,(A - KC)' + B\Sigma_\epsilon B' + KD\Sigma_\epsilon D'K' - B\Sigma_\epsilon D'K' - KD\Sigma_\epsilon B' \tag{6}$$

$$K = \left(A\Sigma_s C' + B\Sigma_\epsilon D'\right)\left(C\Sigma_s C' + D\Sigma_\epsilon D'\right)^{-1} \tag{7}$$

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

$$x_t = C\widehat{s}_{t-1} + u_t \tag{8}$$

$$u_t = C\,(s_{t-1} - \widehat{s}_{t-1}) + D\epsilon_t \tag{9}$$

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

$$x_t = (A - KC)^{t-1}\widehat{s}_0 + C\sum_{j=0}^{t-1}(A - KC)^j\,Kx_{t-1-j} + u_t \tag{10}$$

Under weak conditions, Hansen and Sargent (2007) show that $(A - KC)$ is a stable matrix, so that the $(A - KC)^{t-1}\,\widehat{s}_0$ 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:

$$\Sigma_u = C\Sigma_s C' + D\Sigma_\epsilon D' \tag{11}$$

Fernandez-Villaverde, Rubio-Ramirez, Sargent and Watson (2007) show that the eigenvalues of $\left(A - BD^{-1}C\right)$ 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 Σ_s – i.e. how hidden the state is. This has a number of implications. First, even if the condition for invertibility fails, Σ_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 Σ_s and a closer mapping between VAR innovations and structural shocks. Finally, adding more observable variables should always work to lower Σ_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 6 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:

$$\max_{c_t, n_t, I_t, k_{t+1}, M_{t+1}, B_{t+1}} \quad E_0 \sum_{t=0}^{\infty} \beta^t \left(\ln\left(c_t - \gamma c_{t-1}\right) - \theta \frac{n_t^{1+\xi}}{1+\xi} + \chi \frac{\left(\frac{M_{t+1}}{p_t}\right)^{1-\nu}}{1-\nu} \right)$$

s.t.

$$c_t + I_t + \frac{B_{t+1} - B_t}{p_t} + \frac{M_{t+1} - M_t}{p_t} \leq w_t n_t + R_t k_t + i_{t-1} \frac{B_t}{p_t} + \frac{\Pi_t}{p_t} + \frac{T_t}{p_t}$$

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

γ governs the degree of internal habit formation in consumption, ξ is the inverse Frisch labor supply elasticity, and ν 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, Π_t is nominal profits distributed lump sum from firms, and T_t is nominal lump sum tax/transfers from the government. τ is a parameter governing the cost of adjusting investment, with Δ_I the (gross) balanced growth path growth rate of investment.

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

$$\lambda_t = \frac{1}{c_t - \gamma c_{t-1}} - \beta\gamma E_t \frac{1}{c_{t+1} - \gamma c_t} \tag{12}$$

$$\theta n_t^{\xi} = \lambda_t w_t \tag{13}$$

$$\lambda_t = \beta E_t \lambda_{t+1}(1+i_t)\frac{p_t}{p_{t+1}} \tag{14}$$

$$\mu_t = \beta E_t \left(\lambda_{t+1} R_{t+1} + (1-\delta)\mu_{t+1}\right) \tag{15}$$

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

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

λ_t is the current value Lagrange multiplier on the flow budget constraint and μ_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:

$$y_t = \left(\int_0^1 y_{j,t}^{\frac{\epsilon-1}{\epsilon}} dj\right)^{\frac{\epsilon}{\epsilon-1}}$$

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

$$y_{j,t} = \left(\frac{p_{j,t}}{p_t}\right)^{-\epsilon} y_t \tag{18}$$

$$p_t = \left(\int_0^1 p_{j,t}^{1-\epsilon} dj\right)^{\frac{1}{1-\epsilon}} \tag{19}$$

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

$$y_{j,t} = a_t k_{j,t}^{\alpha} n_{j,t}^{1-\alpha} \tag{20}$$

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:

$$\min_{n_{j,t}, k_{j,t}} \quad w_t n_{j,t} + R_t k_{j,t}$$

$$\text{s.t.}$$

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

The first order conditions are:

$$w_t = mc_{j,t}(1-\alpha) a_t \left(\frac{k_{j,t}}{n_{j,t}} \right)^{\alpha} \tag{21}$$

$$R_t = mc_{j,t} \alpha a_t \left(\frac{k_{j,t}}{n_{j,t}} \right)^{\alpha-1} \tag{22}$$

$mc_{j,t}$, the multiplier on the production constraint, and 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 maximizes the expected present discounted value of future profits, where profits are discounted by both the stochastic discount factor measured in utils, $\Lambda_{t+s} = \beta^s \lambda_{t+s}$, and the probability that a price chosen today is still in effect in the future, ϕ^s. The problem of a firm with the ability to update in date t is:

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

The solution is an optimal reset price, $p_t^{\#}$, satisfying:

$$p_t^{\#} = \frac{\epsilon}{\epsilon - 1} \frac{\sum_{s=0}^{\infty} (\phi\beta)^s \left(mc_{t+s} p_{t+s}^{\epsilon} y_{t+s} \right)}{\sum_{s=0}^{\infty} (\phi\beta)^s \left(p_{t+s}^{\epsilon-1} y_{t+s} \right)}$$

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

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:

$$i_t = \rho i_{t-1} + (1-\rho)\psi_\pi (\pi_t - \pi^*) + (1-\rho)\psi_y \left(\frac{y_t}{y_{t-1}} - \Delta_y \right) \tag{23}$$

$0 \leq \rho \leq 1$ is an interest rate smoothing parameter, π^* is an exogenous inflation target, and ψ_π and ψ_y are response coefficients to inflation and the output "growth gap," where Δ_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:

$$T_t = M_{t+1} - M_t \tag{24}$$

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:

$$\ln a_t = g_a + \ln a_{t-1} + \varepsilon_t + u_{t-q} \tag{25}$$

$$\varepsilon_t \sim N(0, \sigma_e)$$
$$u_t \sim N(0, \sigma_u)$$

ε_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 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:

$$1 + \pi_t = \left((1 - \phi) \left(1 + \pi_t^{\#} \right)^{1-\epsilon} + \phi \right)^{\frac{1}{1-\epsilon}} \tag{26}$$

Here $1 + \pi_t^{\#} = \frac{p_t^{\#}}{p_{t-1}}$. Aggregation of the intermediate firm production functions yields:

$$y_t = \frac{a_t k_t^{\alpha} n_t^{1-\alpha}}{v_t} \tag{27}$$

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:

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

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

$$y_t = c_t + I_t \tag{29}$$

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:

$$\widehat{a}_t = \frac{a_t}{a_{t-1}}, \ \widehat{c}_t = \frac{c_t}{a_t^{\frac{1}{1-\alpha}}}, \ \widehat{y}_t = \frac{y_t}{a_t^{\frac{1}{1-\alpha}}}, \ \widehat{I}_t = \frac{c_t}{a_t^{\frac{1}{1-\alpha}}}, \ \widehat{k}_t = \frac{k_t}{a_{t-1}^{\frac{1}{1-\alpha}}}$$

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." 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. There will be q additional states for q periods of anticipation. Define these as $z_{i,t}$ for $i = 1, \ldots q$. Doing so allows one to write a process for technology that satisfies the Markov property:

$$\ln a_t = g_a + \ln a_{t-1} + \varepsilon_t + z_{1,t-1} \tag{30}$$

$$z_{1,t} = z_{2,t-1} \tag{31}$$

$$z_{2,t} = z_{3,t-1} \tag{32}$$

$$\vdots$$

$$z_{q,t} = u_t \tag{33}$$

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

Some authors, most notably Blanchard, L'Hullier, and Lorenzoni (2011), have studied environments in which agents observe noisy signals about the underlying state of the economy and must solve a signal extraction problem. In a sense, the agents in the economy presented in this section also observe noisy signals about the future state of the economy, though the problem is considerably simpler. While the news shock does generate perfect foresight about future z, it does not lead to perfect foresight about the level of future $\ln a_t$. This is because of the presence of the surprise technology shock, ε_t. Agents could observe a positive news shock today, u_t, that does not materialize into higher $\ln a_{t+q}$ because of a low realization of ε at any of the dates in between, $t+1, \ldots, t+q$. On average, of course, the news shock is realized because the expected realizations of the surprise technology shocks are all zero, and this average response is what is reflected in an impulse response function, either one computed directly from the model or identified from a VAR. A feasible modification of the problem, though not considered here, would be to allow agents to observe only a noisy measure of u_t, which would then translate into a noisy forecast of future z and an even noisier forecast of future $\ln a$ relative to the specification described above. This setup would necessitate solving a signal extraction problem in which the agents in the economy effectively choose how much attention to pay to the signal of u_t, with agents reacting less the noisier is the signal. This kind of setup poses no special problems as pertains estimating a VAR, though it does require that special attention be paid to how one defines and interprets an impulse response.[7]

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, $\ln \widehat{a}_t = \ln a_t - \ln a_{t-1}$, and another variable, say output, $\ln \widehat{y}_t$.[8] Recall that because of the normalization, observed output is:

[7] Blanchard, L'Hullier, and Lorenzoni (2011) look at whether a VAR identified with long run restrictions can separately identify "news" (a true signal about permanent productivity) from "noise" (a signal about productivity that is false ex-post). The answer is no, but this is because a long run restriction on an impulse response function is an ex-ante restriction, and the agents in the model only learn that the signal is false ex-post. These authors refer to the problem resulting from this signal extraction problem as a non-invertibility, but this is a bit of a mis-characterization. Since the agents in their economy cannot separate out the true component of a signal from the false component ex-ante, neither can an econometrician estimating a VAR. If, rather than trying to identify IRFs to the "deeper" shocks that the agents cannot separate out, the econometrician seeks only to identify the impulse response to an innovation in the signal, then there may be no invertibility problem. Non-invertibility problems as stressed in most of the recent literature, and in the present paper, ask whether an econometrician can identify shocks observed by agents. The exercise in Blanchard, L'Hullier, and Lorenzoni (2011) tasks the econometrician with identifying shocks that agents cannot observe, which can only be (partially) accomplished with full information methods.

[8] For the remainder of this section I focus on the case in which output is observed. The results are essentially the same conditioning on other observed variables.

$\ln y_t - \frac{1}{1-\alpha} \ln a_t$, so this imposes the cointegrating relationship of the structural model.[9] Suppose a researcher estimates a VAR(p) on the system $x_t = [\ln \widehat{a}_t \;\; \ln \widehat{y}_t]$ (abstracting from the constant term):

$$x_t = A_1 x_{t-1} + \cdots + A_p x_{t-p} + u_t, \;\; E(u_t u_t') = \Sigma_u \tag{34}$$

Where u_t is a vector of innovations with variance-covariance matrix Σ_u, which in general is not diagonal. It is assumed there exists a mapping between orthogonal shocks, e_t, and the reduced form innovations given by: $u_t = A_0^{-1} e_t$, where A_0^{-1} is the "impact matrix". After normalizing the variance of each orthogonal shock to unity, one sees that $\Sigma_u = A_0^{-1} A_0^{-1'}$. With n variables in the system, one needs to impose $n(n-1)/2$ restrictions to uniquely recover A_0^{-1}.

In the proposed VAR system, a recursive orthogonalization of the innovations, with $\ln \widehat{a}_t$ "ordered" first, is apparently consistent with the implications of the model. This amounts to A_0^{-1} being identified as the Choleski decomposition of the estimated variance-covariance matrix of residuals. 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 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, 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.[10]

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, α, is fixed at 1/3. With average productivity growth of one percent, output and its components will grow at 1.5 percent on average,

[9] 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.

[10] The choice of 8 lags is somewhat arbitrary. As shown in Section 2, the mapping from model to data yields a VAR(∞). 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(∞), 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).

which is broadly consistent with the post-war US per capita data. The subjective discount factor, β, 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 ξ, 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, θ, 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 and separability in preferences, the parameters governing the utility from holding real balances, χ and ν, 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.[11] As we will see below, the surprise technology shock being more important than news shocks tends to exacerbate any problems due to non-invertibility, and can therefore be considered relatively conservative. Finally, the time lag between the revelation of news and its affect on productivity, q, is set to 3. This means that there are three quarters 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, ϕ 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.

4.1 Full Model

In the fully parameterized model with both real and nominal frictions, the state vector is:[12]

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

With three periods of anticipation there are three additional state variables – $z_{1,t}$, $z_{2,t}$, and $z_{3,t}$. TFP growth and normalized output are observed. Denote the vectors of observables as

[11] 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.

[12] 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. $\ln \widehat{y}_t = \ln y_t - \frac{1}{1-\alpha} \ln a_t$.

$x_t = [\ln \widehat{a}_t \;\; \ln \widehat{y}_t]'$. The shock vector is $\epsilon_t = [\varepsilon_t \;\; u_t]$. 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 $M = \left(A - BD^{-1}C\right)$. 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.[13] 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, on average, the structural VAR does 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 impulse responses at short forecast horizons 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

[13] The log levels are obtained by (i) cumulating the response of the growth rate of TFP to the shock and (ii) adding $\frac{1}{1-\alpha}$ times the cumulated log level TFP response to the response of normalized output.

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 three 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 small and likely of little practical concern.

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:

$$s_t = \left[\begin{array}{cccc} \ln \widehat{k}_t & z_{1,t} & z_{2,t} & z_{3,t} \end{array} \right]'$$

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 $M = \left(A - BD^{-1}C\right)$ 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 RBC 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 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, ε_t. In this case, the state space of the model would simply be the transformed capital stock, $\ln \widehat{k}_t$. 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 $M = \left(A - BD^{-1}C\right) = 0.9554 - \frac{-1.4331 \times 0.2210}{-0.3315}$, which is equal to 0. Hence, the invertibility condition is satisfied. This means that a univariate autoregression of $\ln \widehat{y}_t$ 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.[14] In contrast, it is straightforward to show that if there are only news shocks and no surprise shocks, a univariate autoregression in either output or TFP growth is non-invertible. It is the presence of anticipation effects that drives the non-invertibility.

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:

$$s_t = \left[\begin{array}{ccccccc} \ln \widehat{y}_t & \ln \widehat{k}_t & i_t & \ln v_t & z_{1,t} & z_{2,t} & z_{3,t} \end{array} \right]'$$

The maximum modulus of the eigenvalues of $M = \left(A - BD^{-1}C\right)$ 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 Using SVARs to Conduct Model Comparisons

One of the main 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 effects of technology shocks on other variables in the short run – RBC models with few real frictions, for example, typically predict that hours should increase following technological improvement, whereas some sticky price models with insufficiently accomodative monetary policy 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.[15]

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

[14] This statement requires some clarification; the impulse response estimated from the autoregression would correctly recover the impulse response of the *normalized* variable $\ln \widehat{y}_t = \ln y_t - \frac{1}{1-\alpha} \ln a_t$. 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).

[15] 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.

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 has the exact opposite prediction: 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 3 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 model, 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.

4.5 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 performance of the VARs.

Figure 4 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: $\sigma_\varepsilon = 0.01$ (which is the benchmark value) and $\sigma_\varepsilon = 0.001$. 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 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): $\Sigma_u = C\Sigma_s C' + D\Sigma_\epsilon D'$. The variance of the forecast of the state conditional on observables, Σ_s, drives a wedge between the VAR innovations, u_t, and the deep economic shocks, ϵ_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 ε_t shocks declines, the observed variables will do a better job of revealing lagged values of the news shock.

Figure 5 plots the Euclidian norm of Σ_s (a measure of the "size" of the variance in forecasting the state) as a function of the relative magnitude of the standard deviations of the two shocks. For this exercise, σ_u is held fixed at 0.005 and σ_ε is varied. One observes that, as $\sigma_u/\sigma_\varepsilon$ gets large, Σ_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.

4.6 The Role of the Anticipation lag

Up to this point the "anticipation lag" between when agents observe a news shock and when it affects technology has been fixed at $q = 3$. This choice is arbitrary though it conforms with much of the theoretical work on news shocks. Nevertheless, the quality of the results varies in an interesting and instructive way with the anticipation lag. The manner in which the results vary reinforces the notion that non-invertibility is fundamentally a problem of informational insufficiency.

First, consider the case in which $q = 1$, so that there is only one period of anticipation. Appendix section A.5 presents the numerical values of the state space matrixes for the "full model" when $q = 1$ and shows that the eigenvalues of $(A - BD^{-1}C)$ are all strictly less than unity in modulus. In other words, the model with one period of anticipation gives rise to an invertible VAR representation in output and technology. This means that an estimated VAR in these variables on a long enough sample with enough lags will exactly recover the impulse responses to both news and surprise technology shocks. Kurmann and Otrok (2011) similarly find an invertible VAR representation based on a Smets and Wouters (2007) type model with one period anticipated growth shocks. The intuition for this is reasonably straightforward – with only one period of anticipation, there is only one additional state, $z_{1,t} = u_t$. The innovation in output (or other variables from the model) orthogonalized with respect to current technology growth reveals that missing state. With multiple periods of anticipation the innovation in the second variable in the system cannot perfectly reveal q missing states if $q > 1$. Hence, for news shocks to generate non-invertibility there must be more than one period of anticipation.

Moving beyond the case of $q = 1$, since non-invertibility is fundamentally the result of informational deficiency, it stands to reason that the more missing state variables there are, the worse will be the performance of VARs. Since the anticipation lag increases the state vector one for one, one should expect that VARs perform more poorly as the anticipation lag increases.

Table 3 shows summary statistics for Monte Carlo exercises using data generated from the full model with different anticipation lags. To focus in on the bias due to non-invertibility, these statistics are based on estimating a VAR in technology growth and output on one data set of 100,000 observations (i.e. the "large sample" exercise). The structure of the table is similar to Table 2 – it shows the average absolute deviation of the responses of each variable to the two identified shocks over a forty period forecast horizon. The final column shows the average of the average deviations, which serves as a crude metric for the overall bias of the VAR impulse responses relative to the model responses. As the final column indicates, the average fit declines as q, the anticipation lag, increases. The intuition for this is straightforward – the bigger is q, the more missing state variables there are, and therefore the bigger is the wedge between the VAR innovations and shocks. Though there are biases in the estimated responses to the surprise technology shock, these are fairly similar for different q. The biases in the estimated responses to news shocks, in contrast, rise with q. For extremely long anticipation lags (e.g. $q = 16$), the VAR biases are quite large.[16] The average absolute deviation of the estimated output response over forty forecast horizons, for example, is 0.33 percent. This is quite large when considering that the long run response of output to a news shock as calibrated is 0.75 percent.

These findings have some implications for the existing empirical literature on new shocks. Beaudry and Portier (2006), for example, present impulse responses to a news shock which have no discernable effect on TFP for a period of four or five years. The Monte Carlo exercise with $q = 16$ suggests that, if the true response to a news shock had a delay of that long, there would be little hope of a structural VAR recovering the impulse responses correctly.[17] This makes interpreting the Beaudry and Portier (2006) impulse responses as responses to news problematic. The impulse responses identified in Barsky and Sims (2011), in contrast, affect measured TFP within a few quarters (and show quite different impact effects on hours and output). If the true news process affects productivity quickly then there is a much better chance that these impulse responses are accurate representations of the true responses from an economic model.

5 Robustness

The Monte Carlo exercises to this point have been conducted in a relatively simple environment. In particular, I have restricted attention to a two variable system in which a recursive identification lines up with the implications of news and surprise technology shocks in the economic model. This has been done with the goal of cleanly focusing in on the role of non-invertibility as a source of bias in impulse responses obtained from estimated VARs. Nevertheless, one might be interested in how non-invertibilities matter in more realistic environments with which researchers typically have to grapple – for example, in situations in which exogenous TFP is not observed, in which there are more than two shocks, or when the VAR is subject to some other kind of specification bias.

[16]These exercises are conducted with a fixed lag length, $p = 8$, as the anticipation lag, q, varies. It bears pointing out that one needs at least as many lags in the VAR as the anticipation lag to have any hope of capturing the dynamics implied by the model. In that sense the "cards are stacked" against the specification with $q = 16$. Nevertheless, allowing for a much longer lag structure preserves the basic pattern evident in the table. I maintain $p = 8$ to (i) maintain congruity with Table 2 and (ii) to be consistent with empirical work using VARs, which almost never uses more than two years worth of lags.

[17]It is worth pointing out that magnitude of the biases depends on the interaction of the discount factor, β, with the anticipation lag. For example, the quality of the Monte Carlo results is essentially the same for a quarterly calibration with $q = 4$ and $\beta = 0.99$, and an annual calibration with $q = 1$ and $\beta = 0.96$.

This section considers several extensions to more realistic environments and confirms that the same basic results from the earlier Monte Carlo analyses obtain.

Though much of the empirical work on news shocks which uses VARs assumes that "true technology" is well-measured by a suitably constructed total factor productivity (TFP) series, in practice this assumption may be problematic.[18] I therefore consider cases in which TFP is unobserved by the econometrician. Separately identifying surprise technology shocks from news shocks is generally accomplished with a short run recursive assumption as implemented throughout Section 4; without a measure of TFP, however, recursive zero restrictions are generally not available. As such, I begin by assuming that there is no surprise technology shock – the only source of movements in $\ln a_t$ in the model are news shocks. In terms of the process for technology, this can be accomplished by setting the variance of ε_t to zero. So as to avoid stochastic singularity and be able to estimate a VAR with more than one variable, I augment the model with a "demand" shock to the monetary policy rule, replacing (23) with:

$$i_t = \rho i_{t-1} + (1-\rho)\psi_\pi (\pi_t - \pi^*) + (1-\rho)\psi_y \left(\frac{y_t}{y_{t-1}} - \Delta_y \right) + \nu_{i,t}, \;\; \nu_{i,t} \sim N(0, \sigma_\nu) \tag{35}$$

For the purposes of the Monte Carlo exercise I simulate data from the full model with the modified shock process using the baseline parameterization for news (i.e. $q = 3$ and $\sigma_u = 0.005$) and set the standard deviation of the policy shock to $\sigma = 0.0015$, or 15 basis points. I conduct the "small sample" exercise of the previous section – I draw shocks and generate 500 samples with 200 observations of data each. I estimate a VAR in output growth and the consumption-output ratio: $x_t = [\Delta \ln y_t \;\; \ln c_t - \ln y_t]$.[19] I use a long run restriction to separate out the technology news shock from the monetary policy shock, imposing that only the news shock may have a permanent effect on the level of output in the long run. This restriction is consistent with the implications of the underlying model.[20]

Figure 6 summarizes the Monte Carlo results. The solid lines show theoretical impulse responses of the levels of output and consumption to a news shock (upper row) and the monetary policy shock (lower row). The dashed lines are the average estimated response over the 500 simulations, and the shaded gray regions the middle 68 percent of the distribution of estimated responses. Qualitatively the estimated VAR does a very good job at replicating the model's dynamics to both shocks. The estimated responses to the news shock are essentially unbiased on impact and for several horizons thereafter; these responses are slightly downward biased at longer horizons but the magnitude of

[18]In practice there is some debate about how to best measure total factor productivity. The measure produced by Basu, Fernald, and Kimball (2006) uses some simple theoretical restrictions to attempt to control for factor hoarding and thus produces a measure of TFP which is immune to criticisms that it largely measures unobserved factor variation due to demand shocks.

[19]This specification imposes the model generated cointegrating relationship, which is necessary to use a long run restriction. A consumption-output VAR is also estimated in Cochrane (1994), which in many ways is an important antecedent to the more recent literature on news shocks.

[20]The long run restriction is implemented as in Gali (1999). The long run response of the cumulated sums of the variables in the VAR to the two orthogonal shocks is given by the matrix $C_0 = \left(I - \sum_{j=1}^{p} A_j\right)^{-1} A_0^{-1}$. Letting the policy shock occupy position 2, the restriction imposes that A_0^{-1} satisfy $A_0^{-1} A_0^{-1'} = \Sigma_u$ and be such that the long run response of the level of output to the second shock be zero, e.g. that the (1,2) element of C_0 be zero.

the bias is not large. The estimated responses to the policy shock are also quite good, on both a qualitative and quantitative dimension. In short, the VAR with the long run restriction does a good job at capturing the dynamic responses to both the news shock and the policy shock, even when TFP is unobserved.

Next I consider the more challenging case in which there are both news and surprise technology shocks (for this exercise there is no monetary policy shock), but continue to assume that the econometrician cannot observe TFP. I continue to generate 500 data sets from the full model with 200 observations each. For this exercise I consider estimating a two variable VAR in the growth rate of average labor productivity and labor hours, $x_t = [\Delta(\ln y_t - \ln n_t) \;\; \ln n_t]$.[21]

The challenging aspect of this system is that "conventional" VAR identifying restrictions are not available. Since both shocks permanently affect the non-stationary variables of the model, a long run restriction cannot be used to differentiate between the two kinds of technology shocks. And since TFP is not observed, and all other variables react on impact to both kinds of technology shocks, short run zero restrictions are not available either. I therefore employ a modified sign/shape restriction to separately identify the news and surprise technology shocks. In particular, I identify the news shock as the orthogonal shock which generates an impulse response of labor productivity with the greatest difference between the response at a horizon of 40 quarters from the impact response. Even though average labor productivity reacts immediately to news shocks because of movements in hours, this restriction gets at the idea that the news shock should have a much bigger longer run effect on average productivity than on impact.

Before discussing the Monte Carlo results I briefly discuss the mechanics of this particular restriction. In the VAR system (34) the reduced form innovations are a function of the structural shocks: $u_t = A_0^{-1} e_t$. For a particular A_0^{-1} satisfying $A_0^{-1} A_0^{-1'} = \Sigma_u$, call it $\widetilde{A}_0^{-1}$, the entire space of permissible impact matrixes can be represented by $\widetilde{A}_0^{-1}\widetilde{D}$, where $\widetilde{D}$ is an orthonormal matrix, e.g. $\widetilde{D}\widetilde{D}' = I$. The sign/shape restriction methodology consists of (i) starting with an arbitrary $\widetilde{A}_0^{-1}$, say a Choleski decomposition and, (ii) searching over the space of orthonormal matrixes, $\widetilde{D}$, for an impact matrix, $\widetilde{A}_0^{-1}\widetilde{D}$ that satisfies the sign/shape restriction. Step (i) ensures that the resulting impact matrix satisfies $A_0^{-1} A_0^{-1'} = \Sigma_u$. In many applications step (ii) results in a *set* of candidate impulse vectors – i.e. there is not point identification, but rather set identification (see, e.g., Faust, 1998, and Uhlig, 2005). Because my restriction is not simply qualitative (e.g. a monetary tightening lowers both output and prices), but can rather be quantified in terms of the difference between the impulse response of labor productivity at horizon 40 relative to impact, my restriction achieves point identification.

Figure 7 summarizes the results from the sign/shape restriction exercise, where again the dashed line represents the average estimated response across the 500 simulations, the solid line the true model response, and the shaded gray region the middle 68 percent of the distribution of estimated responses. The VAR does a very good job at capturing the model's responses to the two technology

[21] In practice the variables in the VAR do not matter much. I consider a different system here than in the previous exercise mainly for breadth of coverage, but also because the average productivity growth-hours VAR is popular in the literature, e.g. Gali (1999).

shocks. In response to the news shock hours rise slightly on impact and remain high until the level of technology improves three periods later, after which time they sharply decline. The response of labor productivity mirrors the movements in hours – it initially falls and then sharply rebounds when technology actually improves, and continues to rise thereafter. The impulse responses from the sign restricted VAR do a good job at capturing these features. The results for the surprise technology shocks are even better – here the biases are very small. In spite of the non-invertibility and in spite of the fact that TFP is not observed, the VAR does a very good job at capturing the dynamic response to both technology shocks.

As a final robustness exercise I consider the realistic modification of the underlying DSGE model in which there are more than two shocks. In addition to both kinds of technology shocks (news and surprise), I assume that there are two "demand shocks" – a monetary policy shock as in (35), and a preference shock that shifts the disutility of labor. The preference shock obeys a stationary AR(1):

$$\ln \theta_t = (1 - \rho_\theta)\theta^* + \rho_\theta \ln \theta_{t-1} + \varepsilon_{\theta_t} \tag{36}$$

Here $\rho_\theta \in (0, 1)$ and θ^* is chosen to generate steady state labor hours of one-third. For the Monte Carlo exercises the parameterization is the same as in the benchmark, with the volatility of the monetary policy shock of 15 basis points and the volatility of the preference shock of 1 percent, with $\rho_\theta = 0.9$. As earlier, I generate 500 data sets with 200 observations each.

I conduct two separate Monte Carlo exercises on data generated from the four shock system. For simplicity I assume that the researcher can observe TFP, though the results using average labor productivity and sign restrictions are similar. In one case I assume that a researcher estimates a four variable VAR in TFP growth, output and consumption normalized relative to the level of TFP (which imposes the model based cointegrating relationships), and hours worked: $x_t = [\ln \widehat{a}_t \;\; \ln \widehat{y}_t \;\; \ln \widehat{c}_t \;\; \ln n_t]$. In the other case I assume that the researcher estimates a three variable VAR in TFP growth, normalized output, and hours: $x_t = [\ln \widehat{a}_t \;\; \ln \widehat{y}_t \;\; \ln n_t]$. The second specification introduces another bias in addition to non-invertibility, since there are four economic shocks but the estimated VAR model includes only three variables.

With four shocks, even though TFP is observed, a recursive identification is inconsistent with the predictions of the model. This is because the innovation in any of the other three variables orthogonalized with respect to the TFP innovation will be driven by both the news shock and the two demand shocks. As such, I implement a combined short run/long run identification. I use two long run restrictions to separate out the two technology shocks from the two demand shocks, and use a short run zero restriction to identify the surprise technology shock from the news shock. In particular, the restrictions are (i) only the two technology shocks have a long run impact on $\ln a_t$, and (ii) the news shock has no immediate impact on $\ln a_t$.[22] These restrictions are sufficient

[22] Using similar notation to above, the long run responses of the cumulated sums of the variables in the VAR are given by: $C_0 = \left(I - \sum_{j=1}^{p} A_j\right)^{-1} A_0^{-1}$. With TFP growth in the first position in the system and the surprise and news shocks indexed by 1 and 2, the two long run restrictions impose that the (1,3) and (1,4) elements of C_0 be zero. The short run restriction that news not immediately effect TFP requires that the (1,2) element of A_0^{-1} be 0.

to uniquely identify the two technology shocks, but do not identify the other shocks. The same restrictions can be used in either the three or the four variable VAR systems.

Figures 8 and 9 show the Monte Carlo results; Figure 8 for the four variable system and Figure 9 for the mis-specified three variable system. Focusing first on the four variable system, one sees that the estimated responses to the news shock align quite closely with the true model responses, particularly at shorter forecast horizons. The transitional dynamics to the new steady state are also estimated quite well, with only evidence of some small bias in the output response at longer forecast horizons. Focusing on the responses to the surprise technology shock, there is some more evidence of bias in the responses, particularly at longer forecast horizons. It should be noted, however, that VAR impulse responses are quite good on impact and at short horizons, and the hours response is estimated very well over all horizons.

Figure 9 summarizes the Monte Carlo results for the three variable system. The identification is the same as in the four variable system. As noted, in addition to the non-invertibility due to news, an additional bias is introduced due to the fact that the VAR is estimated with three variables, whereas there are four shocks in the model. The estimated impulse responses to both kinds of technology shocks nevertheless remain quite good. Qualitatively the estimated responses to the news shock are in line with the responses from the model. Relative to the four variable VAR, however, the three variable system does not do as good of a job at capturing the large drop in hours when productivity improves, and the longer horizon biases in the output and TFP responses are somewhat larger. The quality of the estimated responses to the surprise technology actually appears somewhat better than in the three variable system. In terms of mean biases this is illusory, as the distribution of estimated responses is quite a bit wider in the three variable system. Nevertheless, the VAR does a good job at qualitatively matching the dynamic responses to the surprise technology shock, and the estimated hours response is quite good.

There are a number of different permutations on these Monte Carlo exercises one could consider for the four shock model. One is mentioned above, and that is assuming that the econometrician cannot observe TFP. In that case, one can estimate either a three or four variable model, replacing TFP with average labor productivity. The long run restrictions used to separate out the two technology shocks from the two demand shocks are the same, and one can replace the zero recursive restriction to identify the news shock with the sign/shape restriction implemented above in the two variable example. These results turn out to be quite similar in terms of how closely the estimated responses align with the model responses. I also considered including different observables in the VAR system; here, too, the results are fairly similar. Finally, I considered alternative parameterizations of non-technology shocks. As shown in Section 4.5, the relative magnitudes of the news and surprise technology shocks impact the bias in the estimated responses to the news shock. The relative magnitudes of the unidentified shocks can also matter for the quality of fit of the identified responses (where in these exercises the unidentified shocks are the demand shocks). It turns out that the biases of the identified responses grow with the relative magnitude of the non-identified shocks, as suggested by Canova and Paustian (2011). However, even when increasing the magnitudes of these shocks by a lot – say, a factor of three – the estimated responses to the two identified technology shocks remain qualitatively quite good with relatively small biases, particularly at short horizons.

6 Adding Information

The Monte Carlo results of the previous two sections 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 title 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:

$$s_{i,t} = u_t + v_{i,t}, \qquad i = 1, \dots, Q \tag{37}$$

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 10 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 Section 4, 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, recursively identified VAR estimated on $\ln \widehat{a}_t$ and $\ln \widehat{y}_t$. 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.[23] 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.[24] Table 4 shows the mean absolute deviations of the estimated impulse responses of output and TFP to both news and surprise technology shocks over a forty quarter forecast horizons. The final column shows the average of the mean deviations as a measure of overall fit. The pattern is quite evident – as more information variables are added, the average of the average absolute deviations becomes smaller.

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.[25] These methods 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: $x_t = [\ln \widehat{a}_t \;\; \ln \widehat{y}_t \;\; F_t]'$, where F_t is the first principal component of Q information variables. Isolating just the first principal 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 any biases due to non-invertibility ought to

[23] To be clear, the estimated VARs are then: $X_t = [\ln \widehat{a}_t \;\; \ln \widehat{y}_t \;\; s_{1,t} \ldots s_{r,t}]'$, for $r = 4$ or $r = 8$.

[24] It is worth pointing out that the standard deviation of the noise innovations, σ_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.

[25] For applications in macroeconomics, see, for example Bernanke, Boivin, and Eliasz (2005) and Stock and Watson (2005).

disappear, since the missing state variables will effectively be revealed.

Figure 11 shows impulse responses obtained from both finite and large sample Monte Carlo exercises making use of the first principal component of $Q = 30$ factors (the number 30 is arbitrary; the important point is that it be "large"). I again 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 principal component of the Q information series is obtained, F_t. In the second step, a conventional unrestricted VAR is estimated with $\ln \widehat{a}_t$, $\ln \widehat{y}_t$, 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. No interpretation is given to the innovation in F_t ordered last.

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 mainly a function of Q being finite – conditioning on sufficiently more variables in the first stage would cause any remaining biases to vanish entirely. The last row of Table 4 shows average mean absolute deviations of the responses estimated in the large sample exercise from the true model counterparts and reveals that the factor augmented VAR represents an improvement over both the two variable system and the systems augmented with smaller numbers of information variables. The results are also substantially better in finite samples in comparison with the simple two variable VAR – comparing the finite sample results in Figure 11 with Figure 2, there is a clear improvement. 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 largely unrelated to non-invertibility. The estimated responses are essentially unbiased on impact and for a number of quarters thereafter. 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 factors 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.

7 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 responses 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 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 nonetheless be small. 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

[1] Altig, David, Lawrence Christiano, Martin Eichenbuam, and Jesper Linde. 2011. "Firm-Specific Capital, Nominal Rigidities, and the Business Cycle." *Review of Economic Dynamics*, 14(2), 225-247.

[2] Barsky, Robert and Eric Sims. 2011. "News Shocks and Business Cycles." *Journal of Monetary Economics* 58(3), 273-289.

[3] Barsky, Robert and Eric Sims. 2012. "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence." *American Economic Review* 102(4), 1343-1377.

[4] Basu, Susanto, John Fernald, and Miles Kimball. 2006. "Are Technology Improvements Contractionary?" *American Economic Review* 96(5), 1418-1448.

[5] Beaudry, Paul and Franck Portier. 2006. "Stock Prices, News, and Economic Fluctuations." *American Economic Review* 96(4), 1293-1307.

[6] Bernanke, Ben, Jean Boivin, and Piotr Eliasz. 2005. "Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach." *Quarterly Journal of Economics* 120(1), 387-422.

[7] Bils, Mark and Peter J. Klenow. 2004. "Some Evidence on the Importance of Sticky Prices." *Journal of Political Economy* 112(5), 947-985.

[8] Blanchard, Olivier, Jean-Paul L'Hullier, and Guido Lorenzoni. 2011. "News, Noise, and Fluctuations: An Empirical Exploration." Unpublished manuscript, Massachussetts Institute of Technology.

[9] Calvo, Guillermo. 1983. "Staggered Prices in a Utility Maximizing Framework." *Journal of Monetary Economics* 12(3), 383-398.

[10] Canova, Fabio and Matthias Paustian. 2011. "Business Cycle Measurement with Some Theory." *Journal of Monetary Economics* 58(4), 345-361.

[11] Chari, VV, Patrick Kehoe, and Ellen McGrattan. 2008. "Are Structural VARs with Long-Run Restrictions Useful in Developing Business Cycle Theory?" *Journal of Monetary Economics* 55(8), 1337-1352.

[12] Christiano, Lawrence, Martin Eichenbuam, and Charles Evans. 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113(1), 1-45.

[13] Cochrane, John. 1994. "Permanent and Transitory Components of GNP and Stock Prices." *Quarterly Journal of Economics* 109(1), 241-265.

[14] Dupor, William and Jing Han. 2011. "Handling Non-Invertibility: Theory and Applications." Unpublished manuscript, Ohio State University.

[15] Faust, Jon and Eric Leeper. 1997. "When Do Long-Run Identifying Restrictions Give Reliable Results?" *Journal of Business & Economic Statistics* 15(3), 345-353.

[16] Faust, Jon. 1998. "On the Robustness of Identified VAR Conclusions about Money." *Carnegie-Rochester Conference Series on Public Policy* 49, 207-244.

[17] Fernandez-Villerde, Jesus, Juan Rubio-Ramirez, Thomas Sargent, and Mark Watson. "ABCs (and Ds) of Understaning VARs." *American Economic Review* 97(3): 1021-1026.

[18] Feve, Patrick and Ahmat Jidoud. 2011 "Identifying News Shocks from SVARs." Unpublished manuscript, Toulouse School of Economics, 2011.

[19] Fisher, Jonas. 2006. "The Dynamic Effects of Neutral and Investment-Specific Technology Shocks." *Journal of Political Economy* 114(3), 413-451.

[20] Forni, Mario and Luca Gambetti. 2011. "Fiscal Foresight and the Effects of Government Spending." Unpublished manuscript, Universitat Autonoma de Barcelona.

[21] Forni, Mario, Luca Gambetti, and Luca Sala. "No News in Business Cycles." Unpublished manuscript, Universitat Autonoma de Barcelona.

[22] Forni, Mario, Domenico Giannone, Marco Lippi, and Lucrezia Reichlin. 2009. "Opening the Black Box: Structural Factor Models with Large Cross Sections." *Econometric Theory* 25(5), 1319-1347.

[23] Gali, Jordi. 1999. "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" *American Economic Review* 89(1), 249-271.

[24] Giannone, Domenico and Lucrezia Reichlin. 2006. "Does Information Help Recovering Structural Shocks from Past Observations?" *Journal of the European Economic Association* 4(2-3), 455-465.

[25] Hansen, Lars, William Roberds, and Thomas Sargent. 1991. "Time Series Implications of Present Value Budget Balance and of Martingale Models of Consumption and Taxes." In *Rational Expectations Econometrics*, edited by Lars Peter Hansen, John Heaton, Albert Marcet, William Roberds, and Thomas J. Sargent, Westview Press: Boulder, Co., 121-163.

[26] Hansen, Lars and Thomas Sargent. 1991. "Two Difficulties in Interpreting Vector Autoregressions." In *Rational Expectations Econometrics*, edited by Lars Peter Hansen, John Heaton, Albert Marcet, William Roberds, and Thomas J. Sargent, Westview Press: Boulder, Co., 77-120.

[27] Hansen, Lars and Thomas Sargent. 2007. "Recursive Linear Models of Dynamic Economies." Unpublished manuscript, NYU.

[28] Jaimovich, Nir and Sergio Rebelo. 2009. "Can News about the Future Drive Business Cycles?" *American Economic Review* 99(4), 1097-1118.

[29] Kimball, Miles and Matthew D. Shapiro. 2010. "Labor Supply: Are Income and Substitution Effects Both Large or Both Small?" Unpublished manuscript, University of Michigan.

[30] Kurmann, Andre and Christopher Otrok. 2011. "News Shocks and the Term Structure of Interest Rates: A Challenge for DSGE Models." Unpublished manuscript, University of Missouri.

[31] Leeper, Eric, Todd Walker, and Shu-Chun Yang. 2011. "Foresight and Information Flows." Unpublished manuscript, Indiana University.

[32] Lippi, Marco and Lucrezia Reichlin. "The Dynamic Effects of Aggregate Demand and Supply Disturbances: Comment." *American Economic Review* 83(3), 644-652.

[33] Mountford, Andrew and Harald Uhlig. 2009. "What are the Effects of Fiscal Policy Shocks?" *Journal of Applied Econometrics* 24(6), 960-992.

[34] Quah, Danny. 1990. "Permanent and Transitory Movements in Labor Income: An Explanation for 'Excess Smoothness' in Consumption." *Journal of Political Economy* 98(3), 449-475.

[35] Sims, Christopher and Tao Zha. 2006. "Does Monetary Policy Generate Recessions?" *Macroeconomic Dynamics* 10(2), 231-272.

[36] Sims, Eric. 2011. "Permanent and Tranistory Technology Shocks and the Behavior of Hours: A Challenge for DSGE Models." Unpublished manuscript, University of Notre Dame.

[37] Smets, Frank and Raf Wouters. 2007. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review* 97(3), 586-606.

[38] Stock, James and Mark Watson. 2005. "Implications of Dynamic Factor Models for VAR Analysis." Unpublished manuscript, Princeton University.

[39] Uhlig, Harald. 2005. "What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure." *Journal of Monetary Economics* 52(2), 381-419.

Table 1: Parameters

Full Model	RBC Model	Sticky Price Model
$\beta = 0.99$	$\gamma = 0$	$\gamma = 0$
$\delta = 0.02$	$\tau = 0$	$\tau = 0$
$g_a = 0.0025$	$\phi = 0$	
$\pi^* = 0.005$		
$\alpha = \frac{1}{3}$		
$\xi = 1$		
$\sigma_e = 0.01$		
$\sigma_u = 0.005$		
$\theta \Rightarrow n^* = \frac{1}{3}$		
$\gamma = 0.7$		
$\tau = 2.5$		
$\phi = 0.7$		
$\epsilon = 11$		
$\rho = 0.8$		
$\psi_\pi = 1.5$		
$\psi_y = 1.5$		

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

	Tech. to News	Output to News	Tech. to Surprise	Output to Surprise	Average
Large Sample					
Full	0.0266	0.0698	0.0693	0.0871	0.0632
RBC	0.0676	0.1006	0.0366	0.0596	0.0661
Sticky Price	0.1062	0.1402	0.0403	0.0619	0.0872
Small Sample					
Full	0.0466	0.0522	0.2833	0.3545	0.1841
RBC	0.0974	0.1339	0.2595	0.3375	0.2071
Sticky Price	0.2031	0.2662	0.1278	0.1524	0.1874

These numbers represent the average absolute deviations of the estimated impulse responses relative to the true responses in the models over the first 40 horizons, 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 final column, labeled "Average," shows the average of the mean absolute deviations of all four impulse responses.

Table 3: Mean Absolute Deviations: Different Anticipation Lags

	Tech. to News	Output to News	Tech. to Surprise	Output to Surprise	Average
Anticipation Lag					
$q = 2$	0.0330	0.0273	0.0389	0.0454	0.0362
$q = 8$	0.0996	0.1495	0.0459	0.0544	0.0874
$q = 16$	0.1942	0.3300	0.0485	0.0679	0.1601

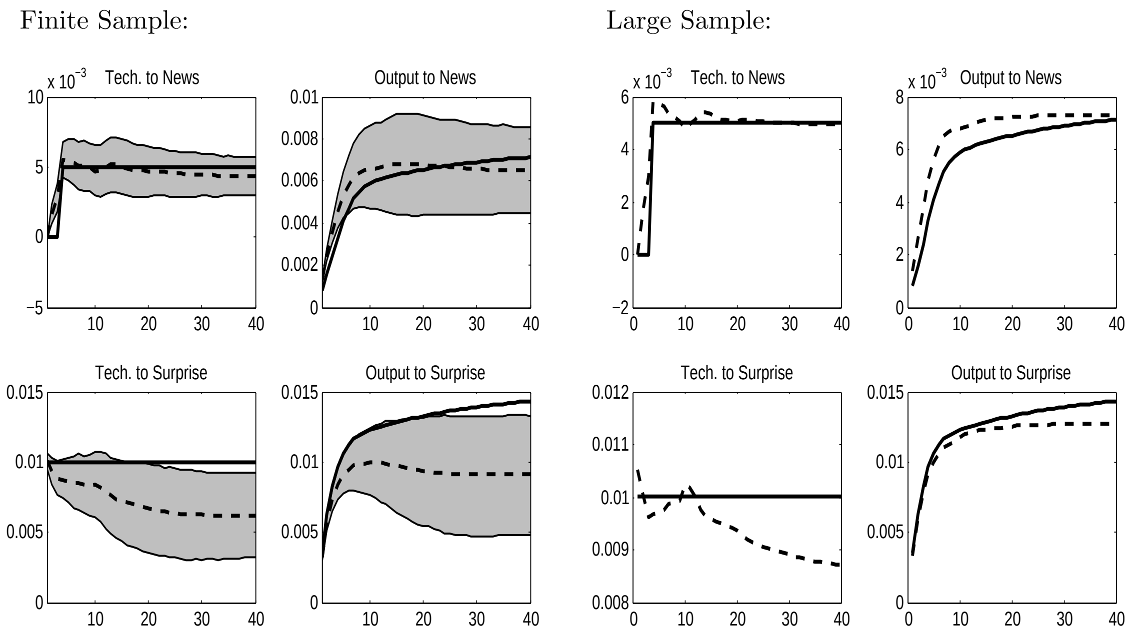
These numbers represent the average absolute deviations of the estimated impulse responses relative to the true responses in the models over the first 40 horizons, times 100, for different news anticipation lags in the full model specification in a two variable VAR in $\ln \widehat{a}_t$ and $\ln \widehat{y}_t$. The final column, labeled "Average," shows the average of the mean absolute deviations of all four presented impulse responses.

Table 4: Mean Absolute Deviations: Adding Information

	Tech. to News	Output to News	Tech. to Surprise	Output to Surprise	Average
Number of Signals					
Standard VAR ($s = 0$)	0.0634	0.0932	0.0298	0.0511	0.0594
$s = 1$	0.0613	0.0924	0.0272	0.0454	0.0566
$s = 2$	0.0500	0.0758	0.0160	0.0288	0.0426
$s = 4$	0.0369	0.0562	0.0056	0.0130	0.0279
$s = 8$	0.0206	0.0317	0.0074	0.0078	0.0169
Factor VAR	0.0068	0.0102	0.0159	0.0200	0.0132

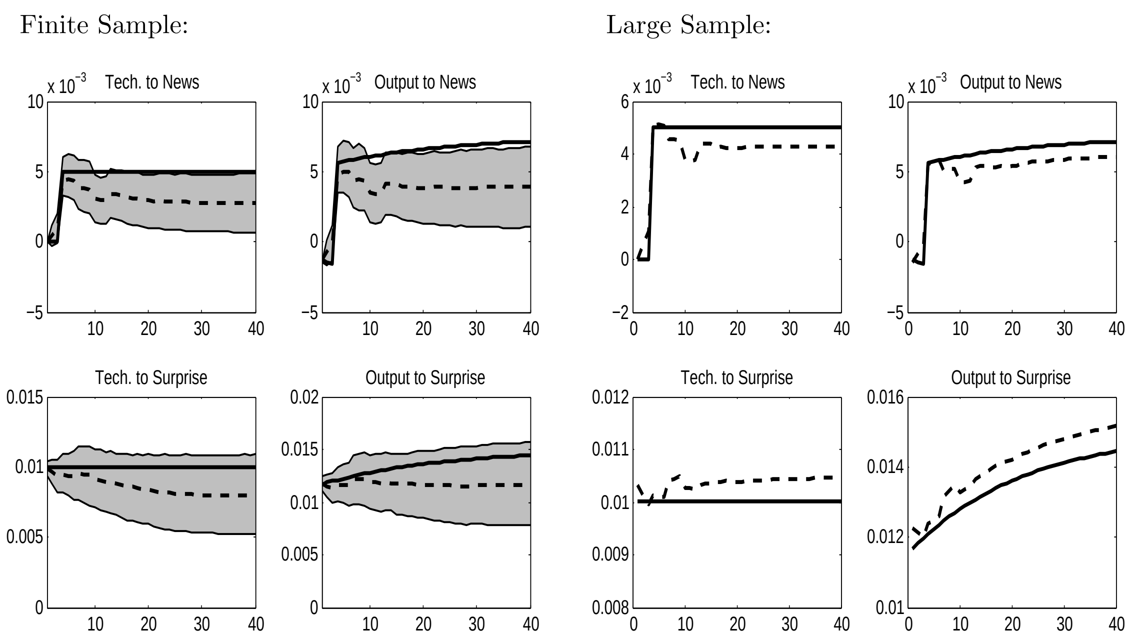
These numbers represent the average absolute deviations of the estimated impulse responses relative to the true responses in the models over the first 40 horizons, times 100, for different numbers of additional information variables included in a two variable VAR in $\ln \widehat{a}_t$ and $\ln \widehat{y}_t$. The final column, labeled "Average" shows the average of the mean absolute deviations of all four presented impulse responses.

Figure 1: Monte Carlo Results in Full Model



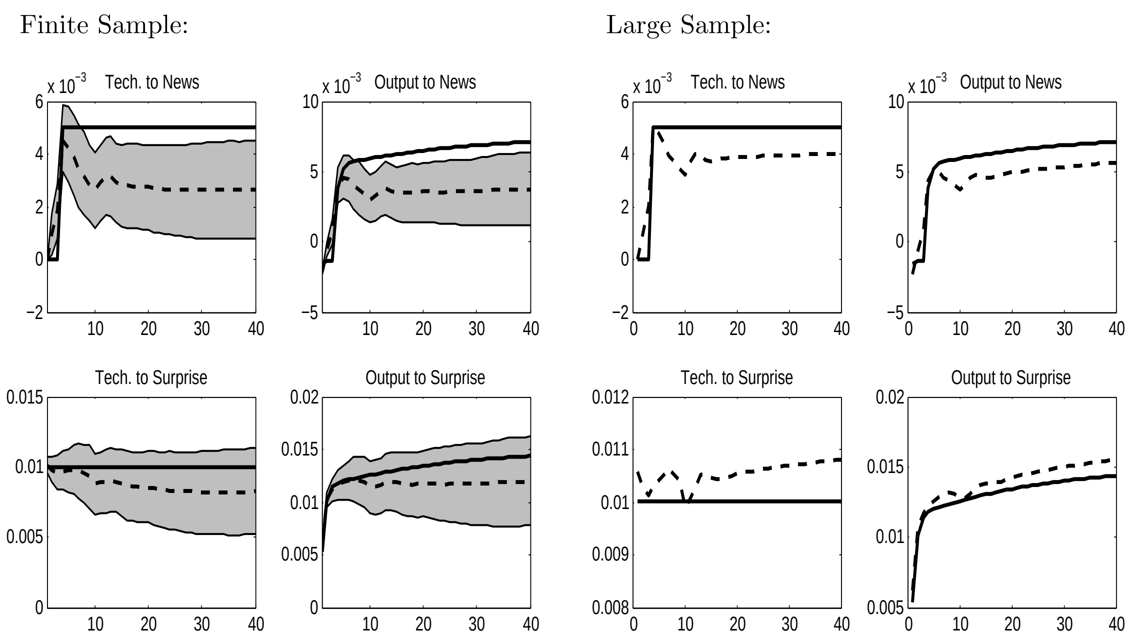
The solid lines are the theoretical impulse responses to news and surprise technology shocks in the "Full Model" using the parameterization as described in the text. For the Monte Carlo exercises the VARs feature $\ln \widehat{a}_t$ and $\ln \widehat{y}_t$ and are estimated with $p = 8$ lags. In the 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.

Figure 2: Monte Carlo Results in RBC Model



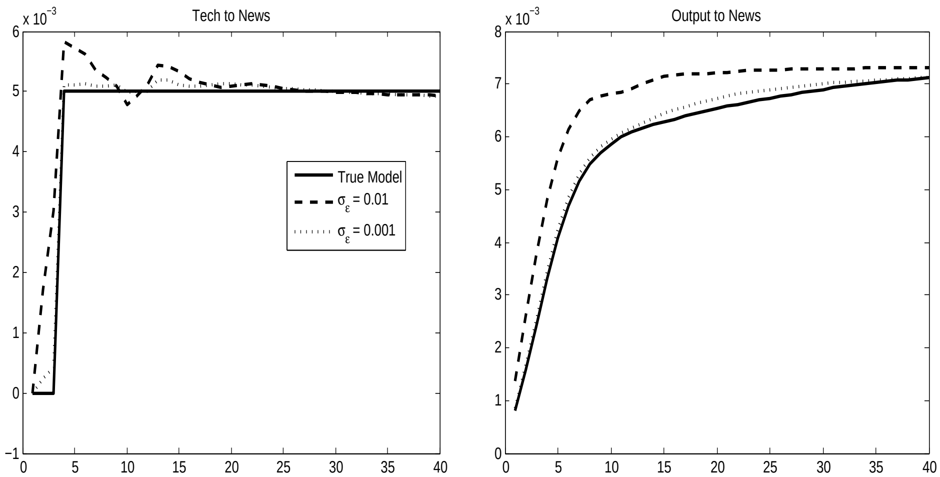
The solid lines are the theoretical impulse responses to news and surprise technology shocks in the "RBC Model" using the parameterization as described in the text. For the Monte Carlo exercises the VARs feature $\ln \widehat{a}_t$ and $\ln \widehat{y}_t$ and are estimated with $p = 8$ lags. In the left 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.

Figure 3: Monte Carlo Results in Sticky Price Model



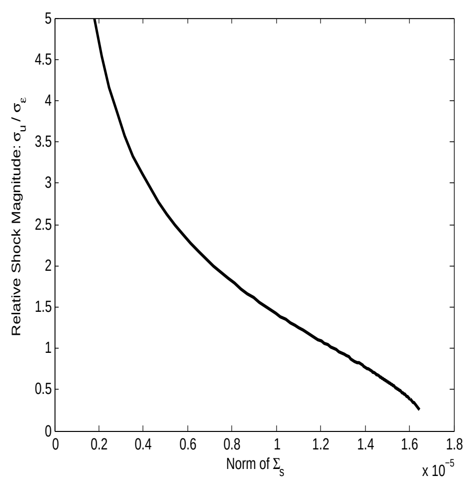
The solid lines are the theoretical impulse responses to news and surprise technology shocks in the "Sticky Price Model" using the parameterization as described in the text. For the Monte Carlo exercises the VARs feature $\ln \widehat{a}_t$ and $\ln \widehat{y}_t$ and are estimated with $p = 8$ lags. 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.

Figure 4: Monte Carlo Results: Large vs. Small Surprise Shocks



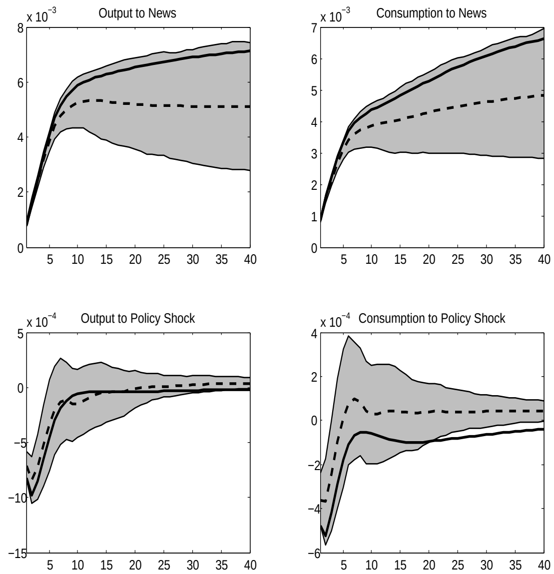
The solid lines are the theoretical impulse responses to a news shock in the "Full Model" using the parameterization as described in the text. The dashed and dotted lines represent the estimated responses from a large sample Monte Carlo exercise using two different values for the standard deviation of surprise technology shocks: $\sigma_\varepsilon = 0.01$ (dashed line, which also corresponds to the baseline used in previous figures) and $\sigma_\varepsilon = 0.001$ (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 5: Theoretical Wedge and Varying Shock Magnitudes



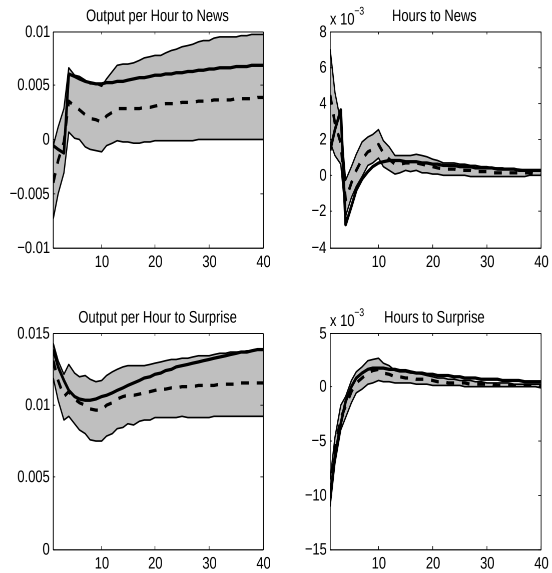
This figure plots the Euclidian norm of Σ_s, the variance-covariance matrix of period t optimal forecast of the state, as a function of the relative shock magnitudes between surprise, ε_t, and news shocks, u_t. This is done in the context of the "Full Model" conditional on observing TFP growth and normalized output. The vertical axis plots the ratio of the the standard deviation of u_t divided by the standard deviation of ε_t against the determinant of Σ_s, which is obtained numerically from solving the Ricatti equations (6)-(7).

Figure 6: Monte Carlo Results: Two Variable Model, TFP Unobserved, LR Restriction



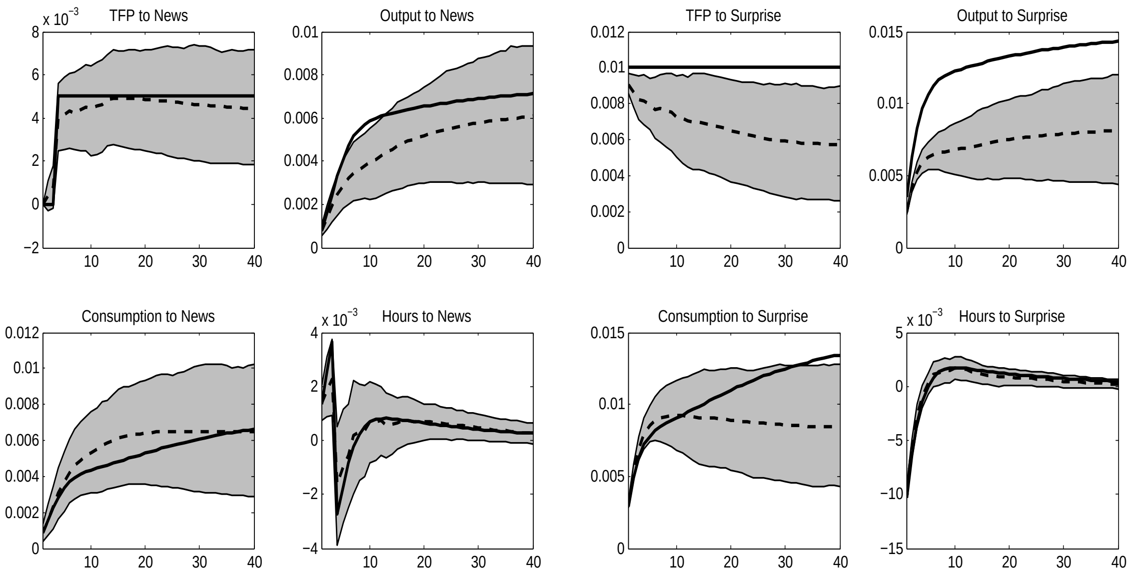
The solid lines are the theoretical impulse responses to a news shock about productivity and to a monetary policy shock in the full specification of the model in which (i) there is no surprise technology shock and (ii) the standard deviation of monetary policy shocks is set to 0.15 basis points. For the Monte Carlo exercise the estimated VAR features $\Delta \ln y_t$ and $\ln c_t - \ln y_t$ and is estimated with $p = 8$ lags. The shocks are identified using a long run restriction such that the news shock is the sole source of the unit root in output. The dashed lines are the mean responses to each shock averaged across 500 simulations of data with 200 observations each, while the shaded gray regions depict the middle 68 percent of the distribution of estimated responses across the 500 simulations.

Figure 7: Monte Carlo Results: Two Variable Model, TFP Unobserved, Sign Restriction



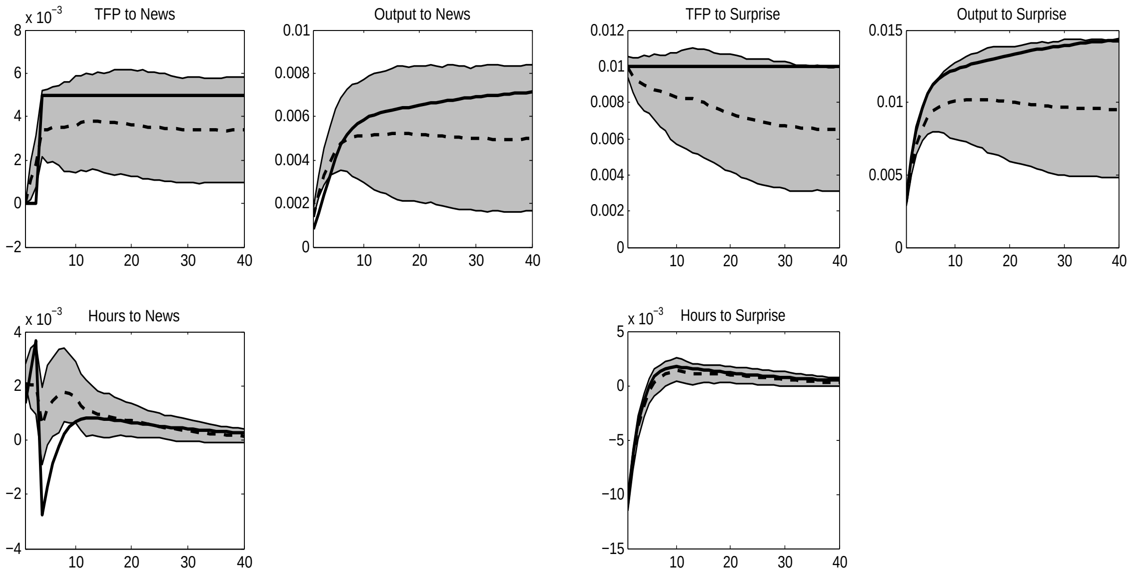
The solid lines are the theoretical impulse responses to a news shock about productivity and to a surprise technology shock in the full model. For the Monte Carlo exercise the estimated VAR features the growth rate of average labor productivity, $\Delta(\ln y_t - \ln n_t)$ and log hours worked, $\ln n_t$, and is estimated with $p = 8$ lags. The shocks are identified with a shape restriction: in particular, the news shock is identified as the shock that results in an impulse response with the largest difference between the level response of average productivity at a 40 period horizon relative to the impact response. The dashed lines are the mean responses to each shock averaged across 500 simulations of data with 200 observations each, while the shaded gray regions depict the middle 68 percent of the distribution of estimated responses across the 500 simulations.

Figure 8: Monte Carlo Results: Four Variable Model, LR and SR Restrictions



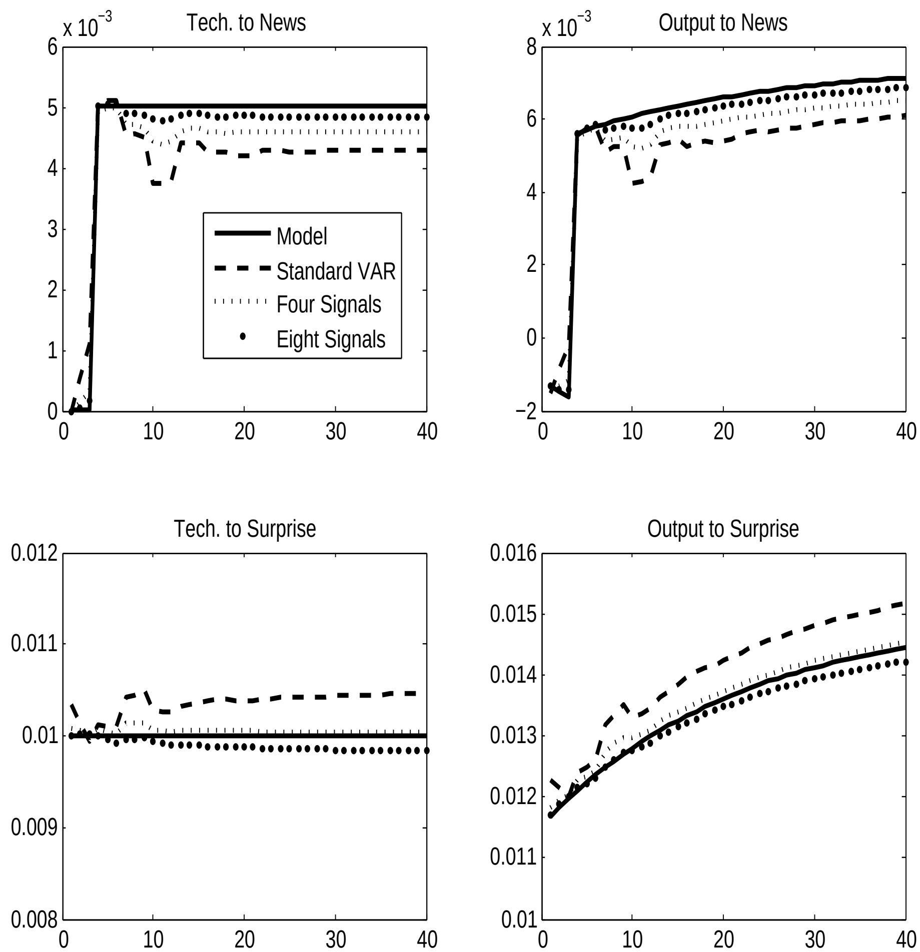
The solid lines are the theoretical impulse responses to a news shock about productivity and to a surprise technology shock in the full model with four shocks: a news shock, a surprise technology shock, a monetary policy shock, and a labor supply preference shock. For the Monte Carlo exercise the estimated VAR features $\ln \widehat{a}_t$, $\ln \widehat{y}_t$, $\ln \widehat{c}_t$, and $\ln n_t$, total hours worked, and is estimated with $p = 8$ lags. The news and surprise technology shocks are identified with a combined short run and long run restriction: the long run restriction says that these two shocks are the sole source of the unit root in observed TFP, and the short run restriction says that news shock does not affect TFP on impact. The remaining two shocks are left unidentified. The dashed lines are the mean responses to each shock averaged across 500 simulations of data with 200 observations each, while the shaded gray regions depict the middle 68 percent of the distribution of estimated responses across the 500 simulations.

Figure 9: Monte Carlo Results: Three Variable Model (Four Shocks), LR and SR Restrictions



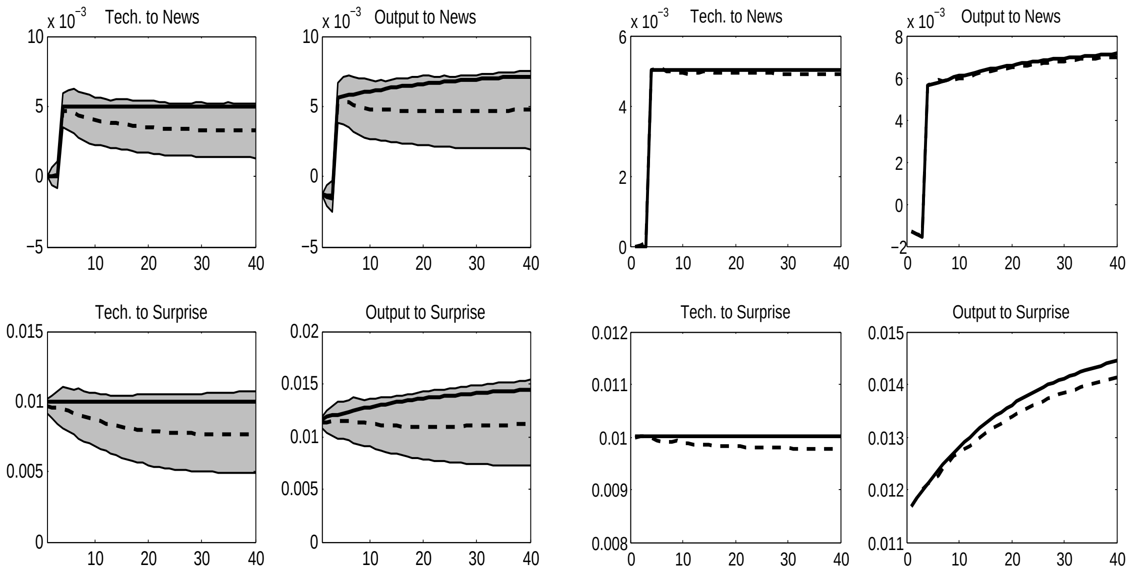
The solid lines are the theoretical impulse responses to a news shock about productivity and to a surprise technology shock in the full model with four shocks: a news shock, a surprise technology shock, a monetary policy shock, and a labor supply preference shock. For the Monte Carlo exercise the estimated VAR features $\ln \widehat{a}_t$, $\ln \widehat{y}_t$, and $\ln n_t$, total hours worked, and is estimated with $p = 8$ lags. The model is misspecified in the sense that there are four shocks in the data generating process but only three variables are included in the VAR. The news and surprise technology shocks are identified with a combined short run and long run restriction: the long run restriction says that these two shocks are the sole source of the unit root in observed TFP, and the short run restriction says that news shock does not affect TFP on impact. The dashed lines are the mean responses to each shock averaged across 500 simulations of data with 200 observations each, while the shaded gray regions depict the middle 68 percent of the distribution of estimated responses across the 500 simulations.

Figure 10: Adding Information to the VAR: RBC Model



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 a VAR featuring $\ln \widehat{a}_t$ and $\ln \widehat{y}_t$, identified using a recursive restriction with TFP growth ordered first and output ordered second. 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.

Figure 11: Monte Carlo Results: Factor Augmented VAR (RBC Model)



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 $\ln \widehat{a}_t$, $\ln \widehat{y}_t$, and the first principal 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:

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

The vector of observables is:

$$x_t = \begin{bmatrix} \ln \widehat{a}_t & \ln \widehat{y}_t \end{bmatrix}'$$

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

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

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

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

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

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

$$\Lambda = \begin{pmatrix} 1.3208 & 1.3208 & 0.5781 & 0.9471 & 0.7239 & 0.8057 & 0.0000 & 0.0000 & 0.0000 \end{pmatrix}$$

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:

$$s_t = \begin{bmatrix} \ln \widehat{k}_t & z_{1,t} & z_{2,t} & z_{3,t} \end{bmatrix}'$$

The vector of observables is the same as above.

The parameters of the state space representation are:

$$A = \begin{bmatrix} 0.9554 & -1.4331 & -0.0745 & -0.0704 \\ 0.0000 & 0.0000 & 1.0000 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & 1.0000 \\ 0.0000 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}$$

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

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

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

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

$$\Lambda = \begin{pmatrix} 0.8987 & 1.0572 & 1.0572 & 0.0000 \end{pmatrix}$$

Again, the conditions required for invertibility are not met.

A.3 The Sticky Price Model

The state vector is:

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

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

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

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

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

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

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

$$\Lambda = \begin{pmatrix} 1.2397 & 1.2397 & 0.8087 & 0.8087 & 0.7638 & 0.0000 & 0.0000 \end{pmatrix}$$

The conditions for invertibility are not satisfied.

A.4 The Full Model with One Period Anticipation

The full model with one period of anticipation is similar to that presented in A.1 but the state space can be reduced because of one period of anticipation. In particular, the process for technology can be written:

$$\ln a_t = g_a + \ln a_{t-1} + \varepsilon_t + z_{1,t-1}$$
$$z_{1,t} = u_{t-1}$$

For this specification the state vector is:

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

The vector of observables is again TFP growth and normalized output, $\ln \widehat{a}_t$ and $\ln \widehat{y}_t$, and the shock vector is $\epsilon_t = [\varepsilon_t \quad u_t]'$. The parameters of the state space representation are:

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

$$B = \begin{bmatrix} -1.1461 & 0.3006 \\ -1.4857 & 0.0111 \\ -0.0452 & 0.0096 \\ -0.0410 & -0.0105 \\ 0.0000 & 1.0000 \\ -0.8958 & 0.4691 \\ -1.2136 & 0.2552 \end{bmatrix}$$

$$C = \begin{bmatrix} -0.0000 & -0.0000 & -0.0000 & 0.0000 & 1.0000 & 0.0000 & 0.0000 \\ 0.0319 & 0.1411 & -0.2543 & -0.0600 & -1.2136 & -0.0153 & 0.6513 \end{bmatrix}$$

$$D = \begin{bmatrix} 1.0000 & -0.0000 \\ -1.2136 & 0.2552 \end{bmatrix}$$

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

$$\Lambda = \begin{pmatrix} 0.5381 & 0.9411 & 0.8043 & 0.7294 & 0.0000 & 0.0000 & 0.0000 \end{pmatrix}$$

Since the maximum modulus of the eigenvalues lies inside of the unit circle, the conditions for invertibility are satisfied with only one period of anticipation.