

**Identifying Noise Shocks**

Luca Benati, Joshua Chan,  
Eric Eisenstat, Gary Koop

18-07

March 2018

**DISCUSSION PAPERS**

# Identifying Noise Shocks<sup>\*</sup>

Luca Benati  
University of Bern<sup>†</sup>

Joshua Chan  
University of Technology Sydney<sup>‡</sup>

Eric Eisenstat  
University of Queensland<sup>§</sup>

Gary Koop  
University of Strathclyde<sup>¶</sup>

## Abstract

We make four contributions to the ‘news *versus* noise’ literature: (I) We provide a new identification scheme which, in population, exactly recovers news and noise shocks. (II) We show that our scheme is not vulnerable to Chahrour and Jurado’s (2018) criticism about the observational equivalence of news and noise shocks, which uniquely holds if the econometrician only observes a fundamental, and agents’ expectations about it. By contrast, we show that observational equivalence breaks down when the econometrician observes macroeconomic variables encoding information about the signal (and therefore about news and noise shocks), because they are chosen by agents conditional on all information, including the signal itself. (III) We propose a new econometric methodology for implementing our identification scheme, and we show, *via* a Monte Carlo study, that it has an excellent performance. (IV) We provide several empirical applications of our identification scheme and econometric methodology. Our results uniformly suggest that, contrary to previous findings in the literature, noise shocks play a minor role in macroeconomic fluctuations.

---

<sup>\*</sup>We wish to thank Harris Dellas, Carlo Favero, Francois Gourio, Luca Sala and Eric Sims for useful discussions and/or suggestions. The usual disclaimers apply.

<sup>†</sup>Department of Economics, University of Bern, Schanzeneckstrasse 1, CH-3001, Bern, Switzerland. Email: luca.benati@vwi.unibe.ch

<sup>‡</sup>Economics Discipline Group, University of Technology Sydney, Sydney, NSW, Australia. Email: joshuacc.chan@gmail.com

<sup>§</sup>School of Economics, University of Queensland, Brisbane St Lucia, QLD 4072, Australia. Email: e.eisenstat@uq.edu.au

<sup>¶</sup>Department of Economics, University of Strathclyde, 199 Cathedral Street, Glasgow G4 0QU, United Kingdom. Email: gary.koop@strath.ac.uk

# 1 Introduction

Noise shocks—defined as announcements about future fundamentals which, in fact, never materialize—offer the possibility of generating macroeconomic fluctuations without any variation in the economy’s fundamentals. They are therefore radically different from news shocks, i.e. announcements about future fundamentals which do indeed materialize at some future date.

Finding a way to distinguish between news and noise shocks has proved difficult. Indeed, Chahrour and Jurado (2018) claim that news and noise shocks are observationally equivalent. In this paper, we develop a new identification scheme which, in population, exactly recovers news and noise shocks, and show that Chahrour and Jurado’s (2018) claim is not generally valid. Our identification scheme cannot be imposed using a structural Vector Autoregressive (SVAR) model, but we show how it can be imposed using a structural Vector Autoregressive Moving Average (SVARMA) model. We develop Bayesian econometric methods for implementing the identification scheme and demonstrate, in a Monte Carlo study, their excellent performance. We then use our methods in several empirical applications. In all of these we find that, contrary to previous findings in the literature, noise shocks play a minor role in macroeconomic fluctuations.

Our identification strategy is based on the general principle that agents’ inability to distinguish news and noise shocks on impact—which is the essence of the entire ‘news *versus* noise’ problem—implies that

- (1) the responses of the economy to the two shocks at  $t=0$  will be the same, whereas
- (2) they will progressively diverge further out.

As time goes by agents gradually learn whether the announcement was news (i.e. that the fundamentals will indeed change) or simply noise, leaving fundamentals unchanged. To provide some intuition, consider an announcement that purports to promise a future positive permanent shock to Total Factor Productivity (TFP). Is this announcement news or noise? Initially, it is impossible to tell. As we illustrate in Section 3 using the real business cycle (RBC) model of Barsky and Sims (2011) augmented with noise shocks about TFP, such an announcement indeed produces, on impact, identical impulse vectors regardless of whether it is news or noise (see Figure 2). The *crux* of our identification strategy is that, although news and noise shocks are observationally equivalent on impact, they are not at all subsequent horizons. The news shock will ultimately cause a permanent increase in TFP, whereas the noise shock will not. This implies that there is no way that the two shocks might be confused once the entire set of their properties is taken into account. Quite simply, within the present example, a permanent shock cannot be observationally equivalent to a transitory one.

This apparently conflicts with Chahrour and Jurado’s (2018; henceforth, CJ) claim that news and noise shocks are observationally equivalent. We show that CJ’s result

crucially hinges on the fact that they cast their entire analysis in terms of the joint process for fundamentals and *agents' expectations* about them. When the analysis is instead cast in terms of the joint process for fundamentals and *signals* about news and noise, observational equivalence breaks down, and the two shocks are, in fact, separately identified.<sup>1</sup> This implies that since, in general, *agents* should be assumed to observe the signal, they are in fact able to disentangle news and noise shocks. We further argue that the fact that agents observe the signal, and the econometrician does not, does not imply that the observational equivalence of news and noise shocks holds from the econometrician's perspective. Econometricians can disentangle news and noise shocks, and the reason for this is straightforward. Although econometricians do not observe the signal, they do observe macroeconomic variables encoding information about the signal—and therefore about news and noise shocks—because they have been chosen by agents conditional on all information, including the signal itself. This implies that the distinction between economic agents (who observe the signal) and econometricians (who do not) is, for discussion of observational equivalence of news and noise shocks, irrelevant. The key point is that both of them have information about the signal, either because they directly observe it (the agents), or because they see how macroeconomic variables react to it (the econometricians). Working with the RBC model with noise shocks about TFP, we then provide a straightforward illustration of how the impulse response functions (IRFs) to news and noise shocks, and the fractions of forecast error variance (FEV) of the variables they explain, can be exactly recovered in population by imposing the model-implied restrictions that (i) news and noise shocks produce identical impulse vectors on impact, and (ii) news and non-news shocks are the only disturbances having a permanent impact on TFP.

Our econometric methodology is based on Bayesian estimation of structural VARMA models.<sup>2</sup> The need for using VARMA instead of VARs originates from the fact that the ‘news *versus* noise’ problem automatically produces a reduced-rank structure for the matrix of the shocks’ impact responses at  $t=0$ .<sup>3</sup> In the econometrics literature, a common finding is that VARMA can be over-parameterized and difficult to estimate, particularly in high dimensional setups such as that used in this paper. Building on previous work with reduced form VARMA,<sup>4</sup> we derive methods for estimating the SVARMA implied by our identification scheme. We use stochastic search variable selection (SSVS) methods—see in particular George *et al.* (2008)—to pick out restrictions and ensure parsimony in an otherwise over-parameterized model. In our

---

<sup>1</sup>In a companion note—see Benati, Eisenstat, and Koop (2018)—we provide a general proof of this point, based on Hilbert spaces arguments.

<sup>2</sup>The general point that SVARMA can be used to estimate (nonfundamental) semi-structural models goes back to Hansen and Sargent (1981) and Ito and Quah (1989).

<sup>3</sup>Intuitively, this has to do with the fact that the very essence of this problem causes the impulse vectors to news and noise shocks at  $t=0$  to be identical. In turn, this automatically implies that, with two identical columns, the matrix of the shocks’ structural impacts at  $t=0$  cannot have full rank.

<sup>4</sup>See Chan *et al.* (2016).

Monte Carlo study, we find our algorithm to work extremely well.

We consider three empirical applications of our identification scheme and econometric methodology. The first ‘news *versus* noise’ problem pertains to TFP. In the second we consider a model involving real dividends and stock prices (with the ‘noisy news’ setup pertaining to real dividends). Finally we consider an application involving defense expenditure. In all cases, we find that noise shocks play a minor role in macroeconomic fluctuations. Our results for the system featuring TFP contrast with those of both Blanchard *et al.* (2013) and Forni *et al.* (2017), who found a significant role for noise shocks pertaining to productivity. Likewise, results for the system featuring dividends and stock prices contrast with the corresponding findings of Forni *et al.* (2017).

The paper is organized as follows. The next section discusses several theoretical implications of the ‘news *versus* noise’ problem. In Section 3 we illustrate the main features of Barsky and Sims’ (2011) RBC model augmented with noise shocks about TFP. We show, working in population, that low-order SVARMAs can approximate very well the model’s theoretical IRFs and fractions of FEV explained by individual shocks. We also discuss our identifying restrictions. In Section 4 we address the potential observational equivalence issue raised by CJ. Section 5 describes our econometric methodology, and discusses the Monte Carlo evidence about its performance. Section 6 contains our three empirical applications. Section 7 concludes, and discusses possible directions for future research.

## 2 Theoretical Implications of the ‘News *versus* Noise’ Problem

In this section we discuss several theoretical implications of the news versus noise problem and illustrate them with reference to a standard present-value model for dividends and stock prices. Three implications which play a crucial role in our identification strategy have already been mentioned:

- (1) *On impact (i.e., at  $t=0$ ) news and noise shocks generate identical IRFs.*
- (2) *After impact (i.e., for all  $t>0$ ) IRFs to news and noise shocks progressively diverge.*
- (3) *The impact matrix of the structural shocks at  $t=0$  has reduced rank.*

Two further theoretical implications of the news versus noise problem are the following:

- (4) *The structural model possesses a non-fundamental representation.*
- (5) *The absolute magnitude of the economy’s response to noise shocks at  $t=0$  is monotonically decreasing in their standard deviation.*

As already discussed, (1) is a direct logical implication of agents’ inability to distinguish news and noise shocks on impact—which is the essence of the news versus noise problem—whereas (2) simply originates from the fact that, as time goes by, the

true nature of the shock is progressively revealed to agents. In particular, in the RBC model of Section 3, the news shock ultimately causes a permanent increase in TFP, whereas the noise shock does not cause any permanent change in any variable. One small qualification to (1) is that exact equality between the IRFs to news and noise shocks at  $t=0$  holds when the standard deviations of the two shocks are identical. In the general case, the two impulse vectors at  $t=0$  will be proportional, with the coefficient of proportionality being equal to the ratio between the two shocks' standard deviations. Our identification strategy will therefore impose proportionality, rather than strict equality, between the two impulse vectors at  $t=0$ . This proportionality implies that the impact matrix of the structural shocks at  $t=0$  will be reduced rank.<sup>5</sup> This will automatically imply (4), i.e., that the model possesses a non-fundamental representation.<sup>6</sup>

The reason for the fifth implication is straightforward. As the volatility of noise shocks increases, the signal becomes less and less informative, and agents therefore react to it less and less. In the limit, as the volatility of noise shocks tends to infinity, so that the signal becomes completely uninformative, agents' reaction to it—and therefore to noise shocks—tends to zero. This implies that the absolute magnitude of the IRFs to noise shocks at  $t=0$  is monotonically decreasing in their standard deviation.<sup>7</sup> With the magnitude of the economy's response decreasing as the volatility of noise shocks increases, an obvious question is whether the importance of these shocks at driving macroeconomic fluctuations (in terms of the fractions of FEV of the variables they explain) is increasing or decreasing with their volatility. Although it is not possible to make general statements on this issue, in the model for dividends and stock prices of the next sub-section, the relationship is hump-shaped (see panel III of Figure 1 below, whereas panel II presents the corresponding evidence for the IRFs of stock prices at  $t=0$ ). Intuitively, when the standard deviation of noise shocks is small, the increase in the volatility of noise shocks dominates the decrease in agents' response to them. As this standard deviation increases, there becomes a point where the latter effect becomes the dominant one, and the importance of noise shocks in driving stock prices decreases.

---

<sup>5</sup>It is worth mentioning in passing that the proportionality of IRFs (and therefore the corresponding matrix of impacts being reduced-rank) only holds on impact. This is a direct logical implication of the fact that, as time goes by, the true nature of the shock is progressively revealed to agents, so that, at all periods after impact, the economy's responses to news and noise shocks will necessarily diverge.

<sup>6</sup>In Section 5.1 we discuss how we address the issue of non-fundamentality within our econometric framework. In a nutshell, for each draw from the posterior distribution of the coefficients of the VARMA model, we consider all of the possible representations obtained by 'flipping the MA roots' inside and outside of the unit circle.

<sup>7</sup>It is clear that the magnitude of the impact of noise shocks on the economy at  $t=0$  will be zero if their standard deviation is zero (in which case there are no noise shocks to speak of). This implies that the magnitude of the impact of noise shocks on the economy at  $t=0$ , as a function of the shocks' standard deviation, exhibits a discontinuity at zero.

## 2.1 A simple illustration

### 2.1.1 A present-value model for dividends and stock prices

We illustrate the news versus noise problem with reference to a standard present-value model for dividends and stock prices along the lines of Forni *et al.* (2017), in which stock prices are equal to the present discounted value of future expected dividends.<sup>8</sup> Labelling stock prices and log dividends as  $S_t$  and  $d_t$ , respectively, and using notation where  $t+j|t$  subscripts on any variable denote expectations at time  $t$  of the variable at time  $t+j$ , we have:

$$\ln S_t = \sum_{j=0}^{\infty} \beta^j d_{t+j|t} \quad (1)$$

where  $\beta$  is the discount rate. Log dividends are assumed be the sum of two unobserved components, a permanent and a transitory one,

$$d_t = d_t^P + d_t^T \quad (2)$$

which evolve according to

$$d_t^P = d_{t-1}^P + \epsilon_t^{NN} + \epsilon_{t-1}^{NE} \quad (3)$$

$$d_t^T = \rho_T d_t^T + v_t \quad (4)$$

where  $\epsilon_t^{NN}$  and  $\epsilon_t^{NE}$  are the non-news and news shocks, respectively, with  $\epsilon_t^{NN} \sim N(0, \sigma_{NN}^2)$  and  $\epsilon_t^{NE} \sim N(0, \sigma_{NE}^2)$ ;  $0 < \rho_T < 1$ ; and  $v_t \sim N(0, \sigma_v^2)$ .

Specification (2)-(4) is the same as the one used by Blanchard *et al.* (2013), with the crucial difference that, in their case, the equation for the permanent component does not feature a ‘proper’ news shock—defined as a time- $t$  disturbance which impacts upon the relevant variable only at a future date—as it is uniquely driven by a standard ‘surprise’ disturbance which enters contemporaneously. As we shall discuss below, this is the key reason why, on impact, their IRFs to the noise and surprise shocks are not the same.

Although at time  $t$  agents observe  $d_t$ , its two individual components,  $d_t^P$  and  $d_t^T$ , are never observed. Agents however observe a signal,  $s_t$ , which reveals some information about the news shock:

$$s_t = \epsilon_t^{NE} + u_t \quad (5)$$

with  $u_t \sim N(0, \sigma_u^2)$ . Online Appendix A.1 characterizes and solves the agent’s signal extraction problem. Based on (1), (2), and (A.5) the solution for log stock prices is given by

$$\ln S_t = \frac{d_{t|t}^P}{1-\beta} + \frac{d_{t|t}^T}{1-\rho_T\beta} + \frac{\beta}{1-\beta} \epsilon_{t|t}^{NE} = \frac{d_{t|t}^P}{1-\beta} + \frac{d_{t|t}^T}{1-\rho_T\beta} + \frac{\beta K_{32}}{1-\beta} [\epsilon_t^{NE} + u_t], \quad (6)$$

where  $K_{32}$  is defined in online Appendix A.1.

---

<sup>8</sup>Appendix A.2 presents an additional illustration based a standard New Keynesian model in which the news versus noise problem pertains to the natural rate of interest.

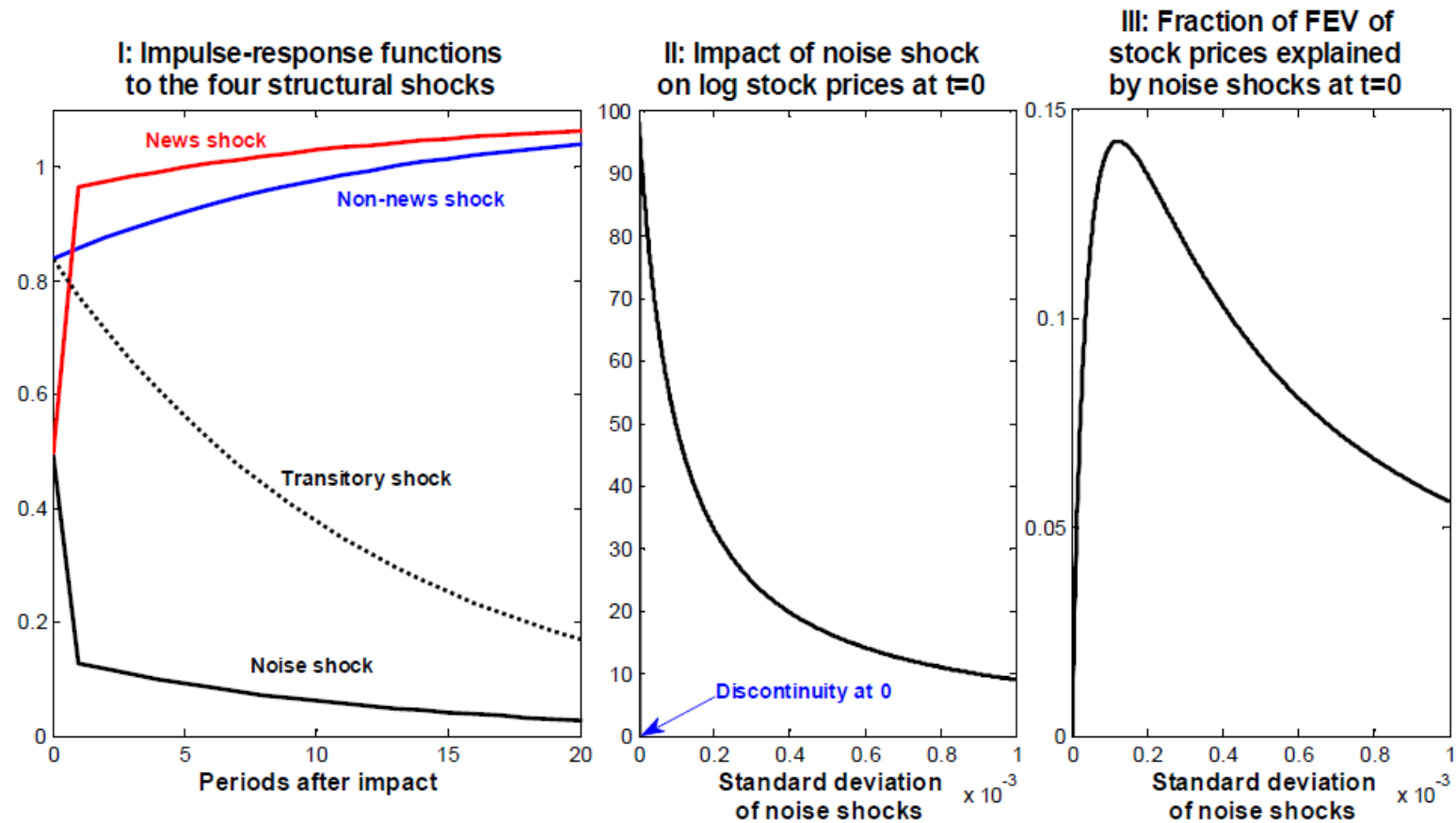


Figure 1 Impulse-response functions of log stock prices to the structural shocks, and fractions of the FEV of stock prices explained by noise shocks, as a function of their standard deviation



### 2.1.2 Illustrating the theoretical implications of the ‘news *versus* noise’ problem

Since, on impact, neither news nor noise shocks affects either  $d_{t|t}^P$  or  $d_{t|t}^T$  it follows that:

$$\left[ \frac{\partial \ln S_t}{\partial \epsilon_t^{NE}} \right]_{t=0} = \left[ \frac{\partial \ln S_t}{\partial u_t} \right]_{t=0} = \frac{\beta K_{32}}{1 - \beta} \quad (7)$$

In words, since on impact agents are unable to distinguish between news and noise shocks, they will react to either disturbance in the same way, and stock prices will jump by exactly the same amount—i.e., implication (1) above. This feature was also noticed, in passing, by Forni *et al.* (2017), who pointed out (see their Section 2.2) that ‘[...] the impact responses are identical, since the agents cannot distinguish between the two shocks immediately.’ Forni *et al.* (2017), however, did not exploit this for identification purposes, as their identification strategy is based on dynamic rotations of the VAR’s residuals.

Blanchard *et al.*’s (2013) IRFs, on the other hand, do not exhibit this property. The reason is that in their case the process for the permanent component of technology does not feature a news shock. Their analogue to our  $d_t^P$  is uniquely driven by a standard surprise disturbance. This implies that, on impact, there cannot be any symmetry between this disturbance and the noise shock. This is because whereas, at  $t=0$ , the noise shock only impacts upon the economy *via* the signal, the surprise shock affects both the signal and (*via* the permanent component) the relevant variable itself. As a result, the impact upon the economy at  $t=0$  of the two shocks considered by Blanchard *et al.* (2013) cannot possibly be the same. Noise shocks and news shocks, on the other hand, are perfectly symmetrical at  $t=0$ , as they impact upon the economy uniquely *via* the signal extraction problem, within which they cannot be distinguished. As a result, their impact on the endogenous variables at  $t=0$  must necessarily be identical.

The first panel of Figure 1 plots IRFs to unitary innovations in the four shocks, by setting  $\beta=0.99$ ,  $\rho_T=0.9$ ,  $\sigma_{NN}=\sigma_{NE}=\sigma_u=\sigma_v=0.01$ . As implied by (7), the impact responses to news and noise shocks are identical.<sup>9</sup> As time goes by, agents observe dividends, and therefore learn whether the signal at  $t=0$  was news or noise, and the two IRFs therefore progressively diverge: The IRF to news shocks converges to the new permanent level of stock prices, whereas that to noise shocks converges to zero.

Another point to stress is that, at all horizons beyond  $t=0$ , the magnitude of the IRF associated with noise shocks is smaller than the one of the IRF associated with news shocks. The reason is that agents are gradually learning about the nature of the shock. At  $t=0$  they cannot distinguish the two shocks (and so they react in exactly the same way), whereas at  $t=\infty$  they can perfectly disentangle them (and

---

<sup>9</sup>Notice that IRFs in Figure 1 are to shocks of size one, so that the impact responses to news and noise shocks are identical for either variable. As mentioned, in general the impacts are proportional, with the factor of proportionality being equal to the ratio between  $\sigma_u$  and  $\sigma_{NE}$ .

their reaction is therefore positive to the news, and zero to the noise). At all other horizons their reaction is intermediate between these two extremes. As we will see in Section 3.2 based on Barsky and Sims' (2011) RBC model augmented with noise shocks about TFP, this feature is a remarkably robust one.

As for implication (3), with four shocks, and just three observed variables (dividends, stock prices, and the signal), the present model does not allow to properly illustrate the reduced-rank structure of the matrix of the shocks' structural impacts at  $t=0$ . For this, we will have to wait for the RBC model of Section 3. As mentioned, non-fundamentality—implication (4)—is a direct implication of the reduced-rank structure of the matrix of structural impacts at  $t=0$ .

With regards to implication (5), the second panel of Figure 1 shows the impact response at  $t=0$  of stock prices to noise shocks as a function of their standard deviation. Apart from the discontinuity at zero, the impact response is monotonically decreasing, reflecting the fact that, as the signal becomes noisier and noisier, agents react to it less and less. Finally, the last panel of Figure 1 illustrates how, in this model and for these parameter values, the fraction of the FEV of stock prices at  $t=0$  explained by noise shocks, as a function of their standard deviation, exhibits a 'Laffer curve-type' shape. Over the initial portion of the parameter space, the increase in the volatility of noise shocks dominates the decrease in agents' response to them. Beyond a certain point, however, the latter effect becomes the dominant one, and the importance of noise shocks at driving stock prices decreases more and more.

We now turn to Barsky and Sims' (2011) RBC model, which we augment with noise shocks about future TFP.

### 3 Barsky and Sims' (2011) RBC Model Augmented With Noise Shocks About TFP

The preceding section examined the properties of news and noise shocks within a simple theoretical model. In our empirical work we will use SVARMA methods, which are atheoretical except for imposing identifying restrictions motivated by theory. In this section we motivate the use of SVARMA methods based on the RBC model of Barsky and Sims (2011), which we augment with noise shocks about TFP. In the next section, working with the theoretical MA representation of the model, we will show that our identifying restrictions can exactly recover news and noise shocks.

The basic elements of the RBC model of Barsky and Sims (2011) can be succinctly described as follows (for details, see Barsky and Sims, 2011, Section 2.2.1). Consumers maximize the discounted sum, with discount factor  $0 < \beta < 1$ , of the period- $t$  utility flows  $U_t \equiv \ln(C_t - bC_{t-1}) - E_t^N(1 + 1/\eta)^{-1}N_t^{1+1/\eta}$ , for  $t = 0, 1, 2, 3, \dots, \infty$ .  $C_t$  is consumption,  $N_t$  is hours worked,  $E_t^N$  is a random process whose logarithm,  $\epsilon_t^n$ , is  $N(0, \sigma_n^2)$ ,  $0 < b < 1$  is a parameter capturing the strength of habit formation and  $\eta$  captures the curvature of the labor supply function. Output,  $Y_t$ , is the sum

of consumption, investment ( $I_t$ ), and public expenditure ( $G_t$ ) and it is produced *via* the production function  $Y_t = A_t K_t^\theta N_t^{1-\theta}$ , where  $K_t$  is the capital stock, and  $\theta$  is the Cobb-Douglas parameter. The capital stock evolves according to the law of motion  $K_{t+1} = K_t(1 - \delta) + I_t[1 - (\gamma/2)(I_t/I_{t-1} - \tilde{g}_I)^2]$ , where  $\delta$  is the depreciation rate,  $\gamma$  is a parameter capturing the magnitude of capital adjustment costs, and  $\tilde{g}_I$  is the gross rate of growth of investment in the steady-state. Finally, public expenditure is postulated to be equal to a stationary random fraction of GDP, that is,  $G_t = g_t Y_t$ , with  $\ln g_t = \ln \bar{g} + \epsilon_t^g$ , with  $\epsilon_t^g$  being  $N(0, \sigma_g^2)$ .

Productivity is captured by  $A_t$ , and we assume that  $a_t = \ln(A_t)$  evolves according to

$$a_t = \tilde{a}_t + v_t \quad (8)$$

where  $\tilde{a}_t$  is the unobserved permanent component of productivity, and  $v_t$  is a white noise disturbance,  $v_t \sim N(0, \sigma_v^2)$ . The ‘news *versus* noise’ problem pertains to the productivity process. In particular, we assume that the permanent component of  $a_t$  evolves according to

$$\tilde{a}_t = \tilde{a}_{t-1} + \epsilon_t^{NN} + \epsilon_{t-\tau}^{NE} \quad (9)$$

where, once again,  $\epsilon_t^{NN}$  and  $\epsilon_t^{NE}$  are a non-news and a news shock, respectively, and  $\tau$  is the anticipation horizon for the news shock. Although at time  $t$  agents observe  $a_t$ , its two individual components,  $\tilde{a}_t$  and  $v_t$ , are never observed. In each period, however, agents receive a signal, which is equal to the sum of the news shock and of a noise component as in (5). Details of the agent’s signal extraction problem, together with the model’s solution, are given in online Appendix C.

Figure 2 shows the IRFs of the logarithms of GDP, consumption, investment, hours and technology to one-standard deviation non-news, news and noise shocks. The model is calibrated using parameter values similar to those in Barsky and Sims (see online Appendix C). As expected, the IRFs to news shocks are qualitatively in line with those reported in Barsky and Sims’ (2011) Figure 1, with consumption increasing on impact; GDP, investment and hours falling on impact; and GDP, consumption, and investment subsequently converging to their new, higher, steady-state values. Once again, for either variable the responses to news and noise shocks at  $t=0$  are identical. The IRFs of GDP, consumption, and investment to noise shocks fade away to zero quite rapidly, whereas those to news shocks converge to the new steady-state values. As for hours, the response to noise shocks quickly fades away to zero, whereas that to news shocks, although it ultimately also converges to zero, exhibits a strong hump-shaped pattern as in Barsky and Sims (2011). Finally, in response to non-news shocks GDP, consumption, investment and hours all increase on impact.

Figure I.1 in the online appendix<sup>10</sup> shows the fractions of FEV explained by non-news, news and noise shocks. Key points to stress are that (i) noise shocks explain almost nothing of the FEV of any variable at all horizons; (ii) non-news and news

---

<sup>10</sup>The online appendix can be found at Luca Benati’s webpage, at: <https://sites.google.com/site/lucabenatiswebpage>.

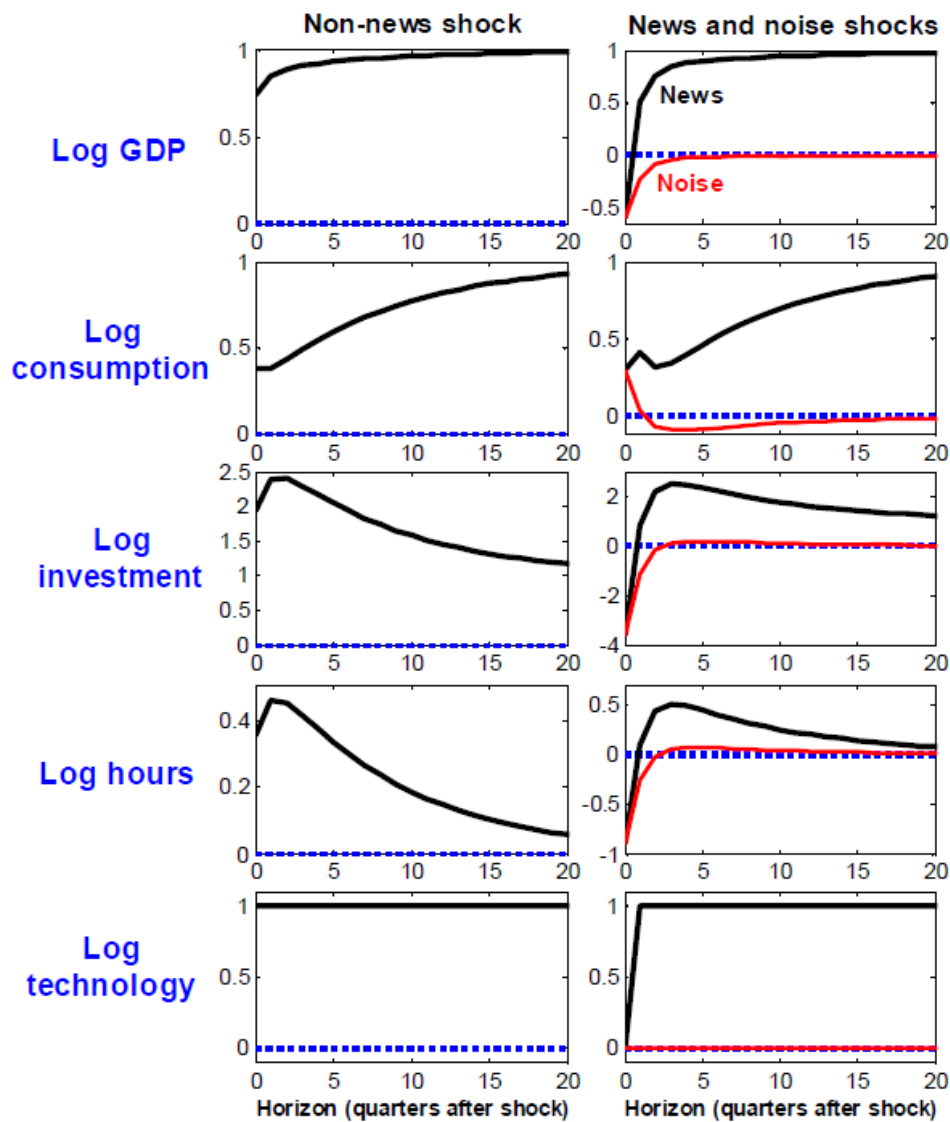


Figure 2 Impulse-response functions to non-news, news and noise shocks for Barsky and Sims' (2011) model augmented with noise shocks about TFP

shocks explain, at long horizons, about 50 per cent each of the FEV of log TFP, and they explain non-negligible-to-sizeable fractions of the FEV of GDP, consumption, and investment; and (iii) none of the three shocks explain an appreciable amount of the FEV of hours.

### 3.1 Reduced-rank structure of the matrix of structural impacts at $t=0$

It can be seen immediately that the matrix of structural impacts at  $t=0$  has a reduced rank. In particular, with two signal-extraction problems (news *versus* noise, and non-news *versus* transitory) there are two pairs of columns (i.e., of impulse vectors at  $t=0$ ) which are proportional to each other: The column generated by noise is proportional to the one generated by news, and the same holds true for the two impulse vectors generated by non-news and transitory.<sup>11</sup> This implies that, within the present context, the matrix of the structural shocks' impacts at  $t=0$  has rank 3. We now turn to the issue of whether low-order SVARMAs can well approximate the main features of the model.

### 3.2 Approximating the theoretical model based on low-order SVARMA( $p,q$ )'s

Figure 3 compares the theoretical IRFs to non-news, news, and noise shocks generated by the model (i.e., the very same IRFs shown in Figure 2) to those produced by either its truncated theoretical SVAR(1) and SVAR(2) representations, or its truncated theoretical SVARMA(1,1) representation.<sup>12</sup> Since, as can be trivially shown, the model possesses a near-exact VARMA(1,1) representation,<sup>13</sup> a truncated VARMA(1,1) recovers the true IRFs almost perfectly.<sup>14</sup> Interestingly, although VARs can never recover, even approximately, the IRFs to noise shocks,<sup>15</sup> a VAR(2) can approximate remarkably well the IRFs to both non-news and news shocks.

---

<sup>11</sup>Within the context of the present-value model for dividends and stock prices of Section 2.1 this was illustrated in the first panel of Figure 1.

<sup>12</sup>The truncated theoretical SVAR and SVARMA representations of the RBC model have been computed *via* linear projections methods starting from the RBC model's structural MA representation, which can immediately be recovered from the model's IRFs (see Appendix D).

<sup>13</sup>In the sense that all MA matrices after the first one have very small entries.

<sup>14</sup>Qualitatively the same evidence holds for the fractions of FEV: We do not report this evidence for reasons of space, but it is available upon request.

<sup>15</sup>The problem is strictly conceptual: Within a VAR framework, there is no way to break the one-to-one mapping at all horizons between the IRFs to news and noise shocks, originating from the fact that they share the very same impulse vector on impact. As a result, a VAR automatically forces the IRFs to news and noise shocks to be identical at *all* horizons. But since the two sets of IRFs are identical only on impact, and then diverge at longer horizons, it is logically impossible for a VAR to correctly recover the true IRFs. This implies that increasing the VAR's lag order cannot solve the problem.

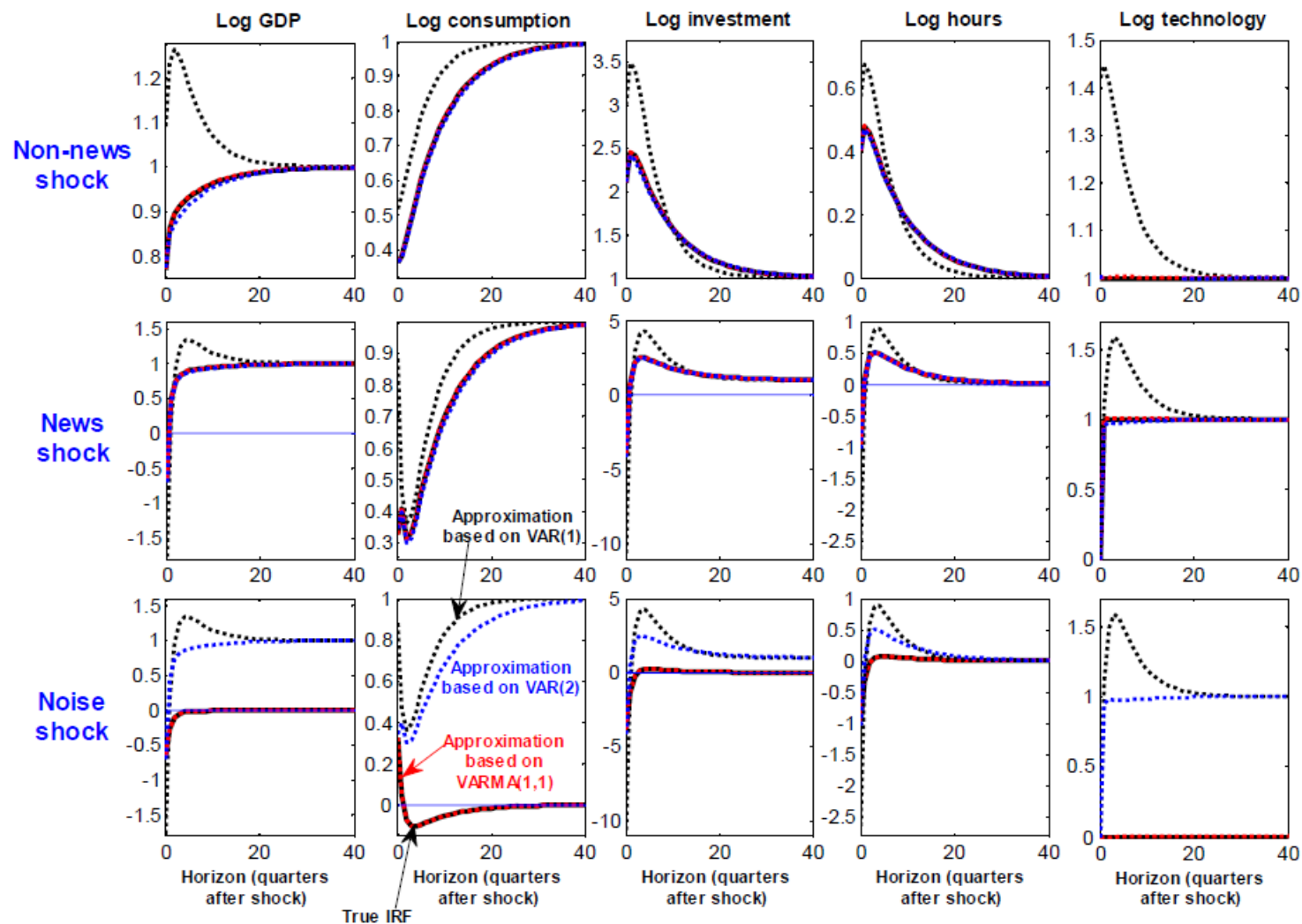


Figure 3 Approximating the IRFs of Barsky and Sims' (2011) RBC model augmented with noise shocks by means of theoretical SVAR(MA) representations

Although these results have been obtained based on a fairly simple DSGE model, at the very least they suggest that, in practice, structural VARMA models with a small MA lag order might be sufficient to accurately approximate the main features of the underlying DGP.

### 3.3 Identifying restrictions

Motivated by the previous discussion, in our empirical work we will impose the following restrictions in order to identify non-news, news, and noise shocks (in Section 4.2 we will show that these restrictions allow indeed to exactly recover the three shocks in population):

(1) At  $t=0$ , the variable of interest—to fix ideas, TFP—is impacted upon by only two disturbances, the non-news shock, and a transitory shock which is disentangled from the non-news shock because it explains the minimum fraction of the FEV of TFP at a specific long horizon (which we will take to be 20 years ahead). The reason for departing slightly from the corresponding restriction used by Barsky and Sims (2011)—who identify the non-news shock as the reduced-form innovation in TFP—is discussed below.

(2) The news shock is the one which, among all of the remaining shocks, explains the maximum fraction of the FEV of TFP at a long horizon (again, 20 years ahead).

In order to isolate the noise shock from the remaining disturbances, we impose the following restriction:

(3) The impulse vectors generated on impact by news and noise shocks are proportional to each other.

Restriction (1) differs from the corresponding restriction used by Barsky and Sims (2011),

(1') the non-news shock is identified as the reduced-form innovation to TFP,

by allowing for an additional transitory shock to impact upon TFP at  $t=0$ . Our preference for scheme (1)-(3), over the alternative scheme defined by (1'), (2), and (3), originates from the fact that the alternative scheme suffers from the following problem. As the previous discussion of the RBC model augmented with noise shocks made clear, for these shocks to play any role it ought to be the case that the permanent component of TFP is not perfectly observed, which is obtained by having a white noise shock impacting upon TFP at  $t=0$ . By postulating that the non-news shock is the only one impacting upon TFP at  $t=0$ , on the other hand, Barsky and Sims' (2011) assumption (1') implicitly postulates that the permanent component of TFP is, in fact, perfectly observed, which rules out the possibility that noise shocks might play a role. Because of this, in what follows we will report and discuss results based on scheme (1)-(3).

Although these restrictions (together with a technical restriction given in (23) and discussed in Section 5.1) are sufficient to identify the three shocks, in what follows we will also consider imposing additional restrictions, in order to get more precise

estimates. We therefore conduct an exercise in the spirit of Canova and Paustian (2011) in order to derive a set of robust restrictions pertaining to news and noise shocks, where ‘robust’ means ‘holding for an overwhelmingly large fraction of plausible random combinations of the model’s parameters’.

We consider the following sets of plausible values for most of the model’s structural parameters:  $b$ :  $[0, 0.99]$ ;  $\delta$ :  $[0.01, 0.08]$ ;  $\eta$ :  $[0.1, \infty)$ <sup>16</sup>;  $\gamma$ :  $[0.01, 0.1]$ ;  $\bar{g}$ :  $[0.15; 0.25]$ ;  $\tilde{g}_A$ :  $[0; 0.04]$  and we set the standard deviations of news and noise shocks to  $\sigma_{NE}=\sigma_u=1$ . The remaining parameters are calibrated as in Barsky and Sims (2011) as described in online Appendix C. Following Canova and Paustian (2011), we take 100,000 draws for the parameters from Uniform distributions defined over these intervals. For each draw of the parameters, we solve the model and compute IRFs to news and noise shocks and the fractions of the FEV of either variable explained by the two shocks.

Table 1: Proportion of Drawn Parameters which Satisfy the Restriction: News Shock IRFs Greater than Noise Shock IRFs (in absolute value)					
<i>Periods after impact</i>	TFP	GDP	Consumption	Investment	Hours
1	1.00	1.00	1.00	0.71	0.98
2	1.00	1.00	1.00	1.00	1.00
3	1.00	1.00	1.00	1.00	1.00
4	1.00	1.00	1.00	1.00	1.00
6	1.00	1.00	1.00	1.00	1.00
8	1.00	1.00	1.00	1.00	1.00

Table 1 reports the proportion of drawn parameters which produce IRFs which satisfy one possible restriction of interest: That IRFs to news shocks are, in absolute value, greater than those to noise shocks.<sup>17</sup> These proportions turn out to be identical to the proportion of draws for which the percentage of the FEV explained by news shocks is greater than that explained by noise shocks. There is strong support for this restriction for all variables and at all horizons. For horizons of two or more this support becomes overwhelming. Accordingly, in our empirical work some of our specifications will impose the restriction that news shocks have larger IRFs (in absolute value) than noise shocks two periods after impact.<sup>18</sup>

We now turn to addressing the issue of whether news and noise shocks can in fact be separately identified.

<sup>16</sup>To be precise, we set the upper value of the interval to be  $\frac{1}{\varepsilon}$  where  $\varepsilon$  is MATLAB’s definition of a small number which is equal to 2.2204e-16.

<sup>17</sup>We do not report results for  $t=0$  because on impact the two sets of IRFs are identical.

<sup>18</sup>This statement assumes that news and noise shocks have the same standard deviation. A slight adjustment is required when they have different standard deviations which is described in section 5.



## 4 Can News and Noise Shocks Be Disentangled?

The possibility of separately identifying news and noise shocks has been recently called into question by CJ (2018), who have argued that the two shocks are, in fact, observationally equivalent, and as such cannot be disentangled. In this section we briefly summarize the main results from a companion note—see Benati, Eisenstat, and Koop (2018)—in which we show that CJ’s observational equivalence result holds only under a specific assumption about what the econometrician can observe and does not hold in more general (and more realistic) frameworks. We make three main points:

(I) CJ’s result about observational equivalence crucially hinges on their casting the analysis from the perspective of the joint distribution of fundamental variables and *agents’ expectations* about them. We demonstrate that observational equivalence breaks down if the analysis is instead cast from the perspective of the joint distribution of fundamental variables and *signals* about them.<sup>19</sup> This implies that since *agents* are typically assumed to observe the signal, they should be able to disentangle news and noise shocks.

(II) The fact that econometricians do not observe the signal does not imply that observational equivalence of news and noise shocks holds from the econometricians’ perspective. The reason for this is straightforward. Although econometricians do not observe the signal, they observe ‘what the signal does to the economy’. Within any relevant macroeconomic model—e.g., Blanchard *et al.*’s (2013) model, or the RBC model with noise about TFP of Section 3—economic agents react to the signal, and their actions have an impact on the economy. This implies that macroeconomic variables contain information about the signal, and therefore about news and noise shocks. Within the RBC model of Section 3, for example, all variables (with the obvious exception of TFP itself) react to the signal, and therefore contain information about it. This is what a general equilibrium, rational expectations framework implies. The crucial question therefore becomes whether econometricians—armed with (i) the information about the signal contained in macroeconomic data, and (ii) appropriate identifying assumptions—can perform the same task as agents can, that is: Can they disentangle news and noise shocks? The answer to this question depends on the macroeconomic model the econometrician is working with, but for the sorts of models used in the noisy-news literature we argue that the answer is yes.

(III) We provide a simple illustration of this point based on the RBC model of Barsky and Sims (2011), which we augment with noise shocks about future TFP. Working in population, we show that the impulse response functions (IRFs) to news and noise shocks, and the fractions of forecast error variance (FEV) of the variables they explain, can be exactly recovered by imposing the model-implied restrictions that (i) news and noise shocks produce identical impulse vectors on impact, and (ii)

---

<sup>19</sup>Benati, Eisenstat, and Koop (2018) contains a general proof of this based on Hilbert spaces arguments.

news and non-news shocks are the only disturbances having a permanent impact on TFP.

We start by illustrating point (I) based on one of CJ's simple mathematical examples.

## 4.1 A simple theoretical example

Throughout the entire paper, CJ (2018) state their results about the observational equivalence of news and noise shocks from the perspective of the bivariate process for fundamentals and *agents' expectations* about them (i.e., within the context of the model of Section 3,  $a_t$ , and either  $a_{t|t-1}$  or  $a_{t+1|t}$ ). Based on a theoretical example from their paper,<sup>20</sup> we now show that, although news and noise shocks are indeed observationally equivalent when a researcher adopts that particular perspective, this is not true in general. In particular, we show that if a researcher adopts instead the perspective of the bivariate process for fundamentals and *signals*, the observational equivalence result disappears. In online Appendix H we show that the same holds true for all of CJ's (2018) theoretical examples.

Let

$$x_t = \mu_{t-1} + \eta_t \quad (10)$$

$$s_t = \mu_t + \xi_t \quad (11)$$

where  $x_t$  is the observed process,  $s_t$  is the signal, and  $\eta_t$ ,  $\mu_t$ , and  $\xi_t$  have the interpretation of non-news, news, and noise shocks, with  $\eta_t \sim N(0, \sigma_\eta^2)$ ,  $\mu_t \sim N(0, \sigma_\mu^2)$ , and  $\xi_t \sim N(0, \sigma_\xi^2)$ . We start by showing that, from a perspective of  $[x_t, x_{t|t-1}]'$ , news and noise shocks cannot be disentangled.

### 4.1.1 Observational equivalence conditional on $[x_t, x_{t|t-1}]'$

From (10) we have that  $x_{t|t-1} = \mu_{t-1|t-1}$ . By applying the Kalman filter to (11) we have

$$\mu_{t|t} = \underbrace{\frac{\sigma_\mu^2}{\sigma_\mu^2 + \sigma_\xi^2}}_{\kappa} [\mu_t + \xi_t] = \kappa s_t \quad (12)$$

so that

$$x_{t|t-1} = \kappa [\mu_{t-1} + \xi_{t-1}] \quad (13)$$

The process  $[x_t, x_{t|t-1}]'$  is fully characterized by the following non-zero moments:<sup>21</sup>

$$\text{Var}[x_t] = \sigma_\mu^2 + \sigma_\eta^2 \quad (14)$$

$$\text{Cov}[x_t, x_{t|t-1}] = \text{Var}[x_{t|t-1}] = \frac{\sigma_\mu^4}{\sigma_\mu^2 + \sigma_\xi^2} \quad (15)$$

<sup>20</sup>The example is from Section 3.2, pp. 15-16.

<sup>21</sup>All other moments are equal to zero.

With three parameters ( $\sigma_\mu^2$ ,  $\sigma_\eta^2$ , and  $\sigma_\xi^2$ ) and two equations, the system is not identified, which is another way of stating that the shocks cannot be separately disentangled.

#### 4.1.2 Observational non-equivalence conditional on $[x_t, s_t]'$

Consider instead what happens if we use  $[x_t, s_t]'$  as the data. In this case, the system is identified, and news and noise shocks are no longer observationally equivalent. With  $x_t$  and  $s_t$  given by (10) and (11), the system  $[x_t, s_t]'$  features three distinct non-zero moments, instead of two. The first one is still given by (14) and the other two are given by

$$\text{Var}[s_t] = \sigma_\mu^2 + \sigma_\xi^2 \quad (16)$$

$$\text{Cov}[x_t, s_{t-1}] = \sigma_\mu^2 \quad (17)$$

Equation (17) directly provides  $\sigma_\mu^2$ . Given  $\sigma_\mu^2$ , equations (14) and (16) then allow for the computation of  $\sigma_\eta^2$  and  $\sigma_\xi^2$ .

#### 4.1.3 Discussion

This simple example illustrates how the possibility of separately identifying news and noise shocks crucially hinges on what the researcher regards as the data. Within the present context, the two shocks are indeed observationally equivalent from the perspective of  $[x_t, x_{t|t-1}]'$ , but they are not from the perspective of  $[x_t, s_t]'$ . Rather, based on data for  $[x_t, s_t]'$ , the shocks' IRFs and fractions of FEV can be exactly recovered in population. As shown in online Appendix H, the same holds true for CJ's more complex models. Finally, Benati, Eisenstat, and Koop (2018) contains a general proof (based on Hilbert spaces arguments) that conditional on observing the signal news and noise shocks are not observationally equivalent.

This has two important implications. First, since agents are typically assumed to observe the signal, in principle they should be able to disentangle news and noise shocks. Second, although econometricians do not observe the signal, in fact they do observe the economy's response to it, as encoded in the reaction of macroeconomic variables such as GDP, consumption, investment, and hours, which are all optimally chosen by agents conditional on all information, including the signal itself. This implies that the distinction between economic agents (who observe the signal) and econometricians (who do not) is, for discussion of observational equivalence of news and noise shocks, irrelevant. The key point is that both of them have information about the signal, either because they directly observe it (the agents), or because they see how macroeconomic variables react to it (the econometricians). In turn this logically suggests that, based on the information on the signal contained in macroeconomic data, econometricians armed with appropriate identifying assumptions might well be able to disentangle news and noise shocks—exactly as agents, in principle, can. As we now

show based on the RBC model with noise, this is indeed the case, as our identifying restrictions allow to exactly recover the shocks' IRFs and fractions of explained FEV in population.

## 4.2 Recovering the IRFs and fractions of FEV in population

Let the theoretical structural MA representation of the RBC model's observables be

$$Y_t = A_0\epsilon_t + A_1\epsilon_{t-1} + A_2\epsilon_{t-2} + A_3\epsilon_{t-3} + \dots \quad (18)$$

where  $Y_t = [y_t, c_t, i_t, h_t, a_t]'$ , where  $y_t, c_t, i_t$ , and  $h_t$  are the logarithms of GDP, consumption, investment, and hours;  $\epsilon_t = [\epsilon_t^{NN}, v_t, \epsilon_t^{NE}, u_t, \epsilon_t^g]'$ , where the notation is as before;<sup>22</sup> and  $E[\epsilon_t\epsilon_t'] = I_N$ , with  $I_N$  being the  $(N \times N)$  identity matrix, so that each of the  $i$ -th columns (with  $i = 1, 2, \dots, N$ ) of the MA matrices  $A_0, A_1, A_2, A_3, \dots$  has been divided by the standard deviation of the  $i$ -th shock in  $\epsilon_t$ . Representation (18) can be immediately recovered from the model's IRFs to the structural innovations.

Observationally equivalent reduced-form representations of (18) can be obtained by post-multiplying all of the MA matrices  $A_0, A_1, A_2, A_3, \dots$  by an orthogonal rotation matrix  $R$ , yielding

$$Y_t = \tilde{A}_0\tilde{\epsilon}_t + \tilde{A}_1\tilde{\epsilon}_{t-1} + \tilde{A}_2\tilde{\epsilon}_{t-2} + \tilde{A}_3\tilde{\epsilon}_{t-3} + \dots \quad (19)$$

where  $\tilde{A}_j = A_jR$  and  $\tilde{\epsilon}_{t-j} = R'\epsilon_{t-j}$ ,  $j = 0, 1, 2, 3, \dots$ . We can randomly generate different rotation matrices<sup>23</sup>, thus producing different observationally equivalent VMAs. The question we wish to address is whether imposing restrictions (1)-(3) from subsection 3.3 on different reduced form VMAs allow us to recover the true structural VMA.

We have performed this exercise 100 times, and for all of these, imposing the restrictions (1)-(3) produces IRFs and FEVs which match the true theoretical IRFs and FEVs.<sup>24</sup> We illustrate this in Figure 4 which contains results using one random

<sup>22</sup>Since we have five observables, in this section we set to 0 all of the shocks other than the five collected in the vector  $\epsilon_t$ .

<sup>23</sup>We generate the  $(N \times N)$  random rotation matrix as follows. We start by taking an  $(N \times N)$  draw  $K$  from an  $N(0, 1)$  distribution. Then, we take the  $QR$  decomposition of  $K$ , and we set the rotation matrix to  $Q'$ .

<sup>24</sup>We impose restrictions (1)-(3) *via* rotation matrices as follows. Restrictions (1)'s zero restrictions on impact are imposed *via* a Householder reflection. Restriction (2) is imposed *via* Uhlig's (2003, 2004) 'maximum FEV' approach. Proportionality between the impulse vectors generated on impact by news and noise shocks is imposed *via* numerical methods, based on the methodology proposed by Benati (2014, Section 3.2.2). In brief, the methodology is based on the notion of working with the entire set of possible rotation matrices associated with a given structural impact matrix, optimizing the relevant criterion function over the space of the associated rotation angles (which we do *via* simulated annealing). Section 3.2.3 of Benati (2014) presents Monte Carlo evidence on the excellent performance of this methodology.

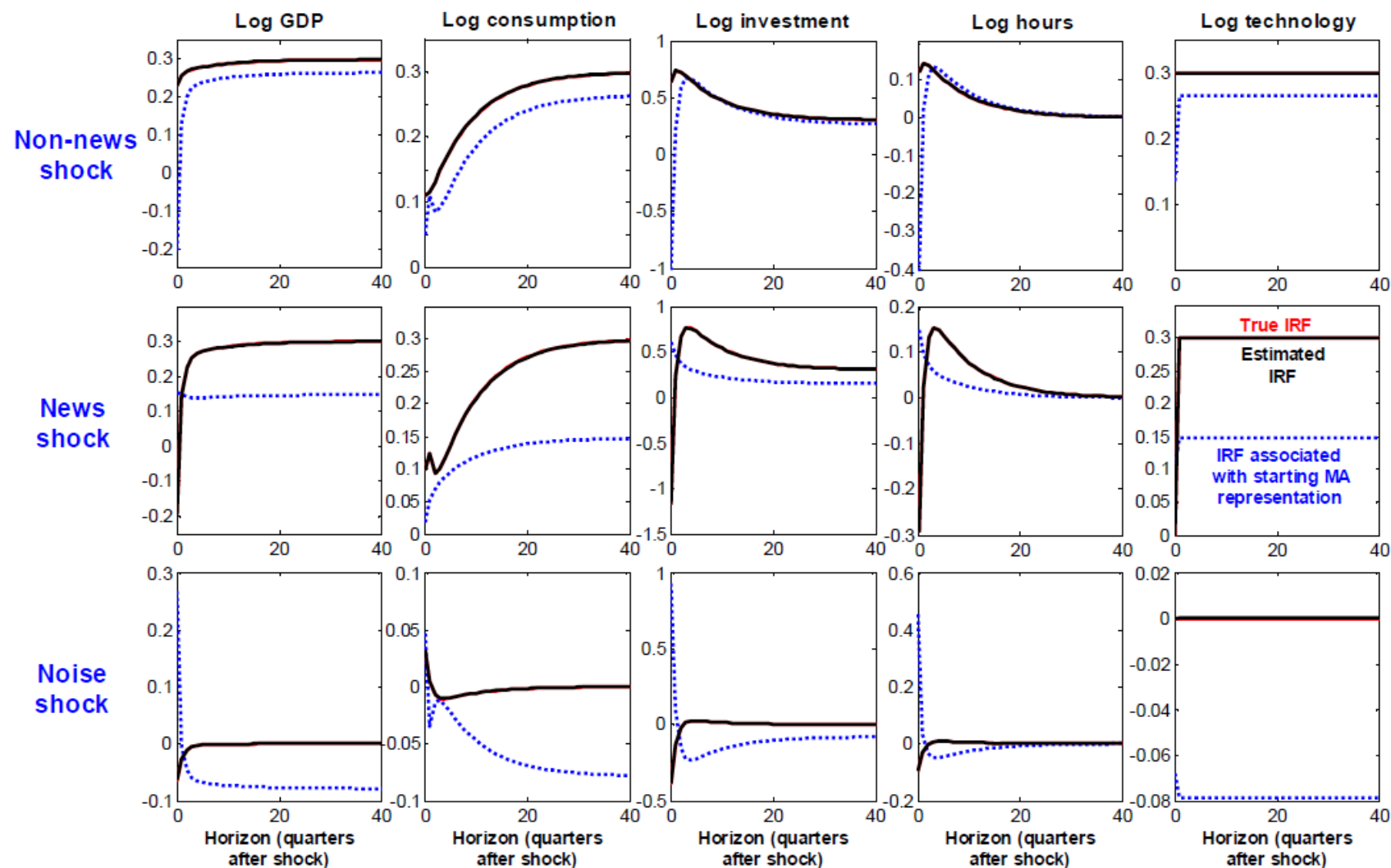


Figure 4 Assessing the identifying restrictions in population: Theoretical and estimated IRFs based on Barsky and Sims' (2011) RBC model augmented with noise shocks

rotation.<sup>25</sup> The figure shows, in red, the true IRFs as captured by the theoretical MA representation (18); in blue (dotted line) the IRFs associated with the initial, randomly rotated MA representation (19); and in black the estimated IRFs, i.e., the IRFs we recovered starting from (19), by imposing identifying assumptions (1)-(3).

In spite of the fact that the IRFs associated with the initial, randomly rotated MA representation were, in general, far away from the true IRFs, our identifying restrictions were able to recover them with great precision (in fact, the true and estimated IRFs are essentially indistinguishable). In the light of the previous discussion this should not be seen as surprising. A crucial difference between news and noise shocks is that whereas the former have a permanent impact on TFP (and GDP, consumption, and investment), the latter do not have a permanent impact on any variable. It is therefore impossible for the two shocks to be observationally equivalent.

#### 4.2.1 Approximating the MA representation with truncated VARMA( $p,q$ )

As we showed in sub-section 3.2, a truncated VARMA(1,1) produces theoretical IRFs and fractions of FEV which are essentially indistinguishable from those associated with the model's theoretical MA representation (18). This logically suggests that, working with the truncated VARMA(1,1), it should be possible to perform an exercise conceptually akin to that we have performed in the previous sub-section based on the RBC model's MA representation, that is: Recovering (near-exactly) news and noise shocks' IRFs and fractions of FEV starting from a reduced-form truncated VARMA(1,1) representation. This is in fact the case. We start from the truncated theoretical VARMA(1,1) representation of the model,

$$Y_t = BY_{t-1} + A_0\epsilon_t + A_1\epsilon_{t-1} \quad (20)$$

and we transform it, again, using randomly generated orthogonal matrices, thus obtaining observationally equivalent VARMA(1,1) representations

$$Y_t = BY_{t-1} + \tilde{A}_0\tilde{\epsilon}_t + \tilde{A}_1\tilde{\epsilon}_{t-1}. \quad (21)$$

We then impose identifying restrictions (1)-(3) and obtain IRFs and FEVs. We perform this exercise 100 times (and results for all repetitions are available on request) and always find that we are able to recover the model's true IRFs and FEVs with great precision. This is illustrated in Benati, Eisenstat, and Koop (2018) in Figure 3, and in Figure A.3 in that note's online Appendix, for the IRFs and the fractions of FEV, respectively. We do not report this evidence here for reasons of space, and because it is, in fact, near-identical to the evidence in Figure 4 we discussed in the previous sub-section.

---

<sup>25</sup>Figure 2 presents results for the IRFs. Results for the FEVs are shown in Figure I.2 in the online Appendix. In addition, we only report results for non-news, news, and noise shocks, but results for the other two shocks are available upon request. They are qualitatively the same to those presented here.

### 4.2.2 Alternative non-fundamental representations

Sub-section 4.2.2 of Benati, Eisenstat, and Koop (2018) also addresses the issue of the observational equivalence of fundamental and non-fundamental MA representations.<sup>26</sup> In brief, two main points ought to be mentioned here.

First, within the context of the RBC model with noise, which yields a non-fundamental representation by construction, the fact that many non-fundamental representations are admissible is not an issue. Imposing the model-implied identifying restrictions allows for the recovery of news and noise shocks for all of the non-fundamental representations of the truncated theoretical VARMA(1,1) representation (see Figure 4 there).

Second, concerning the observational equivalence of non-fundamental representations, there is also a more general, and more important point to be made. For models different from the RBC one we use, admissible non-fundamental representations (that differ by inverted *nonzero* and *finite* characteristic roots) may produce different results. However, this has no implications for the possibility of disentangling news and noise shocks—this possibility hinges solely on whether news and noise shocks are observationally equivalent conditional on a given reduced-form MA representation (or its approximate VARMA representation) regardless of how its characteristic roots are arranged. In other words, the problem of determining the appropriate static orthogonal rotations (to disentangle news and noise shocks from reduced-form innovations) is separate from the problem of determining the correct (nonzero/finite) root structure, which is invariant to static orthogonal rotations.

## 4.3 Summing up

CJ’s point that news and noise shocks are observationally equivalent is correct, if you frame the analysis the way that they do in terms of fundamentals and expectations. But we argue that empirical macroeconomists should not frame their analyses in this way. Within the context of the RBC model with noise about TFP, news and noise shocks are indeed observationally equivalent for an econometrician who based their analysis on TFP and agents’ expectations about it, i.e. on  $X_t = [a_t, a_{t|t-1}]'$ . But, as we have shown, news and noise shocks are not observationally equivalent if the data  $Y_t = [y_t, c_t, i_t, h_t, a_t]'$  are used, and the full set of structural implications used. That is, the very definition of what noisy-news is implies that news shocks have a permanent effect on TFP; noise shocks have no permanent effect on any variable; and news and noise shocks have identical effects on impact. The two shocks are not observationally equivalent once all of their implications for the observed data are taken into account.

---

<sup>26</sup>It is well known (see, e.g. Lippi and Reichlin (1994a)) that alternative representations obtained by “inverting” characteristic MA roots through Blaschke filters preserve the stochastic properties of a multivariate Gaussian times series.

## 5 Econometric Methodology

### 5.1 Justification for SVARMA

In this paper we use SVARMAs, instead of the SVARs which dominate the empirical macroeconomic literature. In general, several authors (e.g. Cooley and Dwyer (1998)) argue that theoretical macroeconomic models such as DSGE models lead to VMA representations which may not be well approximated by VARs, especially parsimonious VARs with short lag lengths. The fundamental concept underlying this argument was formalized by Fernandez-Villaverde *et al.* (2007), who characterize the equilibrium of a general class of DSGE models with linear transition laws in a state-space framework, and show how this can be represented by a finite-order VARMA. The structural VAR approach can then be justified as a truncated approximation to the infinite-order VAR obtained by inverting the MA part of the finite-order VARMA.

The inversion/truncation step, however, entails a number of difficulties. Some researchers—see, e.g., Chari *et al.* (2008) and Poskitt and Yao (2017)—point out that the errors which arise due to this can potentially be substantial, because the typical lag and series lengths used in practice (e.g.  $p < 5, T < 250$  for quarterly data) are much too short.

Moreover, in many cases such as the permanent-income model of Hansen *et al.* (1991), the resulting VARMA representation entails an MA component that is not invertible, and therefore, a VAR approximation to such a model does not exist.<sup>27</sup> VARMA representations admitting an infinite-order VAR representation are commonly known as fundamental, while those that do not are referred to as non-fundamental. The concept of fundamentalness is also related to a subtle, but important identification issue that arises in SVARs and SVARMAs, which we will return to shortly.

The presence of both news and noise shocks in our theoretical framework implies a non-fundamental VARMA representation (see the online Appendix for technical details). This implies that in our specific case it is not possible to impose our key identifying restrictions using an SVAR. An SVARMA is required.

To bring out the main ideas, we use an  $n$ -variate SVARMA(1, 1)

$$\mathbf{y}_t - \mathbf{B}_1 \mathbf{y}_{t-1} = \mathbf{u}_t = \mathbf{A}_1 \boldsymbol{\epsilon}_{t-1} + \mathbf{A}_0 \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim \mathcal{N}(0, \mathbf{I}_n). \quad (22)$$

Everything can easily be extended to a SVARMA( $p, q$ ) with deterministic terms. A simple way to describe the need for a VARMA in our setting is as follows.

Note first that the key aspect of identification relates to the IRFs for  $t=0$  which are in  $\mathbf{A}_0$  and that the error covariance matrix,  $\boldsymbol{\Omega}$ , is  $\boldsymbol{\Omega} = \mathbf{A}_0 \mathbf{A}_0' + \mathbf{A}_1 \mathbf{A}_1'$ . In Section 2, we derived results such as (7) which showed that news and noise shocks had an equal response on impact. In terms of the SVARMA, one might think this restriction

---

<sup>27</sup>Fernandez-Villaverde *et al.* (2007) provide an easily verifiable condition under which the VARMA admits an infinite-order VAR representation.



can be imposed by setting two columns of  $\mathbf{A}_0$  to be equal to one another. As already mentioned, this is not correct, since it would imply that the news and noise shocks have the same standard deviations. The correct way of imposing the equal  $t=0$  impacts of news and noise shocks on impact is to restrict two columns of  $\mathbf{A}_0$  to be proportional to one another with the factor of proportionality being equal to the ratio between the standard deviations of the two shocks.

If we impose the identification restriction that two columns of  $\mathbf{A}_0$  are proportional to one another, then  $\mathbf{A}_0$  will be a reduced-rank matrix. This precludes estimation in the SVAR case since, if  $\mathbf{A}_1 = 0$ ,  $\mathbf{\Omega}$  will be singular. This is why we need the SVARMA. To explain this another way, if we were to work with an SVAR and thus have  $\mathbf{A}_1 = 0$ , the fact that  $\mathbf{B}_1$  will be common to impulse responses at all horizons<sup>28</sup> means that two shocks generating a proportional impulse vector at  $t=0$  will also produce proportional impulse responses at all horizons, with the proportionality factor being exactly the same for all variables and all horizons. This means that within a SVAR framework it is impossible to disentangle news and noise shocks. In Section 3, we demonstrated the theoretical consequences of this within an RBC model. These considerations imply that, unless a researcher is willing to resort to a DSGE-based approach, the SVARMA approach is the only option. Within the SVARMA framework, this problem does not arise since the presence of  $\mathbf{A}_1$  breaks the one-to-one mapping between the impulse responses at  $t=0$  and responses over subsequent horizons.

Writing (22) in terms of polynomial matrices in the lag operator,  $\mathbf{B}(\mathbf{L})\mathbf{y}_t = \mathbf{A}(\mathbf{L})\boldsymbol{\epsilon}_t$ , a reduced-rank  $\mathbf{A}_0$  implies that  $\det \mathbf{A}(\mathbf{0}) = 0$ , i.e.  $\det \mathbf{A}(\mathbf{z})$ , for  $\mathbf{z} \in \mathbb{C}$ , has a root at zero. This means that  $\mathbf{A}(\mathbf{L})$  is not invertible on the closed unit disk, or equivalently, (22) is non-fundamental. As described in online Appendix B.2, the zero root can be ‘flipped’ using an appropriate Blaschke filter (see Lippi and Reichlin (1994b)) to obtain a reduced-form VARMA representation  $\mathbf{B}(\mathbf{L})\mathbf{y}_t = \boldsymbol{\Theta}(\mathbf{L})\boldsymbol{\epsilon}_t$ , in which  $\boldsymbol{\Theta}_0$  is full rank, but  $\boldsymbol{\Theta}_1$  is reduced rank. However, it is not possible to obtain a suitable finite-order VAR approximation to (22).

To summarize, the SVARMA approach allows us to identify non-news, news and noise shocks by imposing the restrictions in our identification scheme (1)-(3) described in Section 3.3. Restriction (3) is the key new restriction for identifying news and noise shocks. In terms of our SVARMA we impose this using the following restrictions (assuming the variable driven by non-news and news shocks, e.g. TFP, is ordered first in the vector of variables and non-news, news and noise shocks are ordered first, second, and third in the vector of shocks, and  $A_{0,ij}$  denotes the  $(ij)^{th}$  element, and  $A_{0,i}$  the  $(i)^{th}$  column of  $\mathbf{A}_0$ ):

- News and noise shocks do not affect TFP at  $t=0$ :  $A_{0,12} = A_{0,13} = 0$ .
- News and noise shocks have a proportional impact on all variables at  $t=0$ :  $A_{0,2} = cA_{0,3}$  for scalar  $c$ .

---

<sup>28</sup>By this we mean that the IRF is based on the VMA representation which involves inverting the VAR which puts  $\mathbf{B}_1$  in every lag of the VMA representation.

To ensure that the impulse responses derived from (22) are uniquely estimated, we also impose the technical restriction that all non-zero roots of  $\det \mathbf{A}(\mathbf{z})$  are outside the unit circle, formally given by

$$\det \mathbf{A}(\mathbf{z}) \neq 0 \text{ for all } 0 < |z| \leq 1. \quad (23)$$

This restriction is necessary because the fundamental and all non-fundamental VARMA representations obtained by flipping MA roots are observationally equivalent (Lippi and Reichlin (1994b)). By imposing (23), we are selecting one of the many (up to  $2^{nq}$ ) possible non-fundamental representations with one (real) root at zero from which to draw inference.

We justify this approach in our setting on three grounds. First, SVAR identification strategies generally rely on a similar restriction, namely imposing the fundamental representation with all MA roots outside the unit circle. Indeed, this was a key point of emphasis in Lippi and Reichlin (1994b), who argued that SVAR inference should be based on a set identification of impulse responses, where each of the fundamental and non-fundamental representations is explicitly considered. In practice, however, this is almost never done and instead SVAR approaches focus on point identification of impulse responses, which are obtained from the fundamental representation alone.

The avoidance of set identification in the SVAR literature is due mainly to the computational challenges involved in deriving impulse responses from as many as  $2^{nq}$  structural VMA representations that would need to be constructed following estimation of the VAR.<sup>29</sup> One complication arises from the fact that information about the MA roots is lost when the AR operator is truncated. Another is related to the fact that the number of representations grows quickly with the size of the system  $n$ , such that enumerating all possible representations becomes computationally infeasible for larger systems (e.g. with  $n = 15$  and  $q = 2$ , there are over 1 billion representations).

In our VARMA setting, restriction (23) can be viewed as the closest departure from the standard SVAR literature and, in this sense, produces results that are more comparable to those of related papers (e.g. Barsky and Sims (2011)).

Nevertheless, in spite of these arguments in favor of working with a point identified VARMA, we also check the robustness of our results by deriving set-identified IRFs and fractions of FEV for models with  $n \leq 8$  in our first application. In particular, for each draw of the parameters in our Bayesian approach, we compute the IRFs and the fractions of FEV for each representation  $i = 1, 2, \dots, K$ . Then, for each variable and each shock, we consider the lower and upper envelopes of the IRFs and of the fractions of FEV across all of the  $K$  representations. This yields draws from the posterior distributions for the lower and upper envelopes of the IRFs and fractions

---

<sup>29</sup>The exact number of non-fundamental representations depends on how many real vs. complex conjugate pairs of roots characterize  $\det A(z)$ , which is a maximum of  $nq$  if all roots are real.

of FEV.<sup>30</sup> In a nutshell, our main conclusions still hold based on the set-identified results, which constitutes further support for employing restriction (23).

Finally, we use this same approach to compute IRFs and fractions of FEV for all of the possible representations in our Monte Carlo study which takes, as DGP, the RBC models with noise shocks of Section 3 (see online Appendix E). We find overwhelming evidence that the representation satisfying (23) produces IRFs and fractions of FEV which closely match the true IRFs and fractions of FEV produced by the RBC model, whereas the other, non-fundamental representations systematically produce different results. We therefore conclude that, at least for the DSGE model considered in this paper, the corresponding VARMA should have all non-zero roots outside the unit circle.

Restriction (23), together with the three theoretical restrictions outlined in Section 3.2, uniquely identify, up to sign normalizations, non-news, news and noise shocks (see Section 4.2). Our main set of results for all of our applications imposes these restrictions along with (in some cases) the over-identifying restriction that, two periods after impact, the IRFs generated by the noise shock are smaller, in absolute value, than the corresponding IRFs generated by the news shock.<sup>31</sup>

## 5.2 Estimating the SVARMA

Despite having many attractive statistical properties, VARMA are rarely used in practice, since they can be over-parameterized, and identification and computation can be difficult. That is, even VARs are often over-parameterized. Ignoring deterministic terms, a VAR( $p$ ) will have  $pn^2$  VAR coefficients, which can be large, even with the moderately sized VARs often used in the structural macroeconomics literature (e.g. Barsky *et al.*, 2014, use a VAR with  $n=9$ ). Adding a VMA component increases this to  $(p+q)n$  making over-parameterization concerns even more relevant. With structural SVARs, there is a need for structural identification restrictions. With SVARMA there is the same need, but even the reduced form VARMA suffers from a lack of identification which arises from the possibility of common factors in the VAR and VMA parts of the model. These issues are discussed in detail in Chan *et al.* (2016). A crucial aspect of this approach is the use of prior shrinkage to mitigate over-parameterization concerns. Chan *et al.* (2016) also develops Bayesian methods for estimating VARMA which work even with large values of  $n$ . We adapt these methods to the SVARMA described in the preceding sub-section, resulting in an ef-

---

<sup>30</sup>The reason for focusing on the lower and upper envelopes of the IRFs and fractions of FEV across all of the  $K$  representations is that these are well-defined objects, for which it is possible to meaningfully define a posterior distribution. This allows us to characterize IRFs and fractions of FEV taking into account of the fact that each draw from the posterior distribution of the parameters of the VARMA is potentially associated with more than one representation satisfying the restrictions.

<sup>31</sup>As previously mentioned, a slight adjustment has to be made when we allow news and noise shocks to have different standard deviations and, hence, the restriction imposed is actually  $|\text{IRF}(\text{news})| > |\text{IRF}(\text{noise})/c|$ .

efficient Markov Chain Monte Carlo (MCMC) algorithm. Complete details of how we carry out Bayesian estimation are given in online Appendix B. Here we provide a basic outline of our strategy.

As discussed in Chan *et al.* (2016), estimation of VARMA models is more easily done when they are put in so-called expanded form. The expanded form writes the VARMA as a state-space model and standard Bayesian Markov MCMC algorithms for state-space models can be used, greatly simplifying computation. In online Appendix B, we show how the SVARMA written in (22), with the identification restrictions imposed on  $A_0$ , can be written in equivalent form as an expanded-form VARMA. We use a prior designed to ensure shrinkage and parsimony in the potentially over-parameterized SVARMA. Several different priors are popular in the Bayesian VAR literature, including the Minnesota prior and various hierarchical shrinkage priors (e.g. various LASSO priors, spike-and-slab priors). In this paper, we use the SSVS prior introduced to the Bayesian VAR literature by George *et al.* (2008), and used in many VAR papers.<sup>32</sup> The basic idea can be explained in terms of a generic VAR or MA coefficient, say  $\theta$ . SSVS specifies a hierarchical prior (i.e., a prior expressed in terms of parameters which in turn have a prior of their own) which is a mixture of two Normal distributions:

$$\theta|\gamma \sim (1 - \gamma) N(0, \kappa_0^2) + \gamma N(0, \kappa_1^2), \quad (24)$$

where  $\gamma \in \{0, 1\}$  is an unknown parameter. If  $\gamma = 0$  the prior for  $\theta$  is given by the first Normal distribution, and if it is  $\gamma = 1$  its prior is given by the second. The prior is hierarchical since  $\gamma$  is treated as an unknown parameter which is estimated in a data-based fashion. The first prior variance,  $\kappa_0^2$ , is chosen to be ‘small’ (so that the coefficient is constrained to be virtually equal to zero) and the second prior variance,  $\kappa_1^2$ , to be ‘large’ (implying a relatively non-informative prior for the corresponding coefficient). Thus, SSVS allows for the data to decide which coefficients should be set to zero, so as to ensure parsimony in the SVARMA. The only subjective prior information that is required is the choice of  $\kappa_0^2$  and  $\kappa_1^2$ , but standard methods exist for their choice. Details of how this is done and the MCMC algorithm which results from use of the SVARMA model with SSVS prior are given in online Appendix B.

We now turn to empirical applications of our econometric methodology and identifying restrictions.

## 6 Empirical Applications

This section contains three different empirical applications involving standard US quarterly macroeconomic variables. The online appendix provides exact definitions and data sources. The sample period goes from 1954Q3 (when the Federal Funds

---

<sup>32</sup>E.g., among many others, Koop (2013), and Korobilis (2013).

rate first becomes available) to 2008Q3 (so that we exclude the period during which the Funds rate has been, in practice, at the zero lower bound).

## 6.1 News and noise shocks to TFP

In this sub-section we provide evidence on the impact upon the economy of TFP non-news, news and noise shocks.

### 6.1.1 Data and specification choices

Our data set includes TFP and several other macroeconomic variables. We estimate SVARMA(4,1) with different VAR and MA lag lengths, and different choices for the macroeconomic variables other than TFP. In this section we report the results produced by a SVARMA(4,1) featuring 8 variables. In Appendices G and I we present additional results for other lag lengths, and choices of  $n$  between 6 and 15, as well as a model comparison exercise which provides evidence that the  $p=4$ ,  $q=1$ ,  $n=8$  choice is supported by the data. A key finding is that results are uniformly robust across the different values of  $n$ ,  $p$  and  $q$ . This is the case, in particular, for one of our main findings, that is, the overall minor-to-negligible role played by noise shocks in driving macroeconomic variables.

Our main results use the following eight variables: log TFP, log hours *per capita*, the Federal Funds rate, GDP deflator inflation, the logarithms of real GDP, consumption, and investment *per capita*, and the spread between the 5-year government bond yield and the Federal Funds rate.<sup>33</sup> The full set of 15 variables (all of which are standard in the literature) is listed in online Appendix F.

We estimate all models in (log) levels. The main reason for estimating the models in levels has to do with robustness with respect to cointegration of unknown order. As discussed by Hamilton (1994), estimating the system in levels is the robust thing to do in this case. In recent years, estimating VARs in levels has become standard practice in the macroeconomic literature.<sup>34</sup> All of the results we report are based on imposing identifying restrictions (1)-(3), outlined in sub-section 3.3. As for the restrictions on the absolute values of the IRFs to news and noise shocks, we report results either imposing, or not imposing this restriction two quarters after impact.

---

<sup>33</sup>Following common practice, for TFP we use John Fernald's purified TFP series. See Fernald (2012). The notion that productivity may contain a transitory component which exhibits some persistence, originally introduced by Blanchard *et al.* (2013) in order to rationalize the existence of a signal-extraction problem, requires some discussion. In univariate models, the log of TFP is typically found to be close to a random walk. But a transitory component can be justified in terms of data revision. That is, even if the true TFP is a random walk (driven by both news and non-news shocks), most of the data underlying the construction of TFP are progressively revised and this will introduce a persistent, but transitory, component to the estimate of TFP. Alternatively, Blanchard *et al.* (2013) showed that, in principle, it is possible to construct a process for TFP with a permanent and a transitory component which looks like a random walk in the univariate sense.

<sup>34</sup>See, e.g., Barsky and Sims (2011) or Kurmann and Otrok (2013).

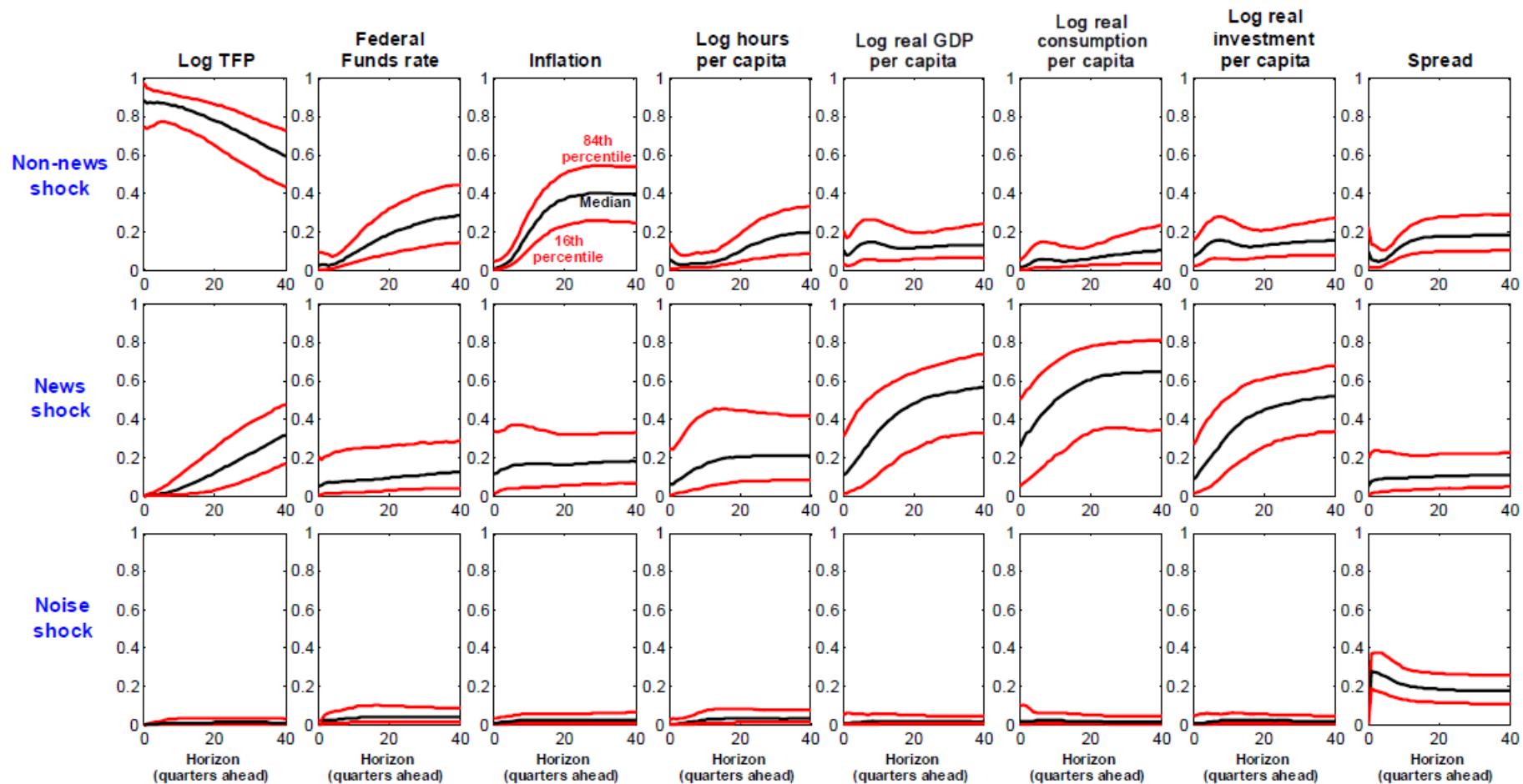


Figure 5 Application with TFP: Fractions of forecast error variance explained by non-news, news, and noise shocks, based on a VARMA(4,1), and point-identification

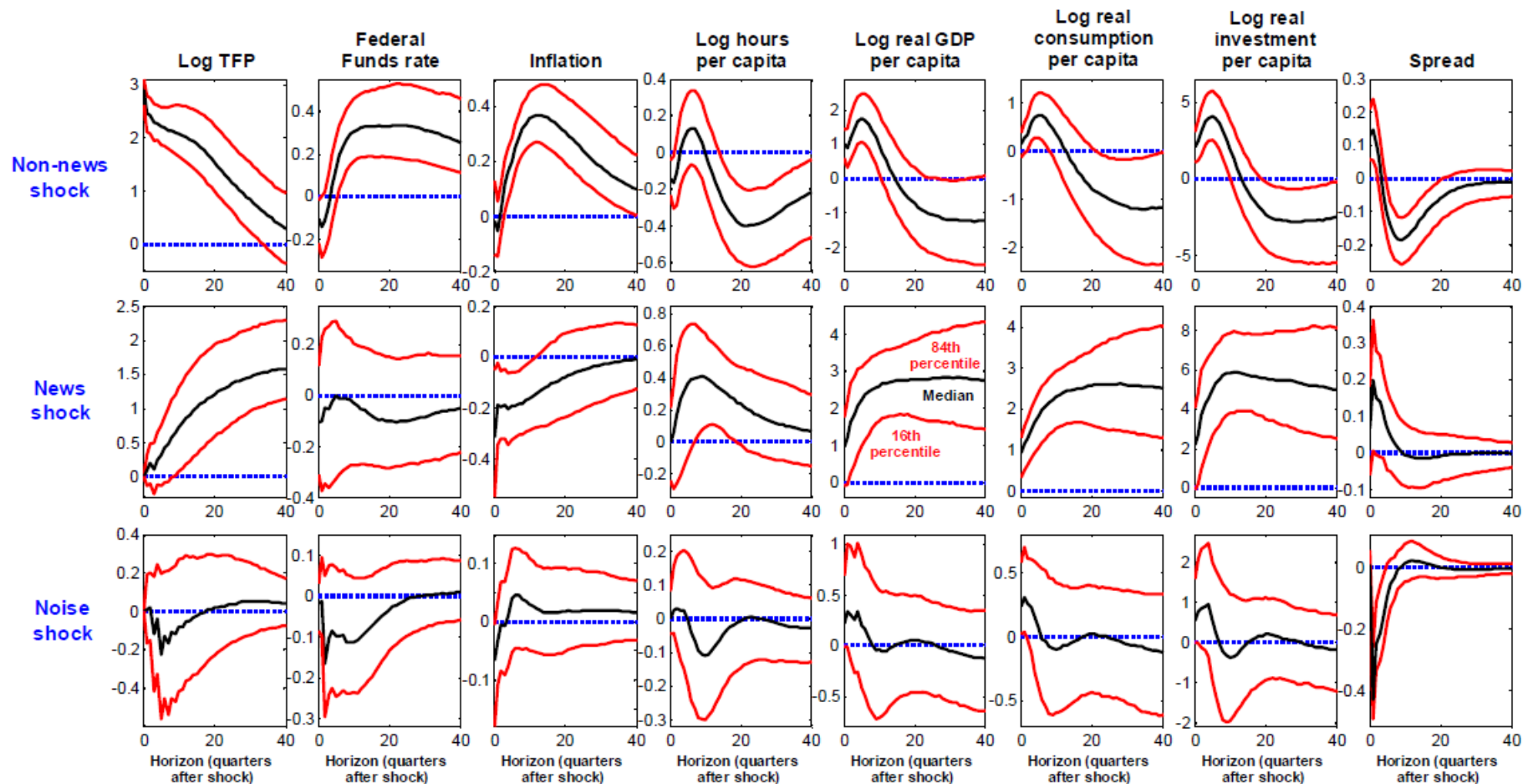


Figure 6 Application with TFP: Impulse-response functions to non-news, news, and noise shocks, based on a VARMA(4,1), and point-identification

The online appendix presents results for a range of choices of  $p$ ,  $q$  and  $n$ . For each specification, we present evidence on the convergence properties of the MCMC algorithm. In particular, for the draws from VARMA's reduced-form parameters, we present the draws' first autocorrelations and the inefficiency factors,<sup>35</sup> which we use in order to assess the convergence of the Markov chain. The first autocorrelations are uniformly very low, whereas the inefficiency factors are typically around one, which is much lower than the 20-25 value which is typically taken as signalling problems in convergence.

### 6.1.2 Evidence

Figures 5 and 6 show, for the baseline 8-variables system, the fractions of FEV explained by non-news, news, and noise shocks, and the IRFs to the three shocks, based on *point identification*, whereas Figures 7 and 8 show the corresponding results based on *set identification* (i.e., taking into account of all of the possible representations obtained by 'flipping the roots'). Figures VII.2-VII.3 in the online appendix report the corresponding results obtained by imposing the additional restriction on the comparative magnitudes of the IRFs to news and noise shocks. Overall, results based on set identification are qualitatively the same as those based on point identification. Further, imposing or not imposing the restrictions on the magnitudes of the IRFs to news and noise shocks does not make any material difference. Accordingly, in what follows we will focus our discussion on the results based on point identification.

Our main substantive finding pertains to the negligible role played by TFP noise shocks in U.S. macroeconomic fluctuations: As shown in Figure 5, these shocks explain small-to-negligible fractions of the FEV of all series at all horizons. This finding is in contrast to those of Blanchard *et al.* (2013) and Forni *et al.* (2017), whereas it is in line with the similar findings of Barsky and Sims (2012). In Figure 6 it can be seen that all of the IRFs to noise shocks have credible intervals that include zero at all horizons. The fact that noise shocks play such a small role, and have so little impact on any of the variables, is the main substantive point we wish to emphasize. Figures 5 and 6, however, show other interesting patterns which we discuss in the remainder of this sub-section.

News shocks explain small fractions of the FEV of TFP at short horizons, but their importance progressively increases further out: E.g., at the 10 year horizon they explain around 40 per cent of the FEV. News shocks also play a very important role for GDP, consumption and investment, especially at long horizons. For the remaining variables, they play a smaller but non-negligible role (e.g., based on median estimates, around 20 per cent of the FEV).

---

<sup>35</sup>The inefficiency factors are defined as the inverse of the relative numerical efficiency measure of Geweke (1992),  $RNE = (2\pi)^{-1} \frac{1}{S(0)} \int_{-\pi}^{\pi} S(\omega) d\omega$ , here  $S(\omega)$  is the spectral density of the sequence of draws from the Gibbs sampler for the quantity of interest at the frequency  $\omega$ . We estimate the spectral densities via the lag-window estimator as described in chapter 10 of Hamilton (1994). (We also considered an estimator based on the fast-Fourier transform, and results were very close.)



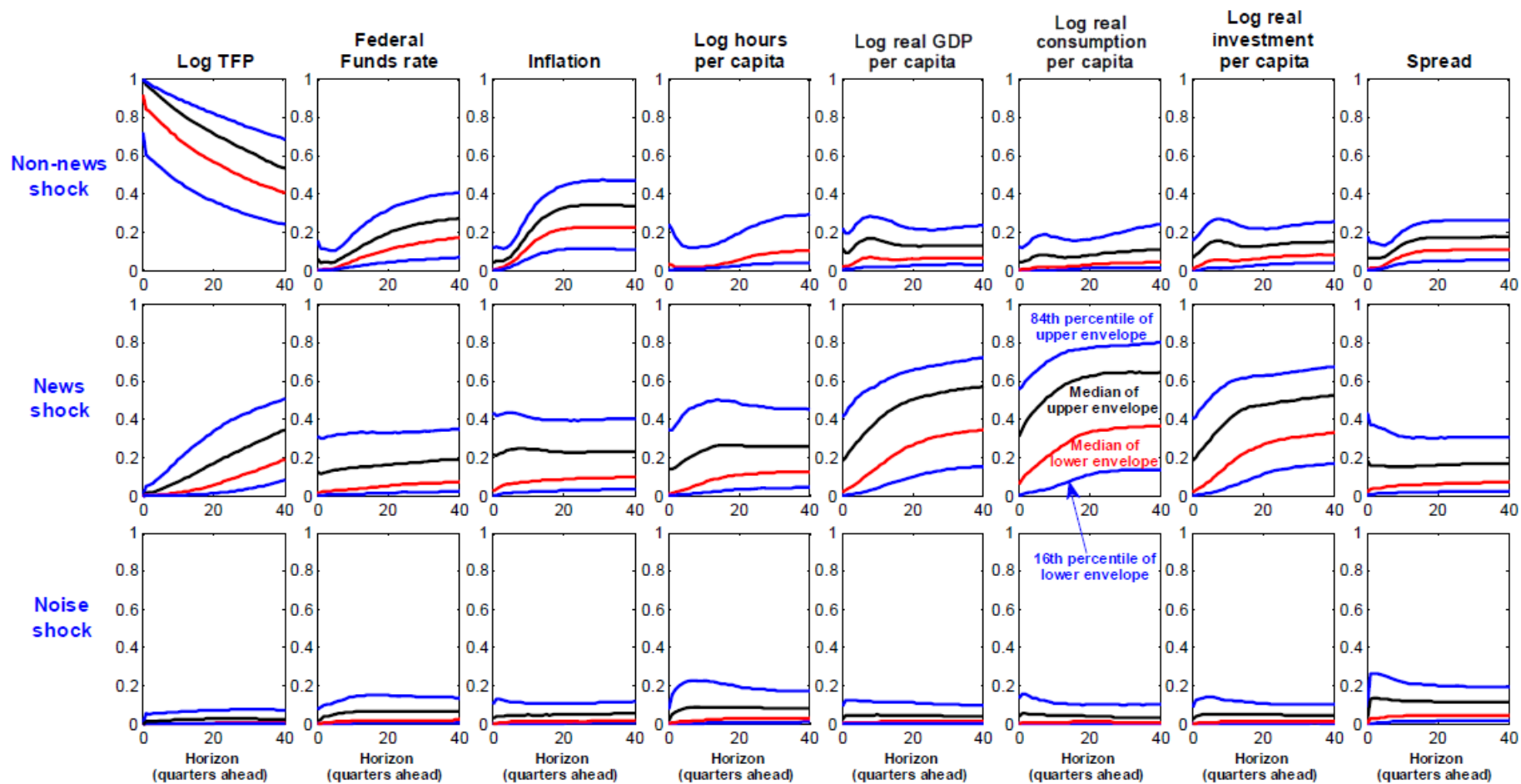


Figure 7 Application with TFP: Fractions of forecast error variance explained by non-news, news, and noise shocks, based on a VARMA(4,1), and set-identification

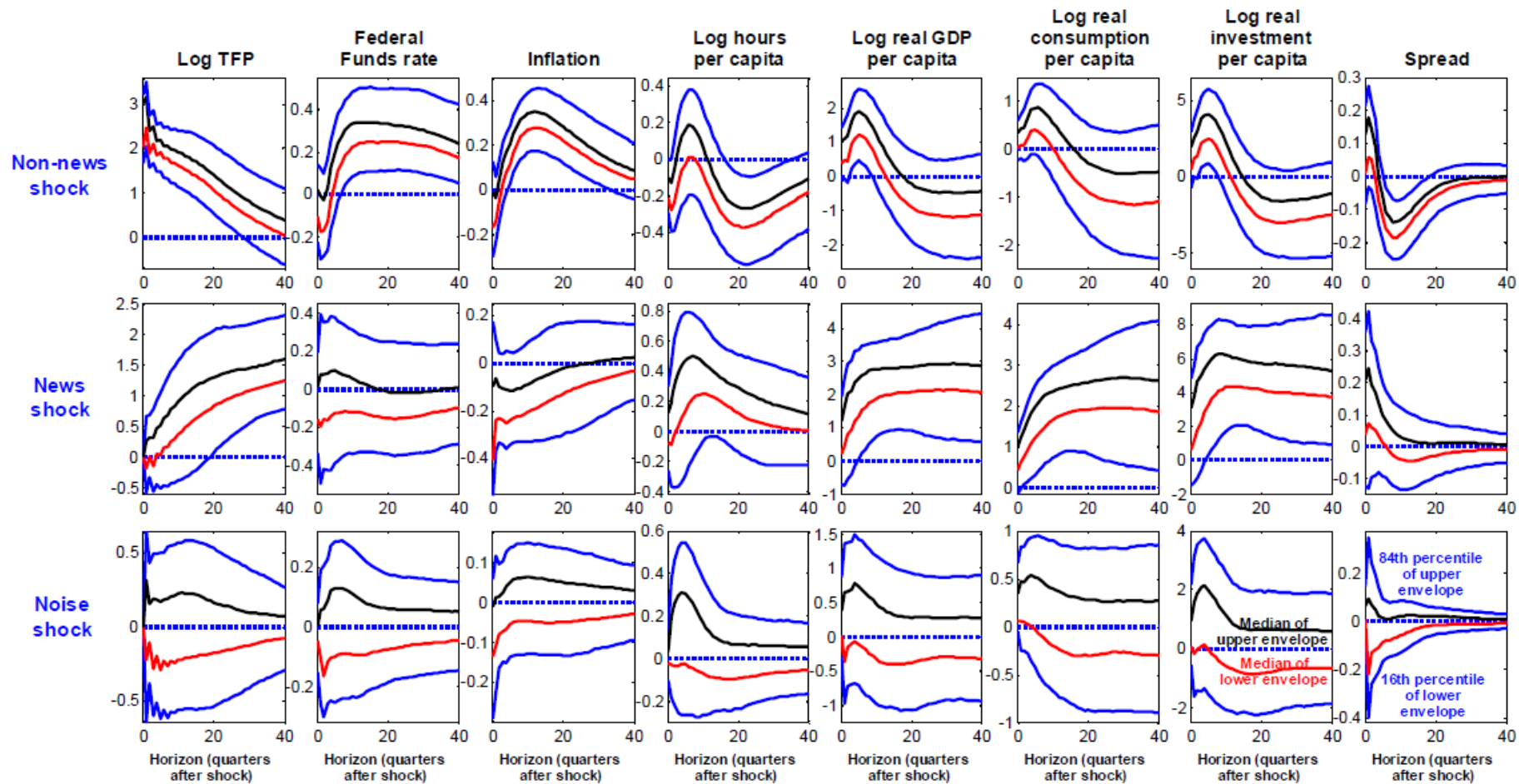


Figure 8 Application with TFP: Impulse-response functions to non-news, news, and noise shocks, based on a VARMA(4,1), and set-identification

Non-news shocks account for almost 100 per cent of the FEV of TFP on impact, but their importance progressively diminishes further out. However, even at the 10 year horizon they still account for roughly 40 per cent of the FEV. As for other series, these shocks tend to play a small role in terms of FEV. Only for inflation and, to a lesser extent, the Funds rate, do they play a non-negligible role (e.g., at longer horizons, median estimates indicate they explain roughly 20-30 per cent of the FEV).

The IRFs of TFP to news and non-news shocks are in line with the previous literature—see e.g. Barsky and Sims (2011)—with log TFP not jumping, and jumping, respectively, on impact, and then slowly converging to its new long-run value. An important point to stress is that, exactly as in Barsky and Sims (2011), the non-news shock is estimated to be transitory, whereas the news shock clearly has a permanent impact on TFP.

The IRFs for other variables are typically sensible. For non-news shocks, the responses of inflation and the Funds rate are similar to one another. As for news shocks, in line with Barsky and Sims (2011), there is some weak support for the notion that they have a negative impact on inflation at short-horizons.

The IRFs of GDP, consumption, and investment to news shocks mimic those of TFP. Their IRFs to non-news shocks have a different shape from that of TFP, but they are, as for TFP, ultimately transitory. The IRFs of hours to non-news and news shocks are different from those in Barsky and Sims (2011), but, as we will discuss in the next sub-section, they are in line with those produced by a structural VAR identified *via* the Barsky-Sims methodology based on this dataset. Finally, the impact response of the interest rate spread to news shocks is—as in Kurmann and Otrok (2013)—positive.

### 6.1.3 Comparison with Barsky and Sims (2011)

How do our results compare to those produced by Barsky and Sims (2011)? Since theirs is an SVAR approach, they can only estimate news and non-news shocks, but not noise shocks. As discussed in Section 3, however, for a sufficiently large lag order, VARs can provide an arbitrarily close approximation to the IRFs for news and non-news shocks produced by a model also featuring noise shocks. This means that a non-negligible discrepancy between our results and those obtained by Barsky and Sims (2011) would cast doubts on the reliability of our approach, and therefore the meaningfulness of our results. In this sub-section we therefore compare our results for news and non-news shocks to those obtained using an econometric model identical to our own but with no noise shock, and therefore  $\mathbf{A}_1 = 0$ .

Figure I.3 in the online appendix reports evidence on this, by showing the IRFs produced by our implementation of Barsky and Sims' (2011) approach for the TFP news and non-news shocks.<sup>36</sup> In order to make these results exactly comparable to

---

<sup>36</sup>We do not show the corresponding results for the fractions of FEV, both for reasons of space, and especially because they are qualitatively the same as those for the IRFs. They are however

those we discussed previously, they have been based on the same eight variables, VAR lag order and estimation sample as was used in the previous sub-section.<sup>37</sup> In line with Barsky and Sims (2011), the two disturbances are identified based on the restrictions that (i) the non-news TFP shock is the only shock which affects TFP on impact, and (ii) the news shock is the one which, among all of the remaining disturbances, explains the maximum fraction of the FEV of TFP at the 20 years ahead horizon. The main finding emerging from Figure I.3 is that the IRFs for the non-news and news shocks produced by the SVAR and by the SVARMA are uniformly very close, which clearly points towards the reliability of the SVARMA-based methodology proposed herein.

We now turn to an application to a model with dividends and stock prices.

## 6.2 Dividends and stock prices

One interesting feature of noise shocks is that, in principle, noise about future dividends might reconcile Fama’s results about market efficiency with Shiller’s findings of excess volatility in stock prices. The reason for this is straightforward, and it can be immediately grasped from the solution for log stock prices in (6). The key point is that, since  $S_t$  depends on noise shocks, a sufficiently large variance of these shocks could cause rationally determined stock prices to exhibit excess volatility compared to their fundamental value. An important *caveat* to this is that, for the reasons we discussed in Section 2.1.2 (see in particular the last panel of Figure 1), there might well be an upper limit to the fraction of FEV of any variable which is explained by noise shocks. Even if that is the case, however, such limit will be, in general, model-specific, so that in principle noise about future dividends might in fact reconcile the two strands of literature.

Figures 9 and XII.2 in the online appendix show results from a VARMA(4,1) for the logarithms of real dividends and real stock prices, and five other standard variables: log labor productivity, log hours *per capita*, the Federal Funds rate, GDP deflator inflation and the spread between the 5-year government bond yield and the Federal Funds rate. The identification scheme assumes that the news *versus* noise problem pertains to dividends. All specification choices are the same as in the previous application. The main substantive finding, which clearly emerges from Figure 9, is, once again, the minor-to negligible role played by noise shocks in explaining the FEV of any of the variables included in the system. It is also worth stressing that news

---

available upon request.

<sup>37</sup>The VAR is estimated *via* Bayesian methods as in Uhlig (1998, 2005). Specifically, Uhlig’s approach is followed exactly in terms of both distributional assumptions—the distributions for the VAR’s coefficients and its covariance matrix are postulated to belong to the Normal-Wishart family—and of priors. For estimation details the reader is therefore referred to either the Appendix of Uhlig (1998), or to Appendix B of Uhlig (2005). Results are based on 10,000 draws from the posterior distribution of the VAR’s reduced-form coefficients and of the covariance matrix of its reduced-form innovations (the draws are computed exactly as in Uhlig (1998, 2005)).

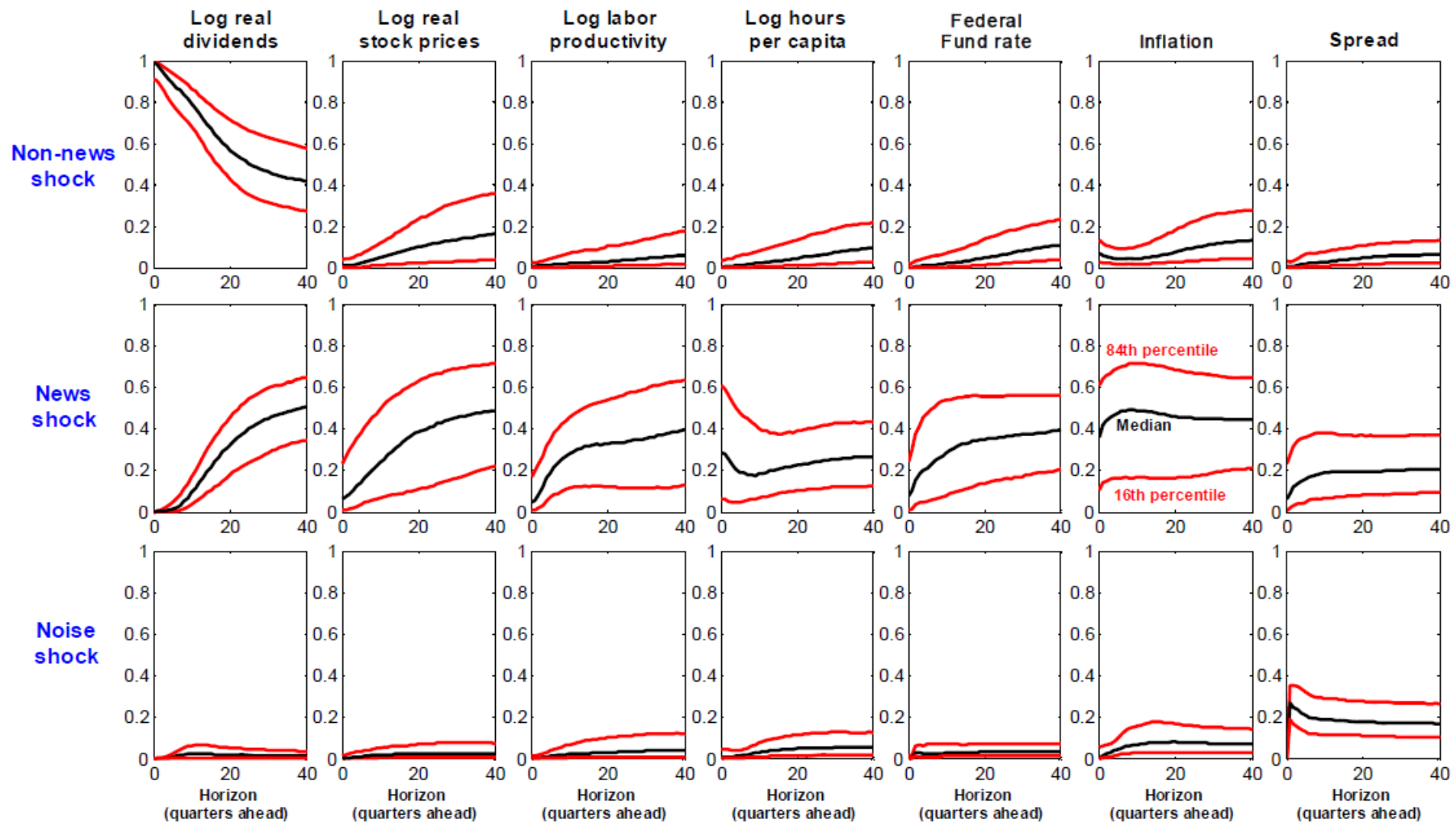


Figure 9 Application with dividends and stock prices: Fractions of forecast error variance explained by non-news, news, and noise shocks, based on a VARMA(4,1), and point-identification

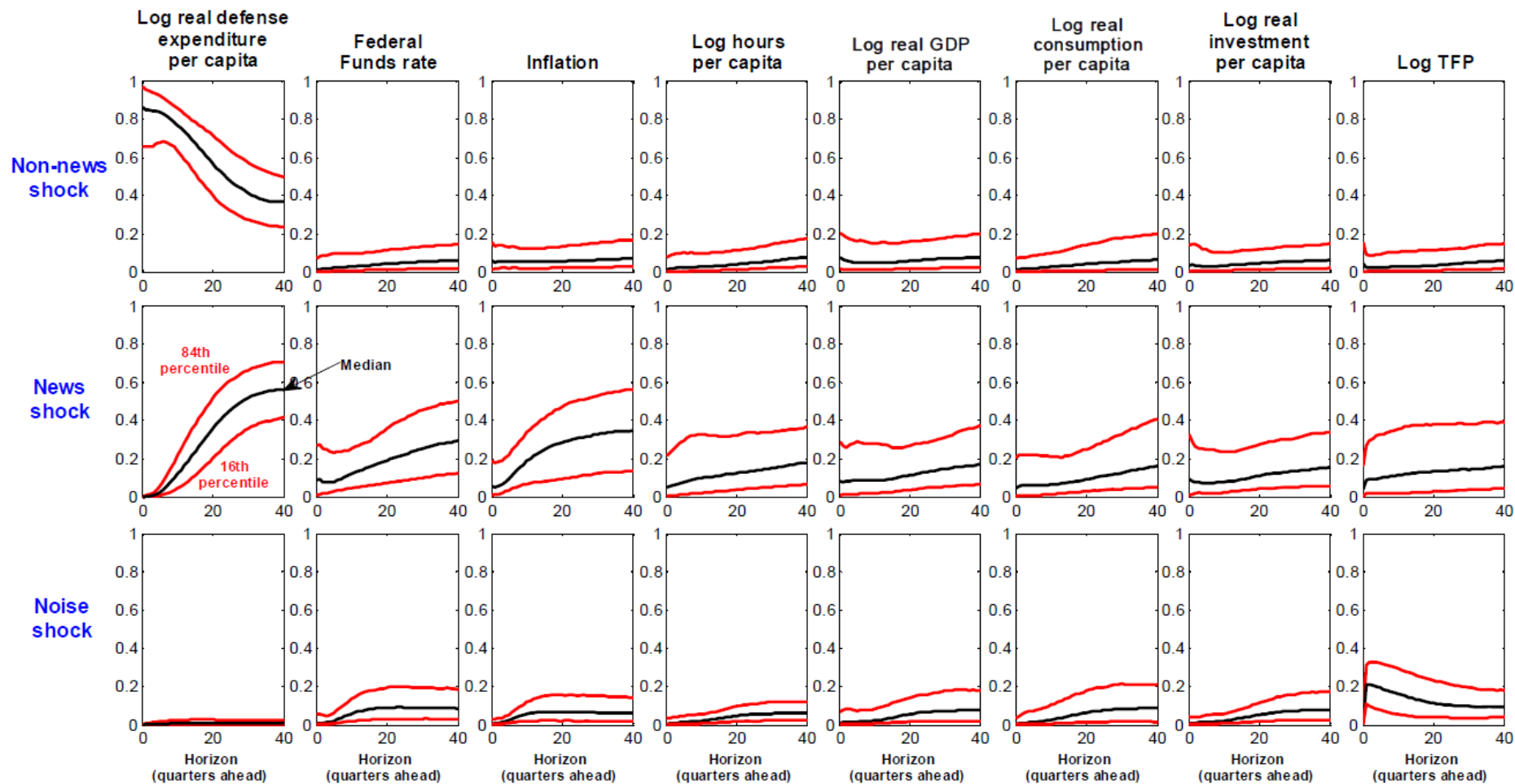


Figure 10 Application with defense expenditure: Fractions of forecast error variance explained by non-news, news, and noise shocks, based on a VARMA(4,1), and point-identification

about dividends are the dominant driver of stock prices at long horizons, and, likewise, they drive about 40 per cent of the FEV of labor productivity.

### 6.3 Defense expenditure

Finally, Figures 10 and XIII.2 in the online appendix show results from a VARMA(4,1) for the logarithms of real defense expenditure *per capita*, and seven other standard variables: log TFP, log hours *per capita*, the Federal Funds rate, GDP deflator inflation and the logarithms of real GDP, consumption, and investment *per capita*. The identification scheme assumes that the news *versus* noise problem pertains to defense expenditure, and all specification choices are the same as in the two previous applications. In line with the results from the previous two applications, the main finding emerging from Figure 10 is, once again, the minor-to negligible role played by noise shocks in driving the series included in the system at all horizons.

Clearly, all of our results suggest that noise shocks play no meaningful role in macroeconomic fluctuations.

## 7 Conclusions

In this paper, we have made four contributions to the ‘news versus noise’ literature: (I) We have provided a new identification scheme which, in population, exactly recovers news and noise shocks. (II) We have demonstrated that Chahrour and Jurado’s (2018) claim about the observational equivalence of news and noise shocks does not, in general, hold. (III) We have proposed a new econometric methodology for implementing our identification scheme, and we have shown, via Monte Carlo, that it has an excellent performance. (IV) We have provided several empirical applications of our identification scheme and econometric methodology. Overall, our results have uniformly suggested that, contrary to previous findings in the literature, noise shocks play a minor role in macroeconomic fluctuations.

The methodology introduced in this paper is potentially useful in a wide variety of contexts. In terms of the econometric methods, we have established that SVARMA models of the large dimension increasingly used in macroeconomics can be easily and robustly estimated. SVARMAs allow a researcher to do things which are impossible to do within the SVAR framework. Our scheme for identifying news and noise shocks should also have wide applicability. For instance, it could be used to explore the role played by news and noise shocks with alternative anticipation horizons (i.e., with  $K$  news shocks anticipating variation in the relevant variable 1, 2, 3, ...,  $K$  periods ahead) or to investigate the comparative role played by surprise, news and noise shocks in driving fluctuations in other variables, such as real exchange rates.

## References

- BARSKY, R., S. BASU, AND K. LEE (2014): “Whither News Shocks?,” in *Jonathan Parker and Michael Woodford, editors, NBER Macroeconomics Annuals*, 29, 225–264.
- BARSKY, R., AND E. SIMS (2011): “News Shocks and Business Cycles,” *Journal of Monetary Economics*, 58(3), 273–289.
- (2012): “Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence,” *American Economic Review*, 102(4), 1343–1377.
- BENATI, L. (2014): “Do TFP and the Relative Price of Investment Share a Common I(1) Component?,” *Journal of Economic Dynamics and Control*, 45(August 2014), 239–261.
- BENATI, L., E. EISENSTAT, AND G. KOOP (2018): “Can News and Noise Shocks Be Disentangled?,” Universities of Bern, Queensland, and Strathclyde, mimeo.
- BLANCHARD, O., J.-P. L’HUILIER, AND G. LORENZONI (2013): “News, Noise, and Fluctuations: An Empirical Exploration,” *American Economic Review*, 103(7), 3045–70.
- CANOVA, F., AND M. PAUSTIAN (2011): “Measurement with Some Theory: Using Sign Restrictions to Evaluate Business Cycle Models,” *Journal of Monetary Economics*, 58, 345–361.
- CELEUX, G., F. FORBES, C. ROBERT, AND D. TITTERINGTON (2006): “Deviance Information Criteria for Missing Data Models,” *Bayesian Analysis*, 1(4), 651–674.
- CHAHROUR, R., AND K. JURADO (2018): “News or Noise? The Missing Link,” *American Economic Review*, Forthcoming.
- CHAN, J., E. EISENSTAT, AND G. KOOP (2016): “Large Bayesian VARMA,” *Journal of Econometrics*, 192(2), 374–390.
- CHARI, V., P. KEHOE, AND E. MCGRATTAN (2008): “Are Structural VARs With Long-Run Restrictions Useful in Developing Business Cycle Theory?,” *Journal of Monetary Economics*, 55(8), 1337–1352.
- COOLEY, T., AND M. DWYER (1998): “Business Cycle Analysis Without Much Theory: A Look at Structural VARs,” *Journal of Econometrics*, 83, 57–88.
- FERNALD, J. (2012): “A Quarterly, Utilization-Adjusted Series on Total Factor Productivity,” Federal Reserve Bank of San Francisco Working Paper Series, 2012-19.



- FERNANDEZ-VILLAYERDE, J., J. RUBIO-RAMIREZ, T. SARGENT, AND M. WATSON (2007): “A’s, B’s, C’s (and D’s) for Understanding VARs,” *American Economic Review*, 97, 1021–1026.
- FORNI, M., L. GAMBETTI, M. LIPPI, AND L. SALA (2017): “Noisy News in Business Cycles,” *American Economic Journal: Macroeconomics*, forthcoming.
- GEORGE, E., D. SUN, AND S. NI (2008): “Bayesian Stochastic Search for VAR Model Restrictions,” *Journal of Econometrics*, 142, 553–580.
- GEWEKE, J. (1992): “Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments,” in *J. M. Bernardo, J. Berger, A. P. Dawid and A. F. M. Smith (eds.), Bayesian Statistics, Oxford University Press, Oxford*, pp. 169–193.
- HAMILTON, J. (1994): *Time Series Analysis*. Princeton, NJ, Princeton University Press.
- HANSEN, L. P., W. T. ROBERDS, AND T. J. SARGENT (1991): “Time Series Implications of Present Value Budget Balance and of Martingale Models of Consumption and Taxes,” in *Lars Peter Hansen and Thomas J. Sargent, editors, Rational Expectations Econometrics, Boulder: Westview Press*, pp. 121–161.
- HANSEN, L.-P., AND T. J. SARGENT (1981): “Exact Linear Rational Expectations Models: Specification and Estimation,” *Federal Reserve Bank of Minneapolis Staff Report No. 71*.
- ITO, T., AND D. QUAH (1989): “Hypothesis Testing with Restricted Spectral Density Matrices, with an Application to Uncovered Interest Parity,” *International Economic Review*, 30(1), 203–215.
- KOOP, G. (2013): “Forecasting with Medium and Large Bayesian VARs,” *Journal of Applied Econometrics*, 28, 177–203.
- KOROBILIS, D. (2013): “VAR Forecasting Using Bayesian Variable Selection,” *Journal of Applied Econometrics*, 28, 204–230.
- KURMANN, A., AND C. OTROK (2013): “News Shocks and the Slope of the Term Structure of Interest Rates,” *American Economic Review*, 103(6), 2612–32.
- LIPPI, M., AND L. REICHLIN (1994a): “VAR Analysis, Nonfundamental Representations, Blaschke Matrices,” *Journal of Econometrics*, 63, 307–325.
- (1994b): “VAR Analysis, Nonfundamental Representations, Blaschke Matrices,” *Journal of Econometrics*, 63(1), 307–325.

- POSKITT, D., AND W. YAO (2017): “Vector Autoregressions and Macroeconomic Modeling: An Error Taxonomy,” *Journal of Business Economic Statistics*, 35(3), 407–419.
- SPIEGELHALTER, D., N. BEST, B. CARLIN, AND A. VANDERLINDE (2002): “Bayesian Measures of Model Complexity and Fit,” *Journal of the Royal Statistical Society Series B*, 64(4), 583–639.
- UHLIG, H. (1998): “Comment On: The Robustness of Identified VAR Conclusions About Money,” *Carnegie-Rochester Conference Series on Public Policy*, 49, 245–263.
- (2003): “What Drives GNP?,” Unpublished manuscript, Euro Area Business Cycle Network.
- (2004): “Do Technology Shocks Lead to a Fall in Total Hours Worked?,” *Journal of the European Economic Association*, 2(2-3), 361–371.
- (2005): “What are the Effects of Monetary Policy on Output? Results from an Agnostic Identification Procedure,” *Journal of Monetary Economics*, 52(2), 381–419.