UNIVERSITÄT BERN

Faculty of Business, Economics and Social Sciences

Department of Economics

Can News and Noise Shocks Be Disentangled?

Luca Benati

18-05

February 2018

DISCUSSION PAPERS

Schanzeneckstrasse 1
Postfach
CH-3001 Bern, Switzerland
http://www.vwi.unibe.ch

Can News and Noise Shocks Be Disentangled?

Luca Benati
University of Bern*

Eric Eisenstat
University of Queensland†

Gary Koop
University of Strathclyde‡

Abstract

Chahrour and Jurado (2018) have shown that news and noise shocks are observationally equivalent when the econometrician only observes a fundamental process and agents' expectations about it. We show that the observational equivalence result no longer holds when the econometrician observes a fundamental process and a noisy signal of it. Working with an RBC model with noise about TFP, we further show that, even if the signal is not directly observed by the econometrician, it can be inferred through its impact on other macroeconomic variables, since they are optimally chosen by agents conditional on all information, including the signal itself. In particular, we show that under these circumstances news and noise shocks can be exactly recovered in population. Our results demonstrate that news and noise shocks are not observationally equivalent for an econometrician exploiting all the information contained in standard macroeconomic time series.

*Department of Economics, University of Bern, Schanzeneckstrasse 1, CH-3001, Bern, Switzerland. Email: luca.benati@vwi.unibe.ch

†School of Economics, University of Queensland, Brisbane St Lucia, QLD 4072, Australia. Email: e.eisenstat@uq.edu.au

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

1 Introduction

In a paper which is forthcoming in this *Review*, Chahrour and Jurado (2018, henceforth CJ) have questioned the logical possibility of separately disentangling news and noise shocks, arguing that they are, in fact, observationally equivalent, and, as such, they cannot be separately identified.

In this note we make three main points:

First, we show that CJ's result about observational equivalence crucially hinges on their casting their 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. An implication of our findings is that, since *agents* should be assumed to observe the signal, they are able to disentangle news and noise shocks.

Second, we 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 'what the signal does to the economy', through the agents' optimal choices in response to the signal. 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.

Third, to reinforce the second point, we show how news and noise shocks can be disentangled within a representative model: the RBC model of Barsky and Sims (2011). We take this model, which involves news shocks about total factor productivity (TFP) and other (non-news) shocks, and we augment it with a noise shock. 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) news and non-news shocks are the only disturbances having a permanent impact on TFP.

The note is organized as follows. In the next section we outline the logical steps which build up our argument. In Section 3 we illustrate our point about the breakdown of observational equivalence conditional on observing the signal based on one of CJ's mathematical examples. Section 4 contains our results using the RBC model of Barsky and Sims (2011) augmented with noise about TFP. Section 5 contains a general mathematical proof of the breakdown of observational equivalence conditional on observing the signal based on Hilbert spaces arguments.

2 A Summary of Our Argument

Before delving into the mathematics, and into DSGE models, it is useful to outline the steps in our argument that, for the applied econometrician, news and noise shocks are not observationally equivalent.

(I) We show, using CJ's simple mathematical examples, that observational equivalence between news and noise shocks breaks down when the analysis is cast from the perspective of the joint distribution for fundamentals and *signals* about them (in Section 5 we provide a general proof based on Hilbert space arguments). This implies that if the signal is observed, the two shocks can be exactly recovered. At first sight this would appear to be irrelevant for the problem at hand since it is only agents who observe the signal, not the econometricians. To show that it is not irrelevant, we develop our argument in the following points.

(II) Within any relevant macroeconomic model—e.g., Blanchard *et al.*'s (2013) model, or the RBC model with noise about TFP we use in Section 4—economic agents will react to the signal, and their actions will have an impact on the economy. This implies that macroeconomic variables will contain information about the signal (and therefore about news and noise shocks). Within the RBC model used in this paper, 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.

(III) The implication of (II) is that although econometricians do not observe the signal, they can learn about it by observing 'what the signal does to the economy'. One way of grasping this point is by analogy with Kalman filtering in a state space model. Although, from the econometrician's point of view, the states are typically unobserved, they do have an impact on the data. This impact is modelled through the measurement equation(s). Knowledge of the form of the state space model then allows the econometrician to estimate the states using standard methods such as Kalman filtering. Here the logic is the same since the most relevant macroeconomic models can be written as state space models, and the signal can be interpreted as a state which is unobserved to the econometrician.

(IV) 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).

(V) The preceding points suggest that econometricians should be able to use the data to learn about the signal and, thus, disentangle news and noise shocks. But it is useful to see how this works in practice in a model of the sort used in this literature. In the RBC model of Barsky and Sims (2011) augmented with noise, we show how this can be done. In particular, we show that it is possible to exactly recover, in population, news and noise shocks' IRFs and fractions of FEV.

(VI) For the purpose of identifying news and noise shocks, the crucial insight is the following. Although, as discussed (e.g.) by Forni, Gambetti, Lippi, and Sala (2017), news and noise shocks are observationally equivalent on impact—in the sense of generating identical impulse vectors at t=0 for all variables[1]—they are *not* observationally equivalent at all subsequent horizons. For example, within the RBC model with noise of Section 4, news shocks ultimately have a permanent impact on TFP, whereas noise shocks do not have a permanent impact on any variable. This implies that the two shocks cannot be observationally equivalent once all of their implications for the observed data are taken into account. Put simply, a shock causing a permanent increase in one or more variables cannot be observationally equivalent to a shock which has no permanent impact on anything.

(VII) We also show how, in this RBC model, news and noise shocks' IRFs and fractions of FEV can be recovered starting from the theoretical reduced form moving average (MA) representation of the model, by imposing the model implied theoretical restrictions that: (*i*) news and noise shocks generate identical impulse vectors at t=0 for all variables; and (*ii*) news shocks have a permanent impact on TFP but noise shocks do not have a permanent impact on it (or any other variable).

(VIII) In order to move from theory to empirical estimation using structural Vector Autoregressive (VAR) methods, point (VII) might appear irrelevant, since, as pointed out (e.g.) by Blanchard *et al.* (2013), VARs cannot be used in order to recover the MA representation of models with noise shocks.[2] We show, however, that low order VARMAs can do so very effectively. For the RBC model with noise, a VARMA(1,1) produces theoretical IRFs and fractions of FEV which are indistinguishable from those obtained from the theoretical MA representation. We show that imposing the model-implied identifying restrictions upon the theoretical truncated VARMA(1,1) representation of the RBC model allows for the recovery with great precision of the model's theoretical IRFs and fractions of FEV.

(IX) We also address the observational equivalence of fundamental and non-fundamental MA representations.[3] We show that within the context of the RBC model, which yields a non-fundamental representation by construction (with one root at zero), the fact that many non-fundamental representations are admissible is not an

[1]See Forni *et al.* (2017, Section 2.2): *'[...] 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.

[2]From a mathematical point of view, this is due to the fact that news and noise shocks generate identical impulse vectors on impact, which makes the matrix of the shocks' structural impacts at t=0 singular.

[3]It is well known (see, e.g. Lippi and Reichlin (1994)) that alternative representations obtained by "inverting" characteristic MA roots through Blaschke filters preserve the stochastic properties of a multivariate Gaussian times series.

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.

(X) 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.

We now turn to illustrating point (I) within the simplest setup used by CJ.

3 Disentangling News and Noise Shocks Conditional on Observing the Signal

The example[4] involves an observed fundamental process, x_t, which is driven by a news shock (μ_t) and a non-news shock (η_t). A news shock is defined as a time-t disturbance which has an impact on x_t only at a future date, which we assume to be one period ahead, and therefore enters lagged into the equation for the fundamental:

$$x_t = \mu_{t-1} + \eta_t. \tag{1}$$

The model also involves an observed noisy signal (s_t) about the news shock:

$$s_t = \mu_t + \xi_t \tag{2}$$

where ξ_t is the noise shock. Thus, the model has three shocks (η_t, μ_t and ξ_t) and two observed variables (x_t and s_t). The three shocks are assumed to be independent of each other and independent over time 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 will use notation where $t|s$ subscripts denote expectations of a variable at time t conditional on information available through time s.

At the most general level, we are interested in the question: Given the data, what can we learn about the unobservables in our model? Observational equivalence arises if the data does not allow for the identification of some of these unobservables (in our case, to distinguish between news and noise shocks). Macroeconomists have come up

[4] The example is from Section 3.2, pp. 15-16.

with standard techniques (such as those used by CJ) to address this question, but the answer can crucially hinge on what the data is.

CJ derive all of their results about the observational equivalence of news and noise shocks from the perspective of the joint distribution of fundamentals and agents' expectations about them (i.e. the data is x_t and $x_{t|t-1}$). They make this clear when they state:

> "To be concrete, imagine an econometrician who is able to observe the entire past, present, and future history of the fundamental process $\{x_t\}$, along with the entire past, present, and future history of people's subjective beliefs regarding $\{x_t\}$. More concisely, we will say that the econometrician observes 'fundamentals and beliefs.' All of our equivalence results are stated from the perspective of such an econometrician, and are to be understood with respect to those observables."

We highlight the fact that CJ are deriving their results from the perspective of the *econometrician*, not of an agent in a theoretical model.

In this paper we argue that—because of points (II) and (III) in the previous section—even from the perspective of the econometrician the more sensible assumption is to treat s_t 'as if' it were observed. To reiterate our main point there, although econometricians, strictly speaking, do not observe the signal, in fact it is 'as if' they did, because they see 'what the signal does' to variables such as GDP, consumption, investment, etc.. In this section we show that adopting the view that the signal is observed overturns CJ's observational equivalence result. In online Appendix A we show that the same holds true in all of CJ's theoretical examples.

First, we show that CJ's result is correct if we assume, as they do, that $[x_t,\, x_{t|t-1}]'$ is the data, and therefore work from the perspective of the joint distribution of these variables.

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

From (1) we have $x_{t-1|t-1} = \mu_{t-1|t-1}$. By applying the Kalman filter to (2) we have that

$$\mu_{t|t} = \underbrace{\frac{\sigma_\mu^2}{\sigma_\mu^2 + \sigma_\xi^2}}_{\kappa}[\mu_t + \xi_t] = \kappa s_t \tag{3}$$

so that

$$x_{t|t-1} = \kappa[\mu_{t-1} + \xi_{t-1}] \tag{4}$$

The assumption of zero-mean shocks means that the only non-zero moments in the joint distribution of $[x_t,\, x_{t|t-1}]'$ are $\mathrm{Var}[x_t]$, $\mathrm{Var}[x_{t|t-1}]$, and $\mathrm{Cov}[x_t,\, x_{t|t-1}]$. But the last two of these are equal to one another, leading to observational equivalence. To be precise,

$$\mathrm{Var}[x_t] = \sigma_\mu^2 + \sigma_\eta^2 \tag{5}$$

$$\text{Cov}[x_t,\, x_{t|t-1}] = \text{Var}[x_{t|t-1}] = \frac{\sigma_\mu^4}{\sigma_\mu^2 + \sigma_\xi^2} \tag{6}$$

With three parameters (σ_μ^2, σ_η^2, and σ_ξ^2) and two equations, the system is not identified and thus news and noise shocks cannot be disentangled. This is CJ's result.

3.2 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. To see this, note that with x_t and s_t given by (1) and (2), the system $[x_t,\, s_t]'$ features three distinct non-zero second moments, instead of two: The first one is still given by (5), whereas the other two are given by

$$\text{Var}[s_t] = \sigma_\mu^2 + \sigma_\xi^2 \tag{7}$$

$$\text{Cov}[x_t,\, s_{t-1}] = \sigma_\mu^2 \tag{8}$$

Equation (8) directly provides σ_μ^2. Given σ_μ^2, equations (5) and (7) can then be used to compute σ_η^2 and σ_ξ^2. Using σ_μ^2, σ_η^2 and σ_ξ^2, we can compute κ. All of the parameters of the model can thus be obtained, which allows us to exactly recover the theoretical IRFs to the three shocks, and the fractions of FEV of the variables they explain at the various horizons.

3.3 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 A, the same holds true for CJ's more complex models.

This is an illustration of the more general (and perhaps obvious) point that the identification of shocks (or parameters) within an empirical model can crucially depend on what data the researcher uses. Theoretical models within the 'news and noise' literature typically feature several observed macroeconomic variables. The DSGE model estimated by Blanchard *et al.* (2013), for example, features, as observables, GDP, consumption, investment, employment, the short-term monetary policy rate, inflation, and wages, whereas the RBC model with noise we will use in the next section features GDP, consumption, investment, hours, and TFP. As discussed in points (II) and (III) in Section 2, this implies that, since most (or all) of these variables react to the signal, they will contain information about it, thus erasing—from a practical standpoint—the distinction between agents (who observe the signal) and econometricians (who do not). In turn this 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 show in the next section based on an RBC model with noise about TFP, this is indeed the case.

4 Recovering News and Noise Shocks Conditional on Observed Macroeconomic Data

4.1 An RBC model with noise shocks about TFP

The RBC model we use in this section is the one proposed by Barsky and Sims (2011), which we augment with noise shocks about TFP. The basic elements of the model 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, ..., ∞. C_t is consumption, N_t is hours worked, E_t^N is a random process whose logarithm, ϵ_t^n, is $N(0,\, \sigma_n^2)$, $0 < b < 1$ is a parameter capturing the strength of habit formation and η 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 θ 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 δ is the depreciation rate, γ 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 ϵ_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

$$a_t = \tilde{a}_t + v_t \tag{9}$$

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

$$\tilde{a}_t = \tilde{a}_{t-1} + \epsilon_t^{NN} + \epsilon_{t-\tau}^{NE} \tag{10}$$

where ϵ_t^{NN} and ϵ_t^{NE} are the non-news and news shocks, respectively, with $\epsilon_t^{NN} \sim N(0, \sigma_{NN}^2)$ and $\epsilon_t^{NE} \sim N(0, \sigma_{NE}^2)$. τ is the anticipation horizon for the news shock, which we set to one period. A direct implication of (10)—which will play a crucial

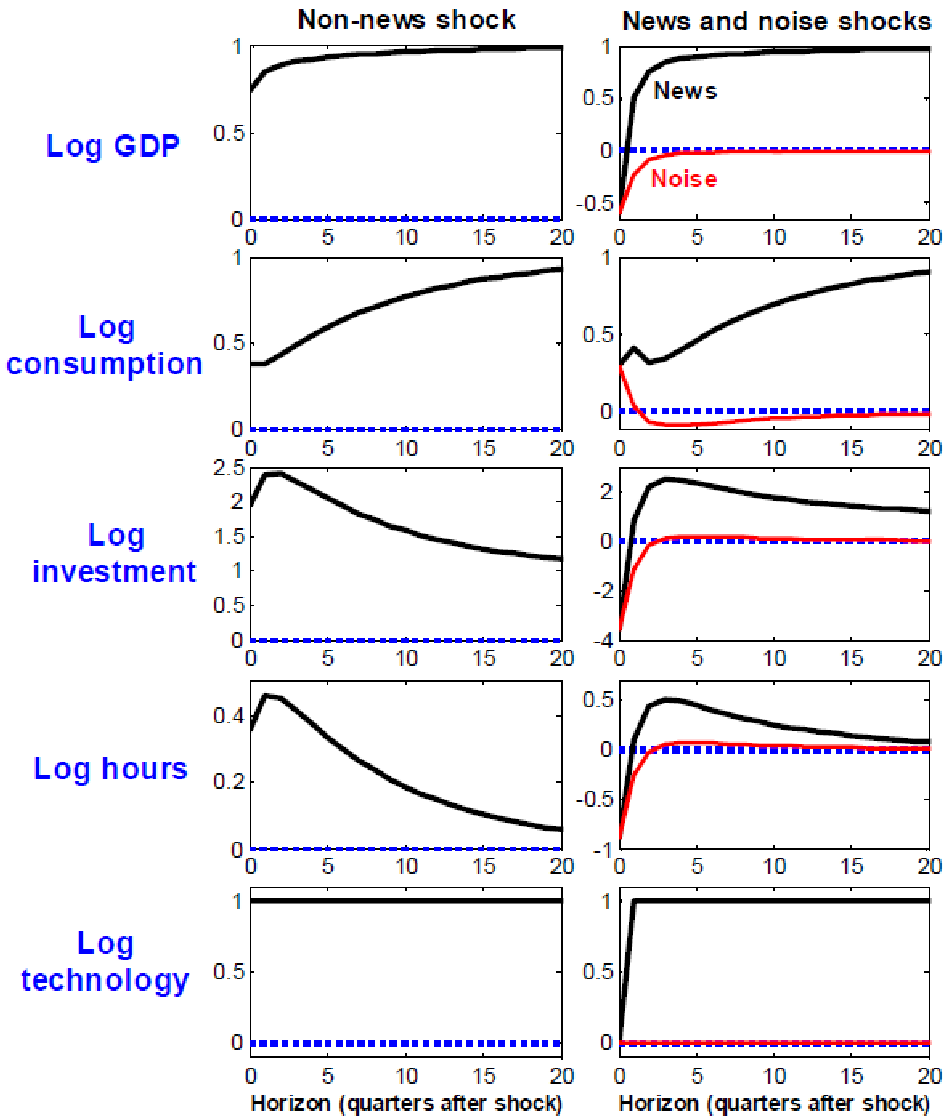

Figure 1 Impulse-response functions to non-news, news and noise shocks for Barsky and Sims' (2011) model augmented with noise shocks about TFP

role in disentangling news and noise shocks—is that non-news and news shocks are the only disturbances having a permanent impact on TFP. Once again, this logically implies that there is no way that news shocks can be confused with noise shocks, for the simple reason that a permanent shock cannot be confused with a disturbance which has no permanent impact on anything.

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, s_t, which is equal to the sum of the news shock and of a noise component

$$s_t = \epsilon_t^{NE} + u_t \tag{11}$$

with $u_t \sim N(0, \sigma_u^2)$. Details of the agent's signal extraction problem, together with the model's solution, are given in online Appendix B.

Figure 1 shows the IRFs of the logarithms of GDP, consumption, investment, hours and technology to one-standard deviation non-news, news and noise shocks, whereas Figure A.1 in the online Appendix shows the fractions of FEV explained by the three shocks. The model is calibrated using parameter values similar to those in Barsky and Sims (for details, see online Appendix B). 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.

4.1.1 Theoretical implications of the 'news *versus* noise' problem

Figure 1 clearly illustrates two key theoretical implications of the 'news *versus* noise' problem:

(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.

Implication (1) follows directly from 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, at t=0, they will respond to the two shocks in exactly the same way.

Implication (2) originates from the fact that, as time goes by, the true nature of the shock is progressively revealed. In particular, within this model, the news shock ultimately causes a permanent increase in TFP, whereas the noise shock does not cause any permanent change in any variable.

4.1.2 Identifying restrictions

Following the reasoning of the preceding sub-section, we impose the following restrictions in order to identify non-news, news, and noise shocks in population (i.e., based on the theoretical MA representation of the RBC model):

(1) At t=0 news and noise shocks have no impact on TFP, whereas they generate identical impulse vectors for all other variables.

(2) At t=0, TFP is impacted upon by only two disturbances, the non-news shock and the transitory shock—see equations (9) and (10)—which generate identical impulse vectors on impact.[5] The former shock is disentangled from the latter because it explains the maximum fraction of the FEV of TFP at a specific long horizon (which we will take to be 40 quarters ahead).

(3) Among all of the shocks which do not impact upon TFP at t=0, the news shock is the one which explains the maximum fraction of the FEV of TFP at a long horizon (again, 40 quarters ahead).

We now show that 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 observables be

$$Y_t = A_0\epsilon_t + A_1\epsilon_{t-1} + A_2\epsilon_{t-2} + A_3\epsilon_{t-3} + ... \tag{12}$$

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; 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, ..., N) of the MA matrices A_0, A_1, A_2, A_3, ... has been divided by the standard deviation of the i-th shock in ϵ_t. Representation (12) can be immediately recovered from the model's IRFs to the structural innovations.

Observationally equivalent reduced-form representations of (12) can be obtained by post-multiplying all of the MA matrices A_0, A_1, A_2, A_3, ... by an orthogonal rotation matrix R, yielding

$$Y_t = \tilde{A}_0\tilde{\epsilon}_t + \tilde{A}_1\tilde{\epsilon}_{t-1} + \tilde{A}_2\tilde{\epsilon}_{t-2} + \tilde{A}_3\tilde{\epsilon}_{t-3} + ... \tag{13}$$

where $\tilde{A}_j = A_jR$ and $\tilde{\epsilon}_{t-j} = R'\epsilon_{t-j}$, j = 0, 1, 2, 3, ... We can randomly generate different rotation matrices[6], thus producing different observationally equivalent VMAs. The question we wish to address is whether imposing restrictions (1)-(3) from the preceding sub-section on different reduced form VMAs allow us to recover the true structural VMA.

[5] The logic here is exactly the same as for news and noise shocks. From (9), although agents observe a_t, they never observe its two individual components, $\tilde{a}_t$ and v_t. As a result, at each time t, they face a signal-extraction problem pertaining to non-news and transitory shocks, which is analogous to that pertaining to news and noise shocks. As a result, on impact agents will react in exactly the same way to non-news and transitory shocks such as v_t.

[6] 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 $Q\prime$.

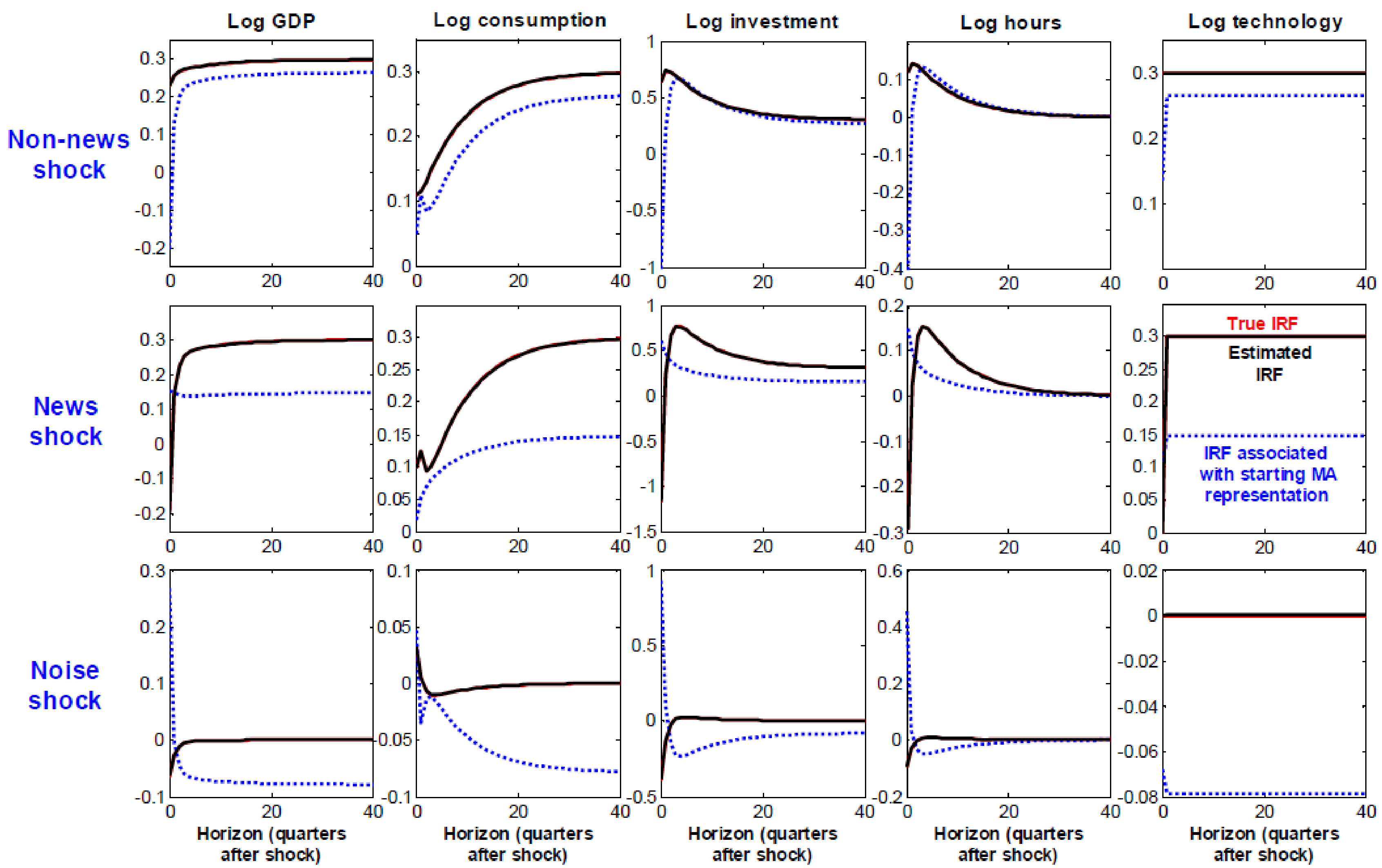

Figure 2 Recovering news and noise shocks in population: Theoretical IRFs, and estimated IRFs based on the theoretical MA representation of the RBC model augmented with noise shocks about TFP

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. We illustrate this in Figure 2 which contains results using one random rotation.[7] The figure shows, in red, the true IRFs as captured by the theoretical MA representation (12); in blue (dotted line) the IRFs associated with the initial, randomly rotated MA representation (13); and in black the estimated IRFs, i.e., the IRFs we recovered starting from (13), 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 discussion of point (VI) of Section 2, 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 VARMAs(p,q)

An issue arises when the econometrician tries to actually go to the data and obtain an estimate of the reduced form VMA representation of the preceding sub-section to disentangle news and noise shocks. A common practice in macroeconomics is to first estimate a reduced form VAR, then invert it to obtain a reduced form VMA representation. In models with noisy news, this is not possible as shown, e.g., in Blanchard *et al.* (2013). However, as we shall show in this sub-section, low-order VARMAs can approximate with great precision the MA representation of the observables. For example, for the RBC model with noise, a truncated[8] VARMA(1,1) already produces theoretical IRFs and fractions of FEV which are indistinguishable from those associated with the model's theoretical MA representation. In fact, as shown in Benati, Chan, Eisenstat, and Koop (2018) *via* extensive Monte Carlo simulations, from an empirical standpoint, an effective way of recovering the VMA representation of the observables is through the use of Bayesian VARMA(p, q) models.

To demonstrate these points, we repeat the exercise of the preceding sub-section where we randomly generate different rotation matrices so as to produce different reduced form specifications. In this sub-section, we use a VARMA(1,1) instead of the VMA used previously. To be precise, we use the truncated theoretical VARMA(1,1) representation of the model,

[7] Figure 2 presents results for IRFs. Results for FEVs are available in the online appendix. Results from all of the 100 runs are available upon request. In addition, throughout the entire paper 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.

[8] We compute truncated VARMA(p,q) approximations to the model's theoretical MA representation *via* linear projection arguments—see Appendix C.

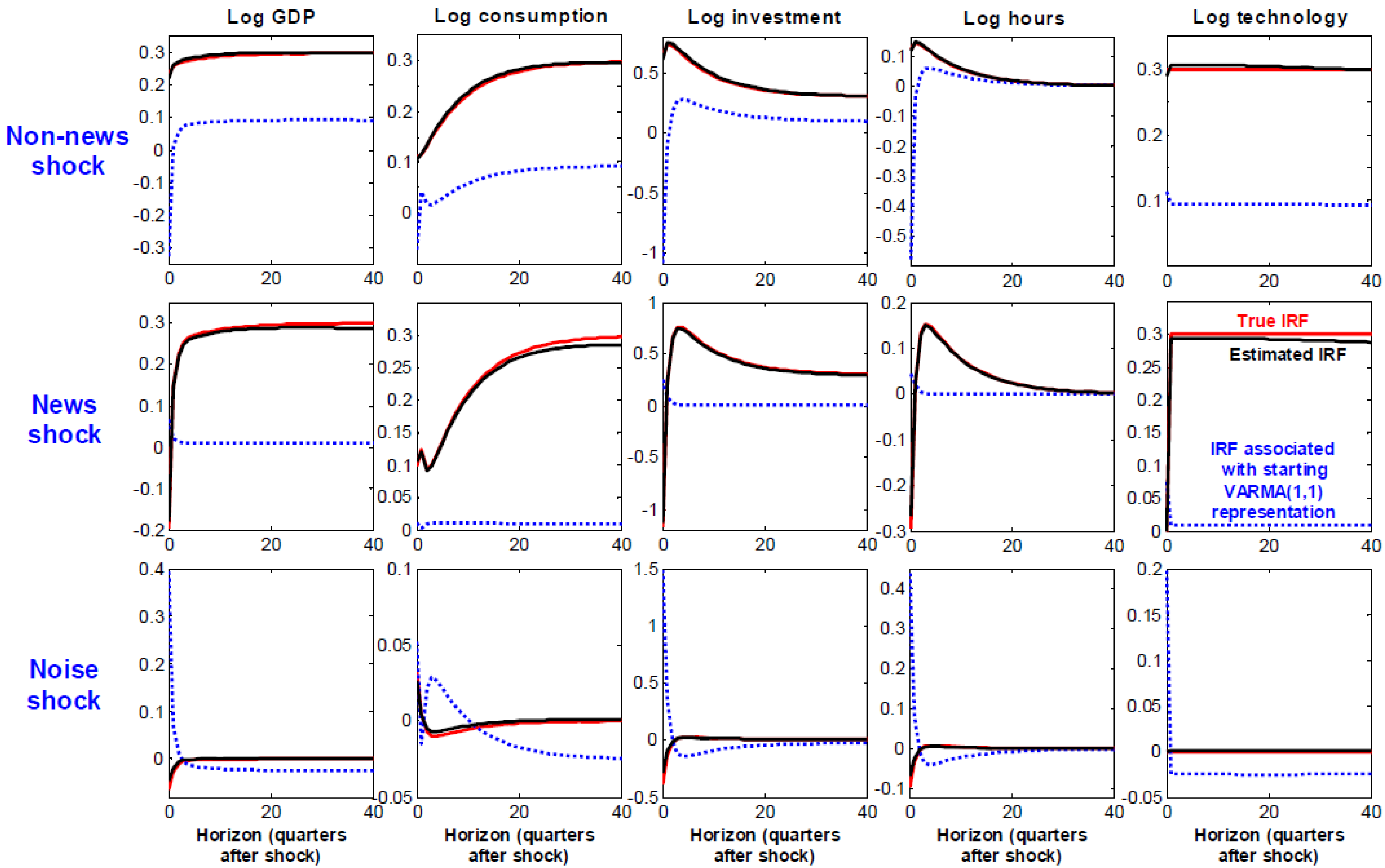

Figure 3 Recovering news and noise shocks in population: Theoretical IRFs, and estimated IRFs based on the theoretical truncated VARMA(1,1) representation of the RBC model augmented with noise shocks about TFP

$$Y_t = BY_{t-1} + A_0\epsilon_t + A_1\epsilon_{t-1} \tag{14}$$

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

$$Y_t = BY_{t-1} + \tilde{A}_0\tilde{\epsilon}_t + \tilde{A}_1\tilde{\epsilon}_{t-1}. \tag{15}$$

We then impose the identifying restrictions of sub-section 4.1.2 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 Figure 3 for the IRFs (the comparable figure for the FEVs is in the online appendix) for one of the random transformations. The red line in the figure shows the true IRFs from RBC model (i.e. the same lines shown in Figure 2). The blue, dotted lines show the IRFs from the reduced form VARMA(1,1) produced by the random transformation. The black line shows the IRFs that result when we impose the identifying restrictions (1)-(3). Again we are finding that imposing the identifying restrictions allows for the recovery of the model's true IRFs.

This suggests that an effective empirical strategy for estimating noisy-news model is to work with low order VARMAs and then impose the restrictions which identify the news and noise shocks. This strategy is implemented in Benati, Chan, Eisenstat, and Koop (2018).

4.2.2 Alternative non-fundamental representations

A possible objection to the conclusion of the preceding sub-section is that considering the set of all admissible non-fundamental representations of the (approximated) VARMA representation of the model may produce different outcomes.[9] In this sub-section we show that, at least for the RBC model augmented with noise, this is not an issue. In particular, we demonstrate that imposing the identifying restrictions of sub-section 4.1.2 allows news and noise shocks to be recovered in all of the non-fundamental representations of the truncated theoretical VARMA(1,1).

Figure 4 reports results from performing the same exercise of the preceding sub-section, but this time based on all of the non-fundamental representations of the truncated VARMA(1,1) we used there. In particular, the theoretically derived VARMA(1,1) representations reported in sub-section 4.2.1 have one characteristic MA root at zero and all nonzero roots outside the unit circle. Starting from this, we compute the set of all non-fundamental representations of the fundamental by inverting combinations of the nonzero roots; the root at zero, which is a direct implication of the theory, is never flipped.

[9] In general, set of admissible fundamental and non-fundamental representations is always countable, with at most 2^{nq} possible representations, although the actual size will vary depending on the number of real roots vs. complex conjugate pairs.

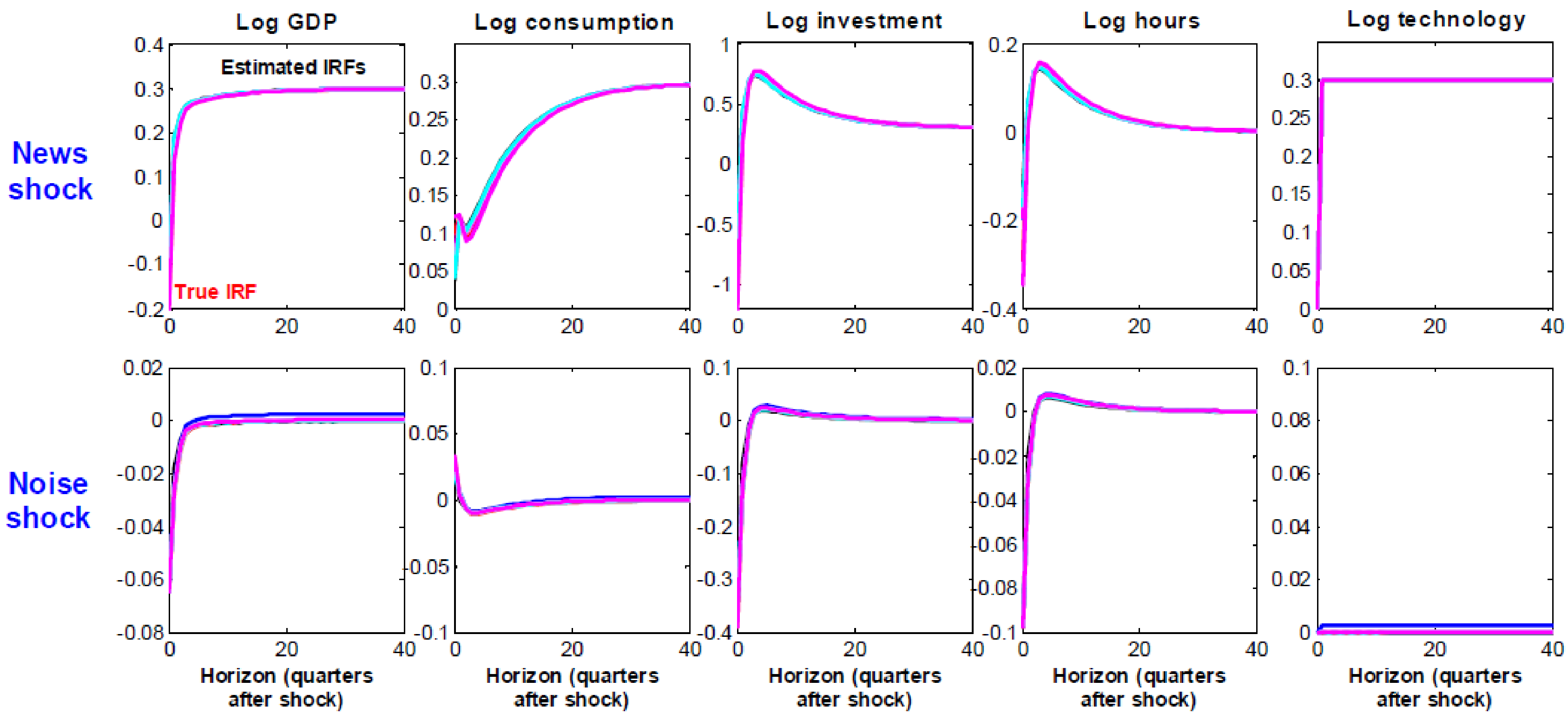

Figure 4 Addressing non-fundamentalness: Recovering news and noise shocks IRFs in population based on the non-fundamental representations of the theoretical truncated VARMA(1,1) representation of the RBC model with noise shocks about TFP

For the sake of brevity, we only report results for the news and noise shocks. As the figure shows, imposing the model-implied identifying restrictions on any of the non-fundamental representations leads to the same results, with the true IRFs to news and noise shocks being exactly recovered (up to numerical approximation errors).[10] So, although in other macroeconomic modeling contexts, it is possible that different non-fundamental representations of a VMA may produce substantially different results, within the present context this is clearly not the case.

A more general point about non-fundamentalness There is however a more general issue to be stressed here. For some models (but not the ones used in this paper), there are some cases where several admissible non-fundamental representations may produce results which differ in some ways. This issue has nothing to do with the possibility of disentangling news and noise shocks. The latter possibility uniquely hinges on the two shocks not being observationally equivalent conditional on a given reduced form VMA (or approximate VARMA) representation, which allows for the construction of appropriate static orthogonal rotations that transform reduced form innovations to structural shocks (including news and noise).

As discussed in Section 4.1.3, conditional on the particular MA representation implied by the RBC model, our model-based identifying restrictions of Section 4.1.2 allow for the recovery of the model's true IRFs and fractions of FEV irrespective of how we randomly (statically) rotate the MA representation. That is, all of the observationally equivalent VMAs constructed there produced the same IRFs and FEVs for news and noise shocks. In Section 4.2.1 we showed that the same result holds for the truncated theoretical VARMA(1,1) representation associated with the MA representation of the RBC model. As a matter of logic, this implies that this will also hold for any individual non-fundamental VARMA representation obtained by inverting nonzero/finite characteristic roots.

The key insight is that while inverting characteristic roots does produce observationally equivalent MA representations, identifying the correct (nonzero) root structure is completely separate from identifying news and noise shocks. Indeed, the two problems can and should be tackled individually. Characteristic MA roots are invariant to static orthogonal rotations, so the problem of identifying the appropriate set of non-fundamental MA representations is the same in the reduced-form case as it is in the structural case. The problem of separating news and noise shocks, in this context, is simply a matter of constructing static orthogonal rotations that recover structural shocks in any MA representation with one characteristic root at zero.

[10] For the comparison with the IRFs produced by the original representation to be meaningful, those produced by the alternative non-fundamental representations ought to be appropriately normalized. The IRFs reported in Figure 4 have been normalized as follows. For the news shock, they have been normalized in such a way that they produce a long-horizon impact on log TFP identical to that produced by the fundamental representation (i.e., 0.3). The IRFs to noise shocks, on the other hand, have been normalized in such a way that they produce an impact on GDP at $t=0$ identical to that produced by the fundamental representation.

All results in this section unequivocally demonstrate that using model-based restriction, static orthogonal rotations can be readily obtained, so that, once again, news and noise shocks are *not* observationally equivalent: the model-implied restriction will always recover the same IRFs and FEVs regardless of how we randomly (statically) rotate the reduced form VARMA representation.

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.

We have made these points in the context of a particular RBC model. But they will hold with much greater generality as we now show. In particular, the next section provides general mathematical proof that observation of the signal allows to exactly recover news and noise shocks.

5 General Proof of Uniqueness When Signals Are Observed

5.1 CJ's Representation Theorem

Before stating the main result regarding the general uniqueness of fundamentals and signals, we first clarify an important issue related to the interpretation of pure news and mixed noise/news representations. Some confusion regarding how these representations are characterized and treated in CJ is related to a small but nontrivial error in the proof of their representation theorem.

To this end, recall that the 'news' representation in CJ (2018) is given by

$$x_t = \sum_{i\in\mathbb{I}} a_{i,t-i}, \qquad a_{i,t-i} \perp a_{j,t-j} \text{ for all } i \neq j. \tag{16}$$

Their representation theorem states that this news representation is observationally equivalent to a 'noise' representation,

$$s_{i,t} = m_{i,t} + v_{i,t}, \qquad m_{i,t} \in \mathcal{H}(x), \qquad v_{i,t} \perp \mathcal{H}(x). \tag{17}$$

The part of the proof aimed at demonstrating this equivalence is contained in the last paragraph of page 44. Specifically, it proceeds by assuming

$$s_{i,t} \equiv a_{i,t} \text{ for all } i \in \mathbb{I},$$

and then making the claim 'because $\mathcal{H}(x) \subset \mathcal{H}(a)$ there exist...' What follows is a decomposition by orthogonal projection of $\mathcal{H}(s) \equiv \mathcal{H}(a)$, which gives rise to a non-trivial 'noise process' $v_{i,t}$.

The key problem here is with the assumption that $\mathcal{H}(x) \subset \mathcal{H}(a)$. Specifically, from the news representation, and by the property $a_{i,t-i} \perp a_{j,t-j}$ for all $i \neq j$, we obtain the orthogonal decomposition

$$\mathcal{H}_t(x) = \oplus_{i \in \mathbb{I}} \mathcal{H}_{t-i}(a_i),$$

where each $\mathcal{H}_{t-i}(a_i)$ is a closed subspace of $\mathcal{H}_t(x)$, or equivalently, $\mathcal{H}_t(a_i) \subset \mathcal{H}_{t+i}(x)$.

Consequently, for every $a_{i,t} \in \mathcal{H}(a)$, we have $a_{i,t} \in \mathcal{H}_t(a_i)$ and therefore $a_{i,t} \in \mathcal{H}_{t+i}(x)$. Carrying this further, since $\mathcal{H}_{t+i}(x) \subset \mathcal{H}(x)$, it holds that $a_{i,t} \in \mathcal{H}(x)$. We conclude from this that every $a_{i,t} \in \mathcal{H}(a)$ is also an element of $\mathcal{H}(x)$. But this (together with the obvious fact that every $x_t \in \mathcal{H}(a)$) implies that $\mathcal{H}(x) = \mathcal{H}(a)$!

From the equivalence of the subspaces $\mathcal{H}(x) = \mathcal{H}(a)$, it follows that $\mathcal{H}(s) \ominus \mathcal{H}(x) = \{0\}$ and $v_{i,t} = 0$ for all i and t. This means that the noise process does not exist, and there is no observational equivalence result in the sense of the representation theorem.

While this does not, per se, invalidate the main result regarding observational equivalence, it is important because it casts into doubt the particular representation (17) CJ refer to as 'noise'. Specifically, CJ appear to rely on the constructive nature of their proof to argue that a noise representation can be derived directly from a pure news representation, and they use this to condemn all other 'mixed news-and-noise' representations in the literature as misleading because they differ from the 'noise' representation they purport to extract.

The above reasoning, however, suggests that their argument for achieving this is incorrect. Indeed, the following section demonstrates that while infinitely many news/noise representations can yield equivalent (almost surely) fundamentals and expectations—so that observing fundamental and expectations is not sufficient to identify any particular such representation—each representation arises from a certain partition of the information space available to agents. Two key conclusions therefore emerge:

(1) A news-and-noise representation exists and is unique given a partition of the agents' information space, whereas a pure news representations arises by assuming agents perfectly observe all news shocks.

(2) Assuming *signals* are observed instead of, or in addition to, agents' expectations yields an observable process which uniquely identifies news and noise shocks. Once again, it is important to keep in mind the previous discussion illustrating why it is unreasonable to disregard the signal and instead focus exclusively on agents' expectations about the fundamental.

5.2 Alternative characterization of news and noise representations

As in CJ, we assume throughout that x_t is a stationary Gaussian stochastic process. Let $\mathcal{H}_t(\hat{x})$ be the space spanned by current and past expectations $\hat{x}_t, \hat{x}_{t-1}$, etc. and let $\mathcal{H}_t(x)$ be the space spanned by current and past fundamentals x_t, x_{t-1}, etc. Define further $\mathcal{H}^t(x) = \mathcal{H}(x) \ominus \mathcal{H}_{t-1}(x)$ to be the space spanned by current and future fundamentals $x_t, x_{t+1}, \ldots$.

Models with news and noise shocks in the literature may be characterized by the following basic assumptions:

Assumption 1 *$\mathcal{H}_t(x) \subseteq \mathcal{H}_t(\hat{x})$ for all $t \in \mathbb{Z}$.*

Assumption 2 *There exists $\mathcal{H}_t(v) \subset \mathcal{H}_t(\hat{x})$ such that $\mathcal{H}_t(v) \perp \mathcal{H}^t(x)$.*

Assumption 1 may be regarded as the typical building block of any model with anticipated shocks as it implies that there exists $\mathcal{H}_t(w) \subseteq \mathcal{H}_t(\hat{x})$ such that $\mathcal{H}_t(w) \subset \mathcal{H}^t(x)$. Assumption 2 defines a space spanned by noise processes and implies that the orthogonal compliment $\mathcal{H}_t(v)^\perp$ of $\mathcal{H}_t(v)$ in $\mathcal{H}_t(\hat{x})$ is the subspace $\mathcal{H}_t(w)$, such that $\mathcal{H}_t(\hat{x}) = \mathcal{H}_t(w) \oplus \mathcal{H}_t(v)$.

This decomposition yields a representation of the fundamentals process as a linear combination of current and past shocks. To complete the specification, assume that in each period t agents receive a signal $s_{i,t} \in \mathcal{H}_t(\hat{x})$ that provides an unbiased estimator of $a_{i,t}$ in the sense that $E(s_{i,t} \,|\, \mathcal{H}_t(w)) = a_{i,t}$. This defines a projection of $s_{i,t}$ on to $\mathcal{H}_t(w)$ such that $s_{i,t} = a_{i,t} + v_{i,t}$, with $v_{i,t} \in \mathcal{H}_t(v)$. Consequently, the representation of fundamentals and beliefs is given by

$$x_t = \sum_{i \in \mathbb{I}} a_{i,t-i}, \qquad s_{i,t} = a_{i,t} + v_{i,t}, \qquad i \in \mathbb{I}, \tag{18}$$

where $\mathbb{I}$ is a set of indices, and all $a_{i,t} \in \mathcal{H}_t(w)$, $v_{i,t} \in \mathcal{H}_t(v)$ are stationary processes such that $a_{i,t} \perp a_{j,t} \perp v_{k,t}$ for all $i \neq j$ and all $k \in \mathbb{I}$ (however, it is possible that $v_{i,t}$ and $v_{j,t}$ are correlated for some $i \neq j$).

When $\mathcal{H}_t(w) = \mathcal{H}_t(\hat{x})$, we have $\mathcal{H}_t(v) = \{0\}$ and therefore (18) is a pure news representation. In this case, expectations at time t are formed entirely from $s_{i,\tau} \equiv a_{i,\tau}$, for $i \in \mathbb{I}$ and $\tau \leq t$—i.e., from perfectly observed future changes to the fundamentals process. Otherwise, expectations are formed from noisy signals.

We can restate the main result in CJ using two theorems as follows.

Theorem 1 *Given the subspace $\mathcal{H}_t(v)$, there exists a unique representation* (18) *of fundamentals and beliefs.*

Proof. Given $\mathcal{H}_t(v)$ the subspace $\mathcal{H}_t(w)$ is uniquely determined and for $x_t \in \mathcal{H}_t(\hat{x})$, it is clear that $x_t \in \mathcal{H}_t(w)$, $x_t \perp \mathcal{H}_t(v)$. Decompose $\mathcal{H}_t(w)$ as

$$\mathcal{H}_t(w) = \bigoplus_{i=0}^{\infty} \mathcal{D}_{t-i}(w),$$

and project x_t into this family of orthogonal subsets to obtain

$$x_t = \sum_{i=0}^{\infty} w_{i,t-i}, \qquad w_{i,t-i} \in \mathcal{D}_{t-i}(w).$$

Following the arguments in CJ, it is possible to apply the Gram-Schmidt procedure to $w_{i,t-i}$ and derive the set of indices $\mathbb{I}$ as well as a collection of stationary processes $a_{i,t-i}$ for $i \in \mathbb{I}$ that satisfy (18). Because all $w_{i,t-i}$ are uniquely determined by the orthogonal projections, so are all $a_{i,t-i}$. The uniqueness of orthogonal projections also guarantees that the representation of signals $s_{i,t}$ is unique. ■

It is important to emphasize that the uniqueness of the representation depends on the choice of $\mathcal{H}_t(v)$. If we consider a different space spanned by noise processes $\mathcal{H}_t(v') \neq \mathcal{H}_t(v)$, then we obtain a different $\mathcal{H}_t(w')$ and an alternative representation

$$x_t = \sum_{i \in \mathbb{I}'} a'_{i,t-i}, \qquad s'_{i,t} = a'_{i,t} + v'_{i,t}, \qquad i \in \mathbb{I}', \tag{19}$$

where $\mathbb{I}'$ is a set of indices, and all $a'_{i,t} \in \mathcal{H}_t(w')$, $v'_{i,t} \in \mathcal{H}_t(v')$ are stationary processes such that $a'_{i,t} \perp a'_{j,t} \perp v'_{k,t}$ for all $i \neq j$ and all $k \in \mathbb{I}'$. This representation is also unique, given $\mathcal{H}_t(v')$. Evidently, there is nothing special about the uniqueness of a pure news representation; it is simply the representation we obtain given $\mathcal{H}_t(v') = \{0\}$, which defines a certain partition of the agents' information space.

We distinguish the set of indices $\mathbb{I}'$ in representation (19) from $\mathbb{I}$ in representation (18) to underscore the fact that for a different specification of $\mathcal{H}_t(v')$, the resulting representation may have completely different news shocks—i.e., the number of news shocks may differ as well as the timing of their impact on x_t. Nevertheless, to generalize the claim stated in CJ, expectations under all such alternative representations remain the same.

Theorem 2 *If $\hat{x}_{i,t}$ is the rational expectation at time t of the fundamental x_{t+i} under representation* (18) *and $\hat{x}'_{i,t}$ is the corresponding expectation under representation* (19)*, then $\hat{x}_{i,t} = \hat{x}'_{i,t}$ almost surely for all i and t.*

Proof. Since $\mathcal{H}_t(s) = \mathcal{H}_t(\hat{x}) = \mathcal{H}_t(s')$, we immediately obtain $\hat{x}_{i,t} = \mathrm{E}(x_{t+i} \,|\, \mathcal{H}_t(s)) = \mathrm{E}(x_{t+i} \,|\, \mathcal{H}_t(\hat{x})) = \mathrm{E}(x_{t+i} \,|\, \mathcal{H}_t(s')) = \hat{x}'_{i,t}$. ■

Theorem 2 implies that observing agents *expectations* about future realizations of the fundamentals process reveal absolutely no useful information about the underlying news/noise structure because all representations yield equal (almost surely) fundamentals and expectations.[11] This conclusion holds even if the expectations are perfectly observed for all horizons $t+1, t+2, \cdots$. Indeed, observing expectations is not sufficient to even identify the number or type of news shocks in a setting where noise shocks are potentially present. Simply, it is altogether irrelevant information for the econometrician.

However, all indeterminacy disappears if signals are observed instead of (or in addition to) expectations. To fix ideas, assume $\mathbb{I}$ contains n elements and therefore at time t agents observe n signals $\{s_{i,t}\}_{i\in\mathbb{I}}$. Let d_t be the $n+1\times 1$ vector containing the fundamental x_t along with the n signals $s_{i,t}$ generated at time t. Likewise, let v_t be the $n\times 1$ vector consisting of noise variables $v_{i,t}$. The question of interest is whether the observed process $\{d_t\}$ can be used to identify the distributions of $\{v_t\}$ and $\{a_{i,t}\}$ for all $i \in \mathbb{I}$.

Recall that the distribution of a n-variate stationary Gaussian process $\{y_t\}$ is uniquely determined by its covariance function. Therefore, two such processes $\{y_t\}$ and $\{y'_t\}$ are equal in distribution, denoted $\{y_t\} \overset{d}{=} \{y'_t\}$, if and only if $\text{Cov}(y_{i,t-|k|}, y_{j,t-|k|}) = \text{Cov}(y'_{i,t-|k|}, y'_{j,t-|k|})$ for all t, i, j, and k, where by convention $\text{Cov}(y_{i,t}, y_{i,t}) \equiv \text{Var}(y_{i,t})$. The following theorem demonstrates that indeed observations of fundamentals and signals uniquely determine the distributions of $\{v_t\}$ and $\{a_{i,t}\}$ for all $i \in \mathbb{I}$.

Theorem 3 *Let $\{d_t\}$ be the fundamentals and signals generated by representation* (18) *and $\{d'_t\}$ be the fundamentals and signals generated by representation* (19)*. Then $\{d_t\} \overset{d}{=} \{d'_t\}$ if and only if $\mathbb{I} = \mathbb{I}'$, $\{a_{i,t}\} \overset{d}{=} \{a_{i,t}\}$ for all $i \in \mathbb{I}$ and $\{v_t\} \overset{d}{=} \{v'_t\}$.*

Proof. The sufficiency is trivial since the assumption that $\{x_t\}$ is a stationary Gaussian process implies $\{a_{i,t}\}$, $\{a'_{i,t}\}$, $\{v_{i,t}\}$ and $\{v'_{i,t}\}$ are stationary Gaussian processes for all $i \in \mathbb{I} = \mathbb{I}'$. Therefore, if $\{a_{i,t}\} \overset{d}{=} \{a'_{i,t}\}$ and $\{v_t\} \overset{d}{=} \{v'_t\}$, identical linear transformations of these processes are also equal in distribution.

To prove the necessity, we first show that $\{d_t\} \overset{d}{=} \{d'_t\}$ implies $\mathbb{I} = \mathbb{I}'$. Assume $\mathbb{I} \neq \mathbb{I}'$ and (without loss of generality) there exists $i \neq 0$ such that $i \in \mathbb{I}$ and $i \notin \mathbb{I}'$. From the definitions of the two representations and the stationarity of both $a_{i,t}$ and $a'_{j,t}$, we obtain

$$
\begin{aligned}
\text{E}(x_t s_{i,t-i}) &= \text{E}(a^2_{i,t-i}) = \text{Var}(a_{i,t}),\\
\text{E}(x_t s'_{j,t-i}) &= \text{E}(a'_{j,t-j} a'_{j,t-i}) = \text{Cov}(a'_{j,t}, a'_{j,t-|k|}),\\
\text{E}(x_t s_{i,t-j}) &= \text{E}(a_{i,t-i} a_{i,t-j}) = \text{Cov}(a_{i,t}, a_{i,t-|k|}),\\
\text{E}(x_t s'_{j,t-j}) &= \text{E}((a'_{j,t-j})^2) = \text{Var}(a'_{j,t}),
\end{aligned}
$$

[11]This type of equality is in fact stronger than *observationally equivalent* in the sense of *equal in distribution*, as stated in CJ.

where $k = i - j$. Since $a_{i,t}$ and $a'_{j,t}$ are stationary Gaussian processes, it must be the case that $\text{Var}(a_{i,t}) > |\text{Cov}(a_{i,t}, a_{i,t-|k|})|$ and $\text{Var}(a'_{j,t}) > |\text{Cov}(a'_{j,t}, a_{j,t-|k|})|$ for all $k \neq 0$. However, $\{d_t\} \stackrel{d}{=} \{d'_t\}$ implies there exist $s_{i,t}$ and $s'_{j,t}$ with $j \in \mathbb{I}'$, such that $\text{E}(x_t s_{i,t-i}) = \text{E}(x_t s'_{j,t-i})$ and $\text{E}(x_t s_{i,t-j}) = \text{E}(x_t s'_{j,t-j})$. But this yields

$$\text{Cov}(a'_{j,t}, a'_{j,t-|k|}) = \text{Var}(a_{i,t}) > \text{Cov}(a_{i,t}, a_{i,t-|k|}) = \text{Var}(a'_{j,t}),$$

which is a contradiction. Therefore, we conclude $\{d_t\} \stackrel{d}{=} \{d'_t\}$ implies $\mathbb{I} = \mathbb{I}'$.

With identical sets of indices, consider the following moments implied by each representation:

$$\begin{aligned}
\text{E}(x_t s_{i,t-i-|k|}) &= \text{Cov}(a_{i,t}, a_{i,t-|k|}), & \text{E}(x_t s'_{i,t-i-|k|}) &= \text{Cov}(a'_{i,t}, a'_{i,t-|k|}),\\
\text{E}((x_t - \sum_{i\in\mathbb{I}} s_{i,t-i})x_{t-i-|k|}) &= \text{Cov}(a_{0,t}, a_{0,t-|k|}) & \text{E}((x_t - \sum_{i\in\mathbb{I}} s'_{i,t-i})x_{t-i-|k|}) &= \text{Cov}(a'_{0,t}, a'_{0,t-|k|}),\\
\text{E}((s_{i,t-i} - x_t)s_{i,t-i-|k|}) &= \text{Cov}(v_{i,t}, v_{i,t-|k|}), & \text{E}((s'_{i,t-i} - x_t)s'_{i,t-i-|k|}) &= \text{Cov}(v'_{i,t}, v'_{i,t-|k|}),\\
\text{E}(s_{i,t} s_{j,t-|k|}) &= \text{Cov}(v_{i,t}, v_{j,t-|k|}), & \text{E}(s'_{i,t} s'_{j,t-|k|}) &= \text{Cov}(v'_{i,t}, v'_{j,t-|k|}),
\end{aligned}$$

for all $k \in \mathbb{Z}$ and $j \neq i$, where we have once again employed the fact that all processes are stationary.

In this case, $\{d_t\} \stackrel{d}{=} \{d'_t\}$ implies

$$\begin{aligned}
\text{E}(x_t s_{i,t-i-|k|}) &= \text{E}(x_t s'_{i,t-i-|k|}),\\
\text{E}((x_t - \sum_{i\in\mathbb{I}} s_{i,t-i})x_{t-i-|k|}) &= \text{E}((x_t - \sum_{i\in\mathbb{I}} s'_{i,t-i})x_{t-i-|k|}),\\
\text{E}((s_{i,t-i} - x_t)s_{i,t-i-|k|}) &= \text{E}((s'_{i,t-i} - x_t)s'_{i,t-i-|k|}),\\
\text{E}(s_{i,t} s_{j,t-|k|}) &= \text{E}(s'_{i,t} s'_{j,t-|k|})
\end{aligned}$$

for all t, i, j, and k. Therefore, the covariance function of each $\{a_{i,t}\}$ is equivalent to that of $\{a'_{i,t}\}$ for all $i \in \mathbb{I}$ and the covariance function of $\{v_t\}$ is equivalent to that of $\{v'_t\}$. Since all processes involved are Gaussian, this implies $\{a_{i,t}\} \stackrel{d}{=} \{a_{i,t}\}$ for all $i \in \mathbb{I}$ and $\{v_t\} \stackrel{d}{=} \{v'_t\}$. ■

CJ also state that any *endogenous* variable $c_t \in \mathcal{H}(\hat{x})$ is observationally equivalent under alternative representations, which implies that marginally $\{s_{i,t}\} \stackrel{d}{=} \{s'_{i,t}\}$ for all $i \in \mathbb{I}$ and $i \in \mathbb{I}'$. This is of course true since the signal itself contains no information regarding what shocks actually impact the fundamentals process. It is, in fact, the dynamic interaction between the fundamentals and signals that convey the roles of news and noise in driving agents' decisions as well as the evolution of the fundamentals process, which allows us to disentangle the two effects and identify a unique representation. In terms of the theory discussed in preceding sections, this is manifested by the news component in the signal ultimately having a permanent impact on the fundamental and the effect of the noise component having no permanent effect on any variable.

6 Conclusions

Noisy-news models involve a signal which contains a combination of news and noise. In this note, we have shown that CJ's finding that news and noise shocks are observationally equivalent holds when the econometrician observes fundamentals and expectations, but does not hold when the econometrician can either directly observe the signal or learn about it using a macroeconomic model. We argue that the latter case is the more realistic one. Macroeconomists do work with structural models such as the RBC model considered in this paper, which allow for signals to be extracted. Thus, the observational equivalence result of CJ, although correct in its own terms, does not have wide applicability for macroeconomists interested in learning about the relative roles of news and noise in driving economic fluctuations.

We conclude by directly addressing the question that forms the title of this note. We believe that, with the models and data sets typically used in this literature, the answer is "Yes".

References

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., J. Chan, E. Eisenstat, and G. Koop (2018): "Identifying Noise Shocks," Universities of Bern, Queensland, Strathclyde, and Australian National University, mimeo.

Chahrour, R., and K. Jurado (2018): "News or Noise? The Missing Link," *American Economic Review*, Forthcoming.

Dadda, C., and A. Scorcu (2003): "On the Time Stability of the Output-Capital Ratio," *Economic Modelling*, 20(6), 1175–1189.

Forni, M., L. Gambetti, M. Lippi, and L. Sala (2017): "Noisy News in Business Cycles," *American Economic Journal: Macroeconomics*, forthcoming, –.

Hamilton, J. (1994): *Time Series Analysis*. Princeton, NJ, Princeton University Press.

Lippi, M., and L. Reichlin (1994): "VAR Analysis, Nonfundamental Representations, Blaschke Matrices," *Journal of Econometrics*, 63, 307–325.

Uhlig, H. (1999): "A Toolkit for Analysing Nonlinear Dynamic Stochastic Models Easily," *in Ramon Marimon and Andrew Scott, eds, Computational Methods for the Study of Dynamic Economies, Oxford University Press*, pp. 30–61.