News or noise? The missing link

Authors: Ryan Chahrour, Kyle Jurado

Boston College Working Papers in Economics, 2016

Originally posted on: http://ideas.repec.org/p/boc/bocoec/917.html

News or Noise? The Missing Link

Ryan Chahrour — Boston College

Kyle Jurado — Duke University

September 7, 2016

Abstract

The macroeconomic literature on belief-driven business cycles treats news and noise as distinct representations of people's beliefs about economic fundamentals. We prove that these two representations are actually observationally equivalent. This means that the decision to use one representation or the other must be made on theoretical, and not empirical, grounds. Our result allows us to determine the importance of beliefs as an independent source of fluctuations. Using three prominent models from this literature, we show that existing research has understated the importance of independent shocks to beliefs. This is because representations with anticipated and unanticipated shocks mix the fluctuations due independently to beliefs with the fluctuations due to fundamentals. We also argue that the observational equivalence of news and noise representations implies that structural vector auto-regression analysis is equally appropriate for recovering both news and noise shocks.

1 Introduction

A large literature in macroeconomics has argued that changes in people's beliefs about the future can be an important cause of economic fluctuations. This idea, which dates at least to Pigou (1927), has been more recently mathematically formalized in two different ways. The first way, which we call a "news representation," models people as perfectly observing some (but not all) parts of an exogenous fundamental in advance. By way of analogy, this is like learning for sure today that in next week's big game your favorite team will win the first half. You don't know whether they will win the game, which is ultimately what you care about, because you are still unsure how the second half will turn out. The second way, which we call a "noise representation," models people as imperfectly observing some (possibly all) parts of an exogenous fundamental in advance. This is like your friend telling you that he thinks your team will win next week's game. He follows the sport much more than you do, and is often right, but sometimes he gets it wrong.

At first glance, these two different ways of representing people's beliefs may seem only superficially similar. In both cases, people are getting some advance information about the future. But on a news view they have perfect information and can fully trust whatever information they receive, while on a noise view they have imperfect information and need to solve a signal extraction problem to determine their best forecast. In their recent review of the literature on belief-driven business cycles, Beaudry and Portier (2014) have this to say about the relationship between the two formulations:

> "While these two formulations may appear almost identical, they are actually quite different...To give an idea of the difference, in the [noise] formulation there is a shock which can be referred to as a noise or error shock...In the [news] formulation there is no direct counterpart: there is an anticipated shock and an unanticipated shock, but no noise shock." (p.998)

This paper argues that, on the contrary, it is actually more accurate to think of news and noise representations as superficially different, but fundamentally the same. Specifically, we prove that from the perspective of data on exogenous fundamentals and people's beliefs about those fundamentals, these two representations are observationally equivalent. There is nothing in the data that can discriminate between them.

We present a constructive proof of this result using Hilbert space methods. The fact that the proof is constructive means that it also provides an explicit method for passing from one representation to the other. We analytically derive these restrictions in several cases of interest from the literature.

The intuition for our main result is that news shocks are linear transformations of the Wold innovations implied by the noise representation. Just as every finite-dimensional signal extraction model has associated with it a unique observationally equivalent "innovations representation" (cf. Anderson and Moore, 1979, ch. 9), so every noise model has a unique observationally equivalent news representation. Because the Wold innovations are contained in the space spanned by the history of variables that agents observe, the news representation is a way of writing models with noise "as if" people have perfect information.

The equivalence between news and noise representations is the key to answering the most basic question in the literature on belief-driven fluctuations: how important are beliefs as an independent source of economic fluctuations? This is a question that the existing literature, which uses models either with news or some combination of news and noise, has not been able to answer. News shocks can change beliefs on impact without any corresponding change in current fundamentals, but they are tied by construction to changes in future fundamentals. As a result, variance decompositions computed in terms of news shocks mix changes due to beliefs with changes due to fundamentals.[1] By contrast, in the type of noise representation we propose, noise shocks capture precisely those movements in beliefs that are independent of fundamentals at all horizons. To isolate the importance of purely belief-driven fluctuations, therefore, we need to compute variance decompositions in terms of these noise shocks.

One caveat is that an infinite number of different observationally equivalent noise representations are associated with any news representation. This is essentially for the same reason that state-space representations of stochastic processes are not unique. Many different signal structures can generate the same Wold representation. The implication is that for any news representation, it is always possible to find an observationally equivalent noise representation, but some statistics of interest, like the impulse responses of endogenous variables to noise shocks, can only be uniquely determined under additional restrictions.

[1]This point has been emphasized in the literature. For example, see the discussion in Section IV.A of Barsky, Basu, and Lee (2015).

Even though a particular news representation does not uniquely determine the signal structure in an observationally equivalent noise representation, we prove that it does uniquely determine the importance of noise shocks. For any signal structure, as long as it generates dynamics that are observationally equivalent to the news representation, the share of the variance of any endogenous variable attributable to noise shocks is uniquely pinned down. This result is important because it means that in order to determine the importance of beliefs in a model with news shocks, we can look at the importance of noise shocks in any observationally equivalent noise version of the model.

We use our equivalence result to determine the importance of beliefs as an ultimate cause of business cycles in three different quantitative models of U.S. business cycles. The three models come from Schmitt-Grohé and Uribe (2012), Barsky and Sims (2012), and Blanchard, L'Huillier, and Lorenzoni (2013). These models all appear to have very different information structures, which — combined with various differences in the rest of the physical environment, estimation procedure, and data sample — has made it difficult to compare results across models. By allowing us to isolate the independent contribution of beliefs in each model, our equivalence result provides a way of coherently comparing them. We perform the analysis using the exact models and estimated parameters from the original studies; our contribution is to determine what those models imply about the independent contribution of beliefs.

In all three cases, the importance of noise shocks has been understated. In the model of Schmitt-Grohé and Uribe (2012), there is no shock labeled "noise," but the actual contribution of independent fluctuations to beliefs is somewhere between 3 and 11 percent depending on the variable. In the model of Barsky and Sims (2012), noise shocks are responsible for 9 percent of the fluctuations in consumption, which is almost an order of magnitude larger than the original estimate of 1 percent. In the model of Blanchard, L'Huillier, and Lorenzoni (2013), the contribution of noise to consumption is 57 percent, compared to an original estimate of 44 percent.

Our results demonstrate that there remains substantial disagreement across models regarding the importance of beliefs. We do not try to settle that disagreement in this paper. However, a consistent result across all three models is that noise shocks are not very important for explaining fluctuations in investment. The contribution of noise shocks for investment never rises above 11 percent, which is the estimate from the model of Schmitt-Grohé and Uribe (2012). This is particularly striking given the fact that forward-looking investment decisions often play an important role in the motivation and discussion of belief-driven business cycles.

The observational equivalence of news and noise representations is also relevant to the discussion of whether structural vector auto-regression (VAR) analysis is appropriate for recovering noise shocks. We show that in principle, structural VARs can be used to recover noise shocks and their associated impulse responses even though noise representations are not invertible. This is because in both cases, the underlying shocks are only one orthogonal transformation away from the reduced-form representation. An implication of our argument is that invertibility should not be viewed as a necessary condition for the applicability of structural VAR analysis.

We provide one orthogonal transformation that is sufficient to uniquely determine noise shocks (and their associated impulse response functions). This transformation is closely related to a popular thought experiment in the literature on news shocks. The thought experiment is as follows: at date t, agents receive advance information concerning fundamentals at some future date $T > t$. But a surprise innovation at that future date T exactly offsets the advance information agents had previously received. So their expectations end up being incorrect after the fact. This experiment is one way that several authors have tried to separate the effect of beliefs from fundamentals. It turns out that under the set of restrictions we provide, noise shocks generate exactly the combination of offsetting news shocks envisioned by this experiment.

2 Observational Equivalence

News and noise representations are two different ways of describing economic fundamentals and people's beliefs about them. "Fundamentals" are stochastic processes capturing exogenous changes in technology, preferences, endowments, or government policy. Throughout this section, fundamentals are summarized by a single scalar process $\{x_t\}$. People's decisions depend on expected future realizations of x_t, so both representations specify what people can observe at each date and how they use their observations to form beliefs about the future.

The main result of the paper, which is presented in this section, is an observational equivalence theorem relating news and noise representations. To facilitate the exposition, the first subsection presents the result in a simple example with news or noise regarding fundamentals only one period in the future while the second subsection presents the general equivalence result.

2.1 Simple Example

In the simplest of news representations, x_t is equal to the sum of two shocks, $a_{0,t}$ and $a_{1,t-1}$, which are independent and identically distributed (i.i.d.) over time, and which are independent of one another:

$$x_t = a_{0,t} + a_{1,t-1}, \qquad \begin{bmatrix} a_{0,t} \\ a_{1,t} \end{bmatrix} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \begin{bmatrix} \sigma_{a,0}^2 & 0 \\ 0 & \sigma_{a,1}^2 \end{bmatrix}\right). \tag{1}$$

At each date t, people observe the whole history of the two shocks up through that date, $\{a_{0,\tau}, a_{1,\tau}\}$ for all integers $\tau \leq t$. Their beliefs regarding fundamentals are rational; the probabilities they assign to future outcomes are exactly those implied by system (1). The shock $a_{1,t}$ is a news or anticipated shock because people see it at date t but it doesn't affect the fundamental until date $t+1$. The shock $a_{0,t}$ is a surprise or unanticipated shock.

Now consider instead a noise representation. The fundamental variable x_t is i.i.d. over time, and there is a noisy signal of the fundamental one period into the future:

$$s_t = x_{t+1} + v_t, \qquad \begin{bmatrix} x_t \\ v_t \end{bmatrix} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}\right). \tag{2}$$

At each date t, people observe the whole history of fundamentals and signals up through that date, $\{x_\tau, s_\tau\}$ for all integers $\tau \leq t$. Even though people only have imperfect information about x_{t+1}, their beliefs are nevertheless still rational. The shock v_t is a noise or error shock because it affects beliefs even though it is totally independent of fundamentals.

Our main point is that these two representations are observationally equivalent. But before making that point explicitly, it is important to be clear about what types of things we are considering to be "observable." 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.

An important feature of our concept of equivalence is that we treat beliefs, as well as fundamentals, as observable. We take this approach for three reasons. First, it is a stronger condition; observational equivalence with respect to a larger set of observables implies observational equivalence with respect to any smaller set of those observables. Second, beliefs are observable in economics, in principle. Beliefs may be measured directly, using surveys, or indirectly, using the mapping between beliefs and actions implied by an economic model. That actions reflect beliefs is, after all, a basic motivation for the literature on belief-driven fluctuations. Third, in a broad class of linear rational expectations models with unique equilibria, endogenous processes are purely a function of current and past fundamentals and beliefs about future fundamentals. So observational equivalence of fundamentals and beliefs implies observational equivalence of the entire economy.

We would also like to emphasize that the observability of beliefs distinguishes our concept of observational equivalence from the typical conception of observational equivalence one often encounters in time series analysis. To use a familiar example (cf. Hamilton, 1994, pp. 64-67) , consider the MA(1) process

$$y_t = \epsilon_t - \theta\epsilon_{t-1}, \quad \epsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\epsilon^2),$$

with $|\theta| < 1$, and view this as a simple full-information rational expectations model for the determination of the process $\{y_t\}$ in terms of the exogenous shocks $\{\epsilon_t\}$. As is well-known, if only $\{y_t\}$ is observable to an econometrician, this invertible representation of the model is observationally equivalent to the non-invertible representation

$$y_t = u_t - \psi u_{t-1}, \quad u_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_u^2),$$

if and only if $\psi = 1/\theta$ and $\sigma_\epsilon^2 = \theta^2\sigma_u^2$. The proof relies on the fact that under those parametric relations, the spectral density function of $\{y_t\}$ is the same under both representations, $f_y(\omega) = \sigma_\epsilon^2(1-\theta e^{-i\omega})(1-\theta e^{i\omega})$. However, under our (stronger) conception of observational equivalence, these two representations are no longer the same. To see why, note that the spectral density function of the one-step-ahead rational forecast $\hat{y}_t \equiv E_t[y_{t+1}]$ is given by $f_{\hat{y}}(\omega) = \theta^2\sigma_\epsilon^2$ under the invertible representation, but $f_{\hat{y}}(\omega) = \sigma_\epsilon^2$ under the non-invertible representation.

The following proposition formally states the observational equivalence result for the simple example of this subsection, and provides the parametric mapping from one representation to the other. Its proof is collected together with all later proofs in Appendix (A).

Proposition 1. *The news representation of fundamentals and beliefs in system* (1) *is observationally equivalent to the noise representation of fundamentals and beliefs in system* (2) *if and only if:*

$$\sigma_x^2 = \sigma_{a,0}^2 + \sigma_{a,1}^2 \quad \textit{and} \quad \frac{\sigma_v^2}{\sigma_x^2} = \frac{\sigma_{a,0}^2}{\sigma_{a,1}^2}.$$

The intuition behind the result comes from the fact that the noise representation implies an observationally equivalent innovations representation (cf. Anderson and Moore, 1979, ch.9) of the form:

$$\begin{aligned} x_t &= \hat{x}_{t-1} + w_{0,t} \\ \hat{x}_t &= \kappa w_{1,t}, \end{aligned} \tag{3}$$

where $\kappa = \sigma_x^2/(\sigma_x^2 + \sigma_v^2)$ is a Kalman gain parameter controlling how much people trust the noisy signal, and $w_t \equiv (w_{0,t}, w_{1,t})'$ is the vector of Wold innovations, which evolves over time according to

$$w_t \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \begin{bmatrix} \kappa\sigma_v^2 & 0 \\ 0 & \sigma_x^2 + \sigma_v^2 \end{bmatrix}\right).$$

But system (3) is the same as the news representation in system (1) when $a_{0,t} = w_{0,t}$ and $a_{1,t} = \kappa w_{1,t}$. The news shocks are linear combinations of the Wold innovations.

A direct implication of Proposition (1) is that the news representation is identified if and only if the noise representation is identified. By observational equivalence, both representations have the same likelihood function. Therefore, because the relations in Proposition (1) define a bijection, it is always possible to go from one set of parameters to the other and vice versa.

Corollary 1. *The parameters of the news representation in system* (1) *are uniquely identified if and only if the parameters of the noise representation in system* (2) *are uniquely identified.*

This corollary hints at how the observational equivalence between news and noise representations is relevant to the question of whether structural VAR analysis can be applied to models with noise shocks. We take up that question in Section (5).

2.2 General Equivalence Result

This subsection generalizes the previous example to allow for news and noise at multiple future horizons, and potentially more complex time-series dynamics. To fix notation, we use $\mathcal{L}^2$ to denote the space of (equivalence classes of) random variables with finite second moments, which is a Hilbert space when equipped with the inner product $\langle a, b \rangle = E[ab]$ for any $a, b \in \mathcal{L}^2$. Completeness of this space is with respect to the norm $\|a\| \equiv \langle a, a \rangle^{1/2}$. For any collection of random variables in $\mathcal{L}^2$,

$$\{y_{i,t}\}, \quad \text{with } i \in \mathcal{I} \subseteq \mathbb{Z} \text{ and } t \in \mathbb{Z},$$

we let $\mathcal{H}_t(y)$ denote the closed subspace spanned by the variables $y_{i,\tau}$ for all $i \in \mathcal{I}$ and $\tau \in \mathbb{Z}$ such that $\tau \leq t$. To simplify notation, we write $\mathcal{H}(y) \equiv \mathcal{H}_\infty(y)$.

Fundamentals are summarized by a scalar discrete-time process $\{x_t\}$. As in the previous subsection, this process is taken to be mean-zero, stationary, and Gaussian. The fact that fundamentals are summarized by a scalar process is not restrictive; we can imagine a number of different scalar processes, each capturing changes in one particular fundamental. In that case it will be possible to apply the results from this section to each fundamental one at a time.

People's beliefs about fundamentals are summarized by a collection of random variables $\{\hat{x}_{i,t}\}$, with $i, t \in \mathbb{Z}$, where $\hat{x}_{i,t}$ represents the forecast of the fundamental realization x_{t+i} as of time t. Under the assumption of rational expectations, which is maintained throughout this paper, $\hat{x}_{i,t}$ is equal to the mathematical expectation of x_{t+i} with respect to a particular date-t information set. This, together with the fact that fundamentals are Gaussian, implies that the collection $\{\hat{x}_{i,t}\}$ fully characterizes people's entire subjective distribution over realizations of the sequence $\{x_t\}$.

A "representation of fundamentals and beliefs" means a specification of the fundamental process $\{x_t\}$ and the collection of people's conditional expectations about that process at each point in time $\{\hat{x}_{i,t}\}$. A typical assumption is that people's information set is equal to $\mathcal{H}_t(x)$, so $\hat{x}_{i,t} \in \mathcal{H}_t(x)$ for all $t \in \mathbb{Z}$. In this case, the process $\{x_t\}$ is itself sufficient to describe both the fundamental and people's beliefs about it. A key departure in models of belief-driven fluctuations is that people may have more information than what is reflected in $\mathcal{H}_t(x)$ alone; as a result, $\mathcal{H}_t(x) \subset \mathcal{H}_t(\hat{x})$. We will therefore maintain this assumption throughout the paper. We also work exclusively with processes that are regular, in the sense of Rozanov (1967).

Definition 1. In a "news representation" of fundamentals and beliefs, the process $\{x_t\}$ is related to a collection of independent, stationary Gaussian processes $\{a_{i,t}\}$ with $i \in \mathcal{I} \subseteq \mathbb{Z}_+$ by the summation

$$x_t = \sum_{i \in \mathcal{I}} a_{i,t-i} \quad \text{for all } t \in \mathbb{Z},$$

where people's date-t information set is $\mathcal{H}_t(a) \supset \mathcal{H}_t(x)$.

The idea behind this representation is that people observe parts of the fundamental realization x_t prior to date t. The variable $\epsilon^a_{i,t} \equiv a_{i,t} - E[a_{i,t}|\mathcal{H}_{t-1}(a)]$ is called the "news shock" associated with horizon i whenever $i > 0$. By convention, $0 \in \mathcal{I}$, and in that case, the variable $\epsilon^a_{0,t}$ is referred to as the "surprise shock." An important aspect of this definition is that all of the news shocks are correlated both with fundamentals and people's beliefs. This is because any increase in fundamentals that people observe in advance must generate a one-for-one increase in fundamentals at some point in the future.

Example 1. In the model of Schmitt-Grohé and Uribe (2012) (see their Section IV), each fundamental process $\{x_t\}$ follows a law of motion of the form (in deviations from its mean) :

$$x_t = \rho_x x_{t-1} + \epsilon^a_{0,t} + \epsilon^a_{4,t-4} + \epsilon^a_{8,t-8}, \quad \begin{bmatrix} \epsilon^a_{0,t} \\ \epsilon^a_{4,t} \\ \epsilon^a_{8,t} \end{bmatrix} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \begin{bmatrix} \sigma^2_{a,0} & 0 & 0 \\ 0 & \sigma^2_{a,4} & 0 \\ 0 & 0 & \sigma^2_{a,8} \end{bmatrix}\right).$$

where $0 < \rho_z < 1$. In terms of Definition (1), this means that $\mathcal{I} \equiv \{0, 4, 8\}$ and

$$\begin{aligned} x_t &= a_{0,t} + a_{4,t-4} + a_{8,t-8} \\ a_{0,t} &= \rho_z a_{0,t-1} + \epsilon^a_{0,t} \\ a_{4,t} &= \rho_z a_{4,t-1} + \epsilon^a_{4,t} \\ a_{8,t} &= \rho_z a_{8,t-1} + \epsilon^a_{8,t}. \end{aligned}$$

Also, it is easy to see that $\mathcal{H}_t(x) \subset \mathcal{H}_t(a)$. △

Definition 2. In a "noise representation" of fundamentals and beliefs, there is a collection of signal processes $\{s_{i,t}\}$ with $i \in \mathcal{I} \subseteq \mathbb{Z}_+$ of the form:

$$s_{i,t} = m_{i,t} + v_{i,t}, \quad \text{for all } t \in \mathbb{Z},$$

where $m_{i,t} \in \mathcal{H}(x)$, $v_{i,t} \perp \mathcal{H}(x)$, and people's date-$t$ information set is $\mathcal{H}_t(s) \supset \mathcal{H}_t(x)$, which satisfies $\mathcal{H}_t(s) = \mathcal{H}_t(\hat{x})$.

The idea behind this representation is that people may receive signals about the fundamental realization x_t prior to date t, but those signals are contaminated with noise. The variable $\epsilon^v_{i,t} \equiv v_{i,t} - E[v_{i,t}|\mathcal{H}_{t-1}(v)]$ is called the "noise shock" associated with signal i. The variable $\epsilon^x_t \equiv x_t - E[x_t|\mathcal{H}_{t-1}(x)]$ is called the "fundamental shock." An important aspect of this definition is that all of the noise shocks are completely independent of fundamentals, but because people cannot separately observe $m_{i,t}$ and $v_{i,t}$ at date t, their beliefs are still affected by noise. The condition that $\mathcal{H}_t(s) = \mathcal{H}_t(\hat{x})$ simply rules out redundant or totally uninformative signals.

Example 2. In the numerical implementation of their baseline model, Beaudry and Portier (2014) specify the fundamental process $\{x_t\}$ (in deviations from its mean) and signal process $\{s_t\}$ as (see their Section 2.1):

$$
\begin{aligned}
x_t &= \rho_x x_{t-1} + \epsilon_t \\
s_t &= \epsilon_{t+8} + v_{1,t},
\end{aligned}
$$

where $0 \leq \rho_z < 1$ and

$$
\begin{bmatrix} \epsilon_t \\ v_{1,t} \end{bmatrix} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_{v,1}^2 \end{bmatrix}\right).
$$

In terms of Definition (2), this means that $\mathcal{I} \equiv \{0, 8\}$ and

$$
\begin{aligned}
s_{0,t} &= x_t \\
s_{1,t} &= (x_{t+8} - \rho_x x_{t+7}) + v_{1,t}
\end{aligned}
$$

so $m_{0,t} = x_t$, $m_{1,t} = x_{t+8} - \rho_x x_{t+7}$, and $v_{0,t} = 0$. Also, note that $\mathcal{H}_t(x) \subset \mathcal{H}_t(s)$. $\triangle$

With these definitions in hand, we are ready to state the main result of the paper.

Theorem 1. *Fundamentals and beliefs always have both a news representation and a noise representation. Moreover, the news representation is unique.*

This theorem clarifies the sense in which news and noise representations of fundamentals and beliefs are really just two sides of the same coin. It is possible to view the same set of data from either perspective. This result is the basis for our claim that it is more accurate to think of news and noise representations as superficially different, but fundamentally the same. The proof of the theorem is constructive, which means that it also provides an explicit computational method for passing from one representation to the other.

The only asymmetric aspect of the theorem involves the uniqueness of the two representations. Any particular news representation will be compatible with several different noise representations. This is the same sort of asymmetry present between signal models representations and innovations representations in the literature on state-space models. In general there exist infinitely many signal models with the same innovations representation (cf. Anderson and Moore, 1979, pp. 224-226). We argue in the subsequent sections, however, that despite this multiplicity of noise representations, most interesting economic questions still have a unique answer.

An important implication of Theorem (1) is that associated with any model economy in which fundamentals and beliefs are expressed in the form of a news representation is an observationally equivalent economy in which fundamentals and beliefs are expressed in the form of a noise representation, and vice versa. This is because the observational equivalence of fundamentals and beliefs implies the observational equivalence of any endogenous processes that are functions of fundamentals and beliefs. To make this statement more precise, we define here what we mean by an endogenous process, and then present this statement as a proposition.

Definition 3. Given a fundamental process $\{x_t\}$ and a collection of forecasts $\{\hat{x}_{i,t}\}$ satisfying $\mathcal{H}_t(x) \subset \mathcal{H}_t(\hat{x})$, a process $\{c_t\}$ is "endogenous" with respect to $\{x_t\}$ if

$$c_t \in \mathcal{H}_t(\hat{x}) \quad \text{for all } t \in \mathbb{Z}.$$

Proposition 2. *If two different representations of fundamentals and beliefs are observationally equivalent, then they imply observationally equivalent dynamics for any endogenous process.*

3 The Importance of Beliefs

The most basic question in the literature on belief-driven fluctuations is: how important are beliefs? Or more specifically, how important are beliefs relative to fundamentals? Perhaps surprisingly, it turns out that no existing quantitative study in this literature has actually answered that question. Some studies report the importance of news shocks, which combine the contribution due to fundamentals with the contribution due independently to beliefs. Others include noise shocks and news shocks in the same model, and as a result, do not properly report the importance of either one. In this section we argue that the observational equivalence result in Theorem (1) is the key to determining the importance of beliefs as an independent cause of fluctuations.

The first subsection explains the problem with using news shocks to determine the importance of beliefs, and the second subsection clarifies the problems that arise when attempting to include both news and noise shocks in the same model. To keep things clear, the discussion of both of these issues is framed in terms of the simple example from Section (2.1). The third subsection establishes an important result regarding the uniqueness of variance decompositions.

3.1 The Problem with News Shocks

In the context of dynamic linear models, the importance of a set of exogenous shocks can be determined by performing a variance decomposition. This entails computing the model-implied variance of an endogenous process under the assumption that all shocks other than those in the set of interest are counterfactually equal to zero almost surely, and comparing that variance to the unconditional variance of the process. More nuanced versions of this exercise include only considering variation over a certain range of spectral frequencies, or variation in forecast errors over a certain forecast horizon. Even in those more nuanced cases, however, the basic intuition is the same.

The problem with using news shocks to determine the importance of beliefs is that news shocks mix changes that are due to fundamentals and changes that are independently due to beliefs. This is because a news shock is an anticipated change in fundamentals. Expectations change at the time the news shock is realized, but then fundamentals change in the future when the anticipated change actually occurs. Of course, people's expectations may not always be fully borne out in the future fundamental, due to other unforeseen disturbances. Nevertheless, the anticipated shock is borne out on average, which is to say that news shocks are related to future fundamentals on average.

A stark way to see this point is to consider the importance of beliefs for driving fundamentals. Because fundamentals are purely exogenous, they are not driven by beliefs at all. However, in the simple news representation from Section (2.1), for example, news shocks can be an arbitrarily large part of fluctuations in the fundamental process $\{x_t\}$. Recall that

$$x_t = a_{0,t} + a_{1,t-1}, \qquad \begin{bmatrix} a_{0,t} \\ a_{1,t} \end{bmatrix} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \begin{bmatrix} \sigma_{a,0}^2 & 0 \\ 0 & \sigma_{a,1}^2 \end{bmatrix}\right).$$

Therefore, the fraction of the variation in $\{x_t\}$ due to news shocks, $\{a_{1,t}\}$ is given by:

$$\frac{\text{var}[x_t|a_{0,t}=0]}{\text{var}[x_t]} = \frac{\text{var}[a_{1,t}]}{\text{var}[x_t]} = \frac{\sigma_{a,1}^2}{\sigma_{a,0}^2 + \sigma_{a,1}^2}.$$

As $\sigma_{a,1}^2$ increases relative to $\sigma_{a,0}^2$, this fraction approaches one, in which case news shocks would explain all the variation in $\{x_t\}$.

To disentangle the importance of beliefs from fundamentals in models with news shocks, we need to use Theorem (1). Specifically, we can write down an observationally equivalent noise representation of the news model, and then use a variance decomposition to compute the share of variation attributable to noise shocks. Because these shocks are unrelated to fundamentals at all horizons, they capture precisely the independent contribution of beliefs.

Returning to the example from Section (2.1), we have already shown that in an observationally equivalent noise representation,

$$s_t = x_{t+1} + v_t, \qquad \begin{bmatrix} x_t \\ v_t \end{bmatrix} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}\right),$$

with $\sigma_x^2 \equiv \sigma_{a,0}^2 + \sigma_{a,1}^2$ and $\sigma_v^2 \equiv \sigma_x^2 \sigma_{a,0}^2 / \sigma_{a,1}^2$. According to this representation, the fraction of variation in $\{x_t\}$ due to noise shocks is:

$$\frac{\text{var}[x_t|x_t=0]}{\text{var}[x_t]} = 0,$$

which is the correct answer to the question of how much beliefs contribute to the fluctuations of fundamentals. This example illustrates the more general point that in order to determine the importance of beliefs, one should perform variance decompositions in terms of noise shocks rather than news shocks.

The fact that variance decompositions in terms of news shocks are not appropriate for determining the importance of beliefs has lead some researchers to conclude that there is a fundamental problem with using variance decompositions for that purpose. For example, Sims (2016) describes the problem of identifying the importance of beliefs (which both he and Barsky, Basu, and Lee (2015) call "pure news") as a fundamental limitation of the traditional variance decomposition:

> "The distinction between pure and realized news is important because one of the promises of the news-driven business cycle literature is to generate "boom-bust" cycles without any observable change in fundamentals ex-post. For understanding whether such "boom-bust" dynamics are quantitatively important it is critical to differentiate between effects of news shocks driven by actual news versus movements in endogenous variables caused by realized changes in fundamentals. A traditional variance decomposition does not make this distinction." (p. 2)

The point we would like to make is that the problem is not with the variance decompositions as such; rather, the problem is with the type of shock one considers. It is noise shocks, not news shocks, that are the appropriate shocks for isolating the independent contribution of beliefs. Once that distinction has been made, traditional variance decomposition methods can be employed as usual.

3.2 Mixing News and Noise Shocks

In some cases, researchers have constructed representations of fundamentals and beliefs that seem to include both news and noise shocks at the same time (e.g. Blanchard, L'Huillier, and Lorenzoni, 2013; Barsky and Sims, 2012). As a simple example, consider the following representation:

$$\begin{aligned} x_t &= \lambda_{t-1} + \eta_t \\ s_t &= \lambda_t + \xi_t, \end{aligned} \tag{4}$$

where

$$\begin{bmatrix} \eta_t \\ \lambda_t \\ \xi_t \end{bmatrix} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \begin{bmatrix} \sigma_\eta^2 & 0 & 0 \\ 0 & \sigma_\lambda^2 & 0 \\ 0 & 0 & \sigma_\xi^2 \end{bmatrix}\right).$$

At each date t, people observe $\{x_\tau, s_\tau\}$ for all integers $\tau \leq t$. The shock λ_t looks like a news shock because it affects people's beliefs at date t (through the signal s_t), but does not affect fundamentals until the following period. Similarly, the shock η_t looks like a surprise shock because it affects people's beliefs and the fundamental at the same time. Finally, the shock ξ_t looks like a noise shock because it affects people's beliefs but is not related to the fundamental at any date.

The problem with this type of representation, at least from the perspective of isolating the importance of beliefs, is that while ξ_t is unrelated to fundamentals, it does not fully capture the contribution of beliefs. This is because λ_t is not purely a function of fundamentals; it also captures changes that are due to beliefs. To see this, notice that in the limiting case $\xi_t = 0$, we have that $s_t = \lambda_t$ and this representation collapses to a news representation with $a_{0,t} \equiv \eta_t$ and $a_{1,t} \equiv \lambda_t$. We have already seen that in such a news representation, the news shock λ_t mixes changes that are due to fundamentals and beliefs. Moreover, we also know that such a news representation has an observationally equivalent noise representation with (non-zero) noise shocks that represent the independent contribution of beliefs (cf. Proposition (1)). Therefore $\xi_t = 0$ does not mean that beliefs do not have an independent role to play as a driver of fluctuations.

However, Theorem (1) implies that a representation of fundamentals and beliefs of the type in (4), which is neither news or noise representation, still has an observationally equivalent noise representation. The noise shocks in the noise representation properly capture the independent contribution of beliefs. The following proposition presents the mapping from one representation to the other.

Proposition 3. *The representation of fundamentals and beliefs in system* (4) *is observationally equivalent to the noise representation of fundamentals and beliefs in system* (2) *if and only if:*

$$\sigma_x^2 = \sigma_\lambda^2 + \sigma_\eta^2 \quad \textit{and} \quad \frac{\sigma_v^2}{\sigma_x^2} = \frac{\sigma_\lambda^2(\sigma_\eta^2 + \sigma_\xi^2) + \sigma_\eta^2\sigma_\xi^2}{\sigma_\lambda^4}.$$

To see why the process $\{\xi_t\}$ does not properly capture the importance of beliefs, consider the endogenous variable $\hat{x}_t = E_t[x_{t+1}]$. Under representation (4),

$$\hat{x}_t = \frac{\sigma_\lambda^2}{\sigma_\lambda^2 + \sigma_\xi^2}(\lambda_t + \xi_t),$$

so the contribution of the process $\{\xi_t\}$ is

$$\frac{\mathrm{var}[\hat{x}_t|\lambda_t = \eta_t = 0]}{\mathrm{var}[\hat{x}_t]} = \frac{\sigma_\xi^2}{\sigma_\lambda^2 + \sigma_\xi^2}.$$

On the other hand, in the observationally equivalent noise representation implied by Proposition (3),

$$\hat{x}_t = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_v^2}(x_{t+1} + v_t).$$

Therefore, the contribution of the process $\{v_t\}$ is

$$\frac{\mathrm{var}[\hat{x}_t|x_t = 0]}{\mathrm{var}[\hat{x}_t]} = \frac{\sigma_v^2}{\sigma_x^2 + \sigma_v^2} = \frac{\sigma_\lambda^2 \sigma_\eta^2}{\sigma_\lambda^2 + \sigma_\eta^2} + \frac{\sigma_\xi^2}{\sigma_\lambda^2 + \sigma_\xi^2}.$$

The second equality uses the parameteric restrictions from Proposition (3). Because the first term in this expression is positive, it follows that $\{\xi_t\}$ understates the importance of beliefs for explaining variations in $\{\hat{x}_t\}$. It is also easy to see how the importance of beliefs can be strictly positive even in the limiting case when $\sigma_\xi^2 = 0$.

3.3 Different Noise Representations

So far we have argued that it is possible to use a noise representation to determine the importance of beliefs as an independent driver of fluctuations. First, one can rewrite any representation of fundamentals and beliefs as a noise representation using the constructive procedure from Theorem (1). Then, one can use a variance decomposition to determine the share of variation in any endogenous variable that is attributable to noise shocks. And this share represents the contribution of beliefs independently of what can be attributed to fundamentals.

But is the variance decomposition in terms of noise shocks unique? As we pointed out in the discussion of Theorem (1), any representation of fundamentals and beliefs is compatible with infinitely many different noise representations. If different noise representations imply different things about the variance decomposition of endogenous variables in terms of beliefs and fundamentals, then the procedure outlined above will not deliver a unique answer.

Fortunately, it turns out that all observationally equivalent noise representations deliver the same answer regarding the importance of beliefs for any endogenous process. This means that when one is interested in the importance of beliefs relative to fundamentals, the fact that noise representations are not unique is not a problem.

Proposition 4. *In any noise representation of fundamentals and beliefs, the variance decomposition of any endogenous process in terms of noise and fundamentals is uniquely determined over any frequency range.*

An immediate corollary of this proposition is that the variance decomposition of people's errors in forecasting an endogenous process is also uniquely determined for any forecast horizon. This is because the forecast errors are themselves endogenous processes to which Proposition (4) applies.

Corollary 2. *In any noise representation of fundamentals and beliefs, the forecast error variance decomposition of any endogenous process in terms of noise and fundamentals is uniquely determined for any horizon, and over any frequency range.*

4 Quantitative Analysis

In this section, we use Theorem (1) and Proposition (4) to empirically quantify the independent contribution of beliefs in driving business-cycle fluctuations. Because several models of belief-driven fluctuations have already been constructed and estimated in the literature, we take something of a meta-analytic perspective. Specifically, we select three prominent theoretical models that have been estimated in the literature and compute the importance of beliefs implied by each of those models for different macroeconomic variables (e.g. output, investment, etc.). The three models are: the model of news shocks from Schmitt-Grohé and Uribe (2012), the model of news and animal spirits from Barsky and Sims (2012), and the model of noise shocks from Blanchard, L'Huillier, and Lorenzoni (2013).

These three models are different in several respects. First, they incorporate different physical environments, including differences in preferences, frictions and market structure. Second, the three models are estimated on different data and with different sample periods. Third, the authors make different assumptions about the information structure faced by agents. While agents in all three models observe current fundamentals and receive advance information about future fundamentals, Schmitt-Grohé and Uribe (2012) take a pure news perspective while the Barsky and Sims (2012) and Blanchard, L'Huillier, and Lorenzoni (2013) offer somewhat different perspectives on combining news and noise within a single model.

Perhaps not surprisingly given the scope of these differences, the authors above come to very different conclusions. Schmitt-Grohé and Uribe (2012) conclude that news shocks explain about one half of aggregate fluctuations, but do not take an explicit stance on the importance of independent fluctuations in beliefs. Barsky and Sims (2012) also conclude that news shocks are important, and that noise shocks explain essentially none of the variation in any variable. However, Blanchard, L'Huillier, and Lorenzoni (2013) conclude that noise shocks play a crucial role in business cycle dynamics, especially for consumption.

In principle, it is possible that these different conclusions are largely a result of the different "normalizations" the authors take with respect to noise shocks. Indeed, our analysis indicates that all authors have (implicitly or explicitly) underestimated the actual of role of independent shocks to beliefs in their estimated economies. For Schmitt-Grohé and Uribe (2012) and Barsky and Sims (2012), we find that the role of noise rises from zero to being small, but non-trivial, generally between 3 and 11 percent at the business cycle frequency. Surprisingly, even Blanchard, L'Huillier, and Lorenzoni (2013) significantly underestimate the role of pure noise in driving their economy, with beliefs about productivity driving endogenous variables more than productivity itself. While our results indicate that noise shocks are more important than previously reported, they do not fully explain the degree of disagreement regarding the independent contribution of beliefs.

4.1 Schmitt-Grohé and Uribe (2012)

The first model we consider comes from Schmitt-Grohé and Uribe (2012), and was constructed to determine the importance of news shocks for explaining aggregate fluctuations in output, consumption, investment, and employment. The main result of their paper is that news shocks account for about half of the predicted aggregate fluctuations in those four variables. As we have seen in the previous section, however, news shocks mix the contribution due to beliefs and fundamentals. As a result, exactly what this model implies about the importance of beliefs is still an unanswered question.

The model is a standard real business cycle model with six modifications: investment adjustment costs, variable capacity utilization with respect to the capital stock, decreasing returns to scale in production, one period internal habit formation in consumption, imperfect competition in labor markets, and period utility allowing for a low wealth effect on labor supply. Fundamentals comprise seven different independent processes, which capture exogenous variation in: stationary and non-stationary neutral productivity, stationary and non-stationary investment-specific productivity, government spending, wage markups, and preferences. The model is presented in more detail in Appendix (B.1).

Each of the seven exogenous fundamentals is described by the following law of motion, as we saw in Example (1) and reproduced here:

$$x_t = \rho_x x_{t-1} + \epsilon^a_{0,t} + \epsilon^a_{4,t-4} + \epsilon^a_{8,t-8}, \qquad \begin{bmatrix} \epsilon^a_{0,t} \\ \epsilon^a_{4,t} \\ \epsilon^a_{8,t} \end{bmatrix} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \begin{bmatrix} \sigma^2_{a,0} & 0 & 0 \\ 0 & \sigma^2_{a,4} & 0 \\ 0 & 0 & \sigma^2_{a,8} \end{bmatrix}\right). \tag{5}$$

where $0 < \rho_z < 1$. The model is estimated using likelihood-based methods on a sample of quarterly U.S. data from 1955:Q2-2006:Q4. The time series used for estimation are: real GDP, real consumption, real investment, real government expenditure, hours, utilization-adjusted total factor productivity, and the relative price of investment.

A variance decomposition shows that news shocks turn out to be very important. The first column of Table (1) shows the share of business-cycle variation in the level of four endogenous variables that is attributable to surprise shocks $\{\epsilon^a_{0,t}\}$, and the second column shows the share attributable to the news shocks $\{\epsilon^a_{4,t}\}$ and $\{\epsilon^a_{8,t}\}$ combined. We define business cycle frequencies as the components of the endogenous process with periods of 6 to 32 quarters, and we focus on variance decompositions over these frequencies to facilitate comparison across the different models in this section. Our results are consistent with the authors' original findings (see their Table V).

However, to determine the contribution of beliefs relative to fundamentals, it is necessary to construct a noise representation that is observationally equivalent to representation (5). To do this, we first present a result that applies to any model with news shocks of this type. It shows that it is always possible to find a particularly simple observationally equivalent noise representation — in particular, one that satisfies Assumption (1) — and explicitly presents the parameteric mapping from one representation to the other.

Proposition 5. *Any news representation in which each process $\{a_{i,t}\}$ is i.i.d. over time is observationally equivalent to a noise representation with $x_t \overset{iid}{\sim} \mathcal{N}(0, \sigma^2_x)$ and*

$$s_{i,t} = x_{t+i} + v_{i,t}, \quad v_{i,t} \overset{iid}{\sim} \mathcal{N}(0, \sigma^2_{v,i}),$$

where $v_{i,t} \perp x_\tau$ and $v_{i,t} \perp v_{j,\tau}$ for any $i \neq j \in \mathcal{I}$ and $t, \tau \in \mathbb{Z}$, if and only if

$$\sigma^2_x = \sum_{i \in \mathcal{I}} \sigma^2_{a,i} \quad \textit{and} \quad \sigma^2_{v,i} = \frac{1}{\sigma^2_{a,i}} \left(\sum_{j<i} \sigma^2_{a,j}\right)\left(\sum_{j \leq i} \sigma^2_{a,j}\right) \quad \textit{for all } i \in \mathcal{I}.$$

Applying this proposition to the model of Schmitt-Grohé and Uribe (2012) requires one small step, which is that Proposition (5) is stated for i.i.d. fundamentals, but the fundamentals in system (5) are not i.i.d. However, because $0 < \rho_x < 1$, observing the current and past history of x_t is equivalent to observing the current and past history of the composite disturbance $\epsilon_t^x \equiv \epsilon_{0,t}^a + \epsilon_{4,t-4}^a + \epsilon_{8,t-8}^a$, which is i.i.d. because each of the news shocks are independent of one another. Therefore, it is possible to treat $\{\epsilon_t^x\}$ as the fundamental process. By doing so, we arrive at the following corollary.

Corollary 3. *The representation of fundamentals and beliefs in system* (5) *is observationally equivalent to the noise representation*

$$\begin{aligned} x_t &= \rho_x x_{t-1} + \epsilon_t^x \\ s_{4,t} &= \epsilon_{t+4}^x + v_{4,t} \\ s_{8,t} &= \epsilon_{t+8}^x + v_{8,t}, \end{aligned}$$

with the convention that $s_{0,t} \equiv x_t$*, and where*

$$\begin{bmatrix} \epsilon_t^x \\ v_{4,t} \\ v_{8,t} \end{bmatrix} \overset{iid}{\sim} \mathcal{N}\left(0, \begin{bmatrix} \sigma_x^2 & 0 & 0 \\ 0 & \sigma_{v,4}^2 & 0 \\ 0 & 0 & \sigma_{v,8}^2 \end{bmatrix}\right),$$

if and only if

$$\begin{aligned} \sigma_x^2 &= \sigma_{a,0}^2 + \sigma_{a,4}^2 + \sigma_{a,8}^2 \\ \sigma_{v,4}^2 &= \frac{1}{\sigma_{a,4}^2}\sigma_{a,0}^2(\sigma_{a,0}^2 + \sigma_{a,4}^2) \\ \sigma_{v,8}^2 &= \frac{1}{\sigma_{a,8}^2}(\sigma_{a,0}^2 + \sigma_{a,4}^2)(\sigma_{a,0}^2 + \sigma_{a,4}^2 + \sigma_{a,8}^2). \end{aligned}$$

We can use the noise representation in Corollary (3) with the same parameter estimates as before, and re-compute the variance decomposition of the seven observable variables in terms of fundamental shocks and noise shocks. This decomposition is unique by Proposition (4). There is no need to re-estimate the model because observational equivalence implies that the likelihood function is the same under both representations. The third column of Table (1) shows the share of variation attributable to fundamental shocks $\{\epsilon_t^x\}$, and the fourth column shows the share attributable to the noise shocks $\{v_{4,t}\}$ and $\{v_{8,t}\}$ combined.

Variable	Surprise	News	Fundamental	Noise
Output	57	43	95	5
Consumption	50	50	95	5
Investment	55	45	89	11
Hours	15	85	97	3

Table 1: Variance decomposition (%) in the model of Schmitt-Grohé and Uribe (2012) over business cycle frequencies of 6 to 32 quarters. All variables are in levels. Estimated model parameters are set to their posterior median values.

The main result is that nearly all of the variation in output, consumption, investment, and hours is due to changes in fundamentals. In terms of differences across the endogenous variables, it is interesting that real investment growth is affected the least by news shocks, but it is affected the most by noise shocks. At the same time, hours worked is affected the most by news shocks and the least by noise shocks. But based on the fact that 90% or more of the variation in every series is attributable to fundamental changes, we conclude that beliefs are not an important independent source of fluctuations through the lens of this model.

4.2 Barsky and Sims (2012)

The second model comes from Barsky and Sims (2012). It was constructed to determine whether measures of consumer confidence change in ways that are related to macroeconomic aggregates because of noise (i.e. "animal spirits") or news. The main result of the paper is that changes in consumer confidence are mostly driven by news and not noise. They also find that that noise shocks account for negligible shares of the variation in forecast errors of consumption and output, while news shocks account for over half of the variation in long-horizon forecast errors. However, as we saw in Section (3.2), including both news and noise shocks in the same model can be misleading when it comes to isolating the importance of beliefs.

The model is a standard dynamic, stochastic, general equilibrium (DSGE) model with real and nominal frictions: one period internal habit formation in consumption, capital adjustment costs (as opposed to investment adjustment costs, according to which costs are expressed as a function of the growth rate of investment rather than the level of investment relative to the existing capital stock), and monopolistic price setting with time-dependent price rigidity. Fundamentals comprise three different independent processes, which capture exogenous variation in: non-stationary neutral productivity, government spending, and monetary policy. The model is presented in more detail in Appendix (B.2).

People only receive advance information about productivity, and not about the other two fundamentals. So it is only beliefs about productivity that can play an independent role in driving fluctuations. Letting x_t denote the growth rate of productivity (in deviations from its mean), and using our notation from Section (3.2), the process $\{x_t\}$ is assumed to follow a law of motion of the form:

$$\begin{aligned} x_t &= \lambda_{t-1} + \eta_t \\ \lambda_t &= \rho\lambda_{t-1} + \epsilon_t^\lambda \\ s_t &= \lambda_t + \xi_t, \end{aligned} \tag{6}$$

where $0 < \rho < 1$ and

$$\begin{bmatrix} \eta_t \\ \epsilon_t^\lambda \\ \xi_t \end{bmatrix} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \begin{bmatrix} \sigma_\eta^2 & 0 & 0 \\ 0 & \sigma_\lambda^2 & 0 \\ 0 & 0 & \sigma_\xi^2 \end{bmatrix}\right).$$

Barsky and Sims (2012) refer to ϵ_t^λ as a news shock, η_t as a surprise shock, and ξ_t as a noise (animal spirits) shock. However, these definitions of news, surprise, and noise shocks are not consistent with the definitions in our paper. To avoid any confusion we will use asterisks to indicate the terminology of Barsky and Sims (2012). So ϵ_t^λ is a news* shock, η_t is a surprise* shock, and ξ_t is a noise* shock.

The model is estimated by minimizing the distance between impulse responses generated from simulations of the model and those from estimated structural vector auto-regressions. The vector auto-regressions are estimated on quarterly U.S. data from 1960:Q1-2008:Q4. The time series used to estimate the vector auto-regression are: real GDP, real consumption, CPI inflation, a measure of the real interest rate, and a measure of consumer confidence from the Michigan Survey of Consumers (E5Y).

A variance decomposition shows that news* shocks are much more important than noise* shocks. The first column of Table (2) shows the share of business-cycle variation in the level of four endogenous variables that is attributable to surprise* shocks $\{\eta_t\}$, the second column shows the share attributable to news* shocks $\{\epsilon_t^\lambda\}$, and the third column shows the share attributable to noise* shocks $\{\xi_t\}$. Due to the presence of government spending and monetary policy shocks, the rows do not sum to 100%; the residual represents the combined contribution of these two additional fundamental shocks. These results are consistent with the authors' original findings, which are stated in terms of the variance decompositions of forecast errors over different horizons, but across all frequency ranges (see their Table 3).

However, to properly isolate the independent contributions of beliefs, it is necessary to construct a noise representation that is observationally equivalent to representation (6). The following proposition presents one such noise representation.

Proposition 6. *The representation of fundamentals and beliefs in system* (6) *is observationally equivalent to the noise representation*

$$\begin{aligned} x_t &= \gamma_0 m_t + \gamma_1 m_{t-1} + \gamma_0 m_{t-2} \\ m_t &= \phi_1 m_{t-1} + \phi_2 m_{t-2} + \epsilon_t^x \\ s_t &= m_t + v_t \\ v_t &= \delta v_{t-1} + \epsilon_t^v - \beta \epsilon_{t-1}^v, \end{aligned}$$

with the convention that $s_{0,t} \equiv x_t$, *and where*

$$\begin{bmatrix} \epsilon_t^x \\ \epsilon_t^v \end{bmatrix} \overset{iid}{\sim} \mathcal{N}\left(0, \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}\right),$$

if and only if δ *is equal to the root of the polynomial*

$$\mathcal{P}(z) = \rho z^2 - \left(1 + \rho^2 + \frac{\sigma_\lambda^2}{\sigma_\eta^2}\right) z + \rho$$

that lies inside the unit circle, β *is the root of the polynomial*

$$\mathcal{P}(z) = \rho z^2 - \left(1 + \rho^2 + \frac{\sigma_\lambda^2(\sigma_\lambda^2 + \sigma_\xi^2)}{\sigma_\eta^2 \sigma_\xi^2}\right) z + \rho$$

that lies inside the unit circle, and

$$\begin{aligned} \gamma_0 &= -\rho \frac{\sigma_\eta^2}{\sigma_\lambda^2} & \gamma_1 &= -\gamma_0 \left(\frac{1+\delta^2}{\delta}\right) \\ \phi_1 &= \rho + \delta & \phi_2 &= -\rho\delta \\ \sigma_x^2 &= \frac{\delta \sigma_\lambda^4}{\rho \sigma_\eta^2} & \sigma_v^2 &= \frac{\delta}{\beta}\sigma_\xi^2. \end{aligned}$$

Using the noise representation in this proposition, we can re-compute the variance decomposition of the endogenous processes in terms of fundamental shocks and noise shocks. The fourth column of Table (2) shows the share of variation attributable to fundamental productivity shocks $\{\epsilon_t^x\}$, and the fifth column shows the share attributable to productivity noise shocks $\{\epsilon_t^v\}$. Again, the rows do not sum to 100% due to the presence of government spending and monetary policy shocks. Conceptually, the contribution of these shocks should also be included under the heading of fundamental shocks, but for comparison with the first three columns, we only include fundamental productivity shocks in the fourth column.

Variable	Surprise*	News*	Noise*	Fundamental	Noise
Output	0.53	0.37	0	89	1
Consumption	0.62	0.35	1	89	9
Investment	0.40	0.43	1	80	4
Hours	0.64	0.14	0	75	3

Table 2: Variance decomposition (%) in the model of Barsky and Sims (2012) over business cycle frequencies of 6 to 32 quarters. All variables are in levels, and estimated parameters are set to their point-estimated values. The rows do not sum to 100% because of other non-technology fundamental processes. Asterisks refer to the authors' terminology.

As in the model of Schmitt-Grohé and Uribe (2012), we find that nearly all of the variation in output, consumption, investment, and hours is due to changes in fundamentals. The contribution of noise shocks is larger than the contribution of noise* shocks, for all variables. However, the bulk of the contribution of news* shocks turns out to be due to fundamentals rather than noise.

4.3 Blanchard, L'Huillier, and Lorenzoni (2013)

The third model we consider comes from Blanchard, L'Huillier, and Lorenzoni (2013), and was constructed "to separate fluctuations due to changes in fundamentals (news) from those due to temporary errors in agents' estimates (noise)."[2] The main quantitative result of their paper is that noise shocks explain a sizable fraction of short-run

[2]This quotation is taken from the article's abstract, which can be found on the AEA's website: `https://www.aeaweb.org/articles?id=10.1257/aer.103.7.3045`.

consumption fluctuations. However, it turns out that what the authors call “noise” shocks do not fully isolate fluctuations due to temporary errors in agents’ estimates. As a result, it is still an unanswered question what exactly this model implies about the importance of beliefs.

The model is a standard DSGE model with real and nominal frictions: one-period internal habit formation in consumption, investment adjustment costs, variable capacity utilization with respect to capital, and monopolistic price and wage setting with time-dependent price rigidities. Fundamentals comprise six different independent processes, which capture exogenous variation in: non-stationary neutral productivity, stationary investment-specific productivity, government spending, wage markups, final good price markups, and monetary policy. For more details, see Appendix (B.3).

People only receive advance information about productivity, and not about the other five fundamentals. So it is only beliefs about productivity that can play an independent role in driving fluctuations. Letting x_t denote the growth rate of productivity (in deviations from its mean), and using our notation from Section (3.2), the process $\{x_t\}$ is assumed to follow a law of motion of the form:[3]

$$\begin{aligned} x_t &= \lambda_t + \eta_t - \eta_{t-1} \\ \lambda_t &= \rho\lambda_{t-1} + \epsilon_t^\lambda \\ \eta_t &= \rho\eta_{t-1} + \epsilon_t^\eta \\ s_t &= -\eta_t + \xi_t, \end{aligned} \tag{7}$$

where $0 < \rho < 1$ and

$$\begin{bmatrix} \epsilon_t^\eta \\ \epsilon_t^\lambda \\ \xi_t \end{bmatrix} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \begin{bmatrix} \sigma_\eta^2 & 0 & 0 \\ 0 & \sigma_\lambda^2 & 0 \\ 0 & 0 & \sigma_\xi^2 \end{bmatrix}\right).$$

Blanchard, L’Huillier, and Lorenzoni (2013) refer to ϵ_t^λ as a permanent shock, ϵ_t^η as a transitory shock, and ξ_t as a noise shock. Taken together, they refer to ϵ_t^λ and ϵ_t^η as

[3]The authors actually present the information structure in terms of non-stationary processes. But because we are working in $\mathcal{L}^2$, we transform that system into stationary form. The variable λ_t represents the growth rate of the permanent component of productivity, and η_t is the transitory component. The signal s_t in system (7) is equal to their signal minus the natural logarithm of the level of productivity.

news shocks. Again, because these definitions are not consistent with the ones in our paper, we will use asterisks to indicate the authors' terminology in contrast to ours.

The model is estimated using likelihood-based methods on a sample of quarterly U.S. data from 1954:Q3-2011:Q1. The time series used for estimation are: real GDP, real consumption, real investment, employment, the federal funds rate, inflation as measured by the implicit GDP deflator, and wages.

A variance decomposition reveals that noise* shocks are important, especially for consumption. The first column of Table (3) shows the share of business-cycle variation in the level of output, consumption, investment, and hours that is attributable to news* shocks, $\{\epsilon_t^\lambda\}$ and $\{\epsilon_t^\eta\}$, and the second column shows the share attributable to noise* shocks $\{\xi_t\}$. Due to the presence of the other five fundamental shocks, the rows do not sum to 100%; the residual represents the combined contribution of these additional fundamental shocks. These results are consistent with the authors' original findings, which are stated in terms of the variance decompositions of forecast errors over different horizons (see their Table 6).

However, to properly isolate the independent contribution of beliefs, it is necessary to construct a noise representation that is observationally equivalent to representation (7). The following proposition presents one such noise representation.

Proposition 7. *The representation of fundamentals and beliefs in system* (7) *is observationally equivalent to the noise representation*

$$\begin{aligned} x_t &= -\phi m_{t+1} + m_t \\ m_t &= \phi m_{t-1} + \epsilon_t^x \\ s_t &= -\phi(m_t - m_{t-1}) + v_t \\ v_t &= \phi v_{t-1} + \epsilon_t^v - (\beta_1 + \beta_2)\epsilon_{t-1}^v + \beta_1\beta_2\epsilon_{t-2}^v, \end{aligned}$$

with the convention that $s_{0,t} \equiv x_t$, *and where*

$$\begin{bmatrix} \epsilon_t^x \\ \epsilon_t^v \end{bmatrix} \overset{iid}{\sim} \mathcal{N}\left(0, \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_v^2 \end{bmatrix}\right),$$

if and only if (β_1, β_2) *are equal to the two roots of the polynomial*

$$\mathcal{P}(z) = \rho^2 z^4 - 2\rho(1+\rho^2)z^3 + \left(1 + \rho^4 + 4\rho^2 + \rho\frac{\sigma_\lambda^2}{\sigma_\xi^2}\right) z - 2\rho(1+\rho^2)z + \rho^2$$

that lie inside the unit circle, $\phi = \rho$, $\sigma_x^2 = \sigma_\eta^2/\rho$, *and* $\sigma_v^2 = \rho^2\sigma_\xi^2/(\beta_1\beta_2)$.

Using the noise representation in this proposition, we can re-compute the variance decomposition of the endogenous processes in terms of fundamental shocks and noise shocks. The fourth column of Table (3) shows the share of variation attributable to fundamental productivity shocks $\{\epsilon_t^x\}$, and the fifth column shows the share attributable to productivity noise shocks $\{\epsilon_t^v\}$. Again, the rows do not sum to 100% due to the presence of fundamental processes other than productivity.

Variable	News*	Noise*	Fundamental	Noise
Output	34	22	26	29
Consumption	40	44	27	57
Investment	6	3	4	5
Hours	17	29	7	39

Table 3: Variance decomposition (%) in the model of Blanchard, L'Huillier, and Lorenzoni (2013) over business cycle frequencies of 6 to 32 quarters. All variables are in levels, and estimated parameters are set to their posterior median values. The rows do not sum to 100% because of other non-technology fundamental processes.

In contrast to both the Schmitt-Grohé and Uribe (2012) and Barsky and Sims (2012) models, we find that a sizable fraction of the variation in output, consumption, and hours worked can be attributed to noise shocks. For example, nearly 60% of the variation in consumption is due to noise shocks. This is over 10% larger than the share Blanchard, L'Huillier, and Lorenzoni (2013) originally attributed to independent fluctuations in beliefs. A result of similar magnitude is true for output and hours worked.

5 Structural VAR Analysis

One issue that several researchers have emphasized as a difference between news and noise representations involves the applicability of structural VAR analysis. A common view is that news shocks can generally be recovered from data using structural VAR analysis, while noise shocks generally cannot. The typical reason given is that news representations are invertible (the shocks can be expressed as a function of current and past observables), while noise representations are not.

In this section, we prove that the observational equivalence of news and noise representations implies that the shocks in any noise representation can be recovered from observables up to an orthogonal transformation. Based on this result, we argue that structural VAR analysis is equally appropriate for recovering the underlying shocks in either a news or noise representation. We also explore one particular orthogonalization that is related to the popular "news-reversal" thought experiment that the existing literature has used to describe boom-bust episodes in models with news shocks.

5.1 Recovering Shocks and Invertibility

When can the underlying shocks in news and noise representations be recovered from the data? That is, given data on fundamentals and beliefs, when are the underlying shocks in these representations uniquely determined? The uniqueness of the news representation according to Theorem (1) implies that in any news representations of fundamentals and beliefs, each underlying shock is uniquely determined.

By contrast, each underlying shock in a noise representation is not uniquely determined. Proposition (4) and its associated Corollary (2) establish the uniqueness of variance decompositions computed in terms of fundamentals and noise, but they do not imply that it is possible to separately recover each individual shock. However, we can prove the following result, which says that the shocks are determined up to an orthogonal transformation.

Proposition 8. *In any noise representation of fundamentals and beliefs, the space spanned by the underlying shocks at each date is uniquely determined.*

An important concept in discussions regarding the applicability of structural VAR analysis is that of invertibility. This has to do with whether or not it is possible to express one collection of stochastic processes as a linear combination of the current and past history of another collection of stochastic processes.

Definition 4. A collection of stochastic processes $\{y_{i,t}\}$, $i \in \mathcal{I}_y \subseteq \mathbb{Z}_+$ is "invertible" with respect to the collection of shock processes $\{\epsilon_{i,t}\}$, $i \in \mathcal{I}_\epsilon \subseteq \mathbb{Z}_+$, if

$$\mathcal{H}_t(\epsilon) = \mathcal{H}_t(y) \quad \text{for all } t \in \mathbb{Z}.$$

Based on this definition, we refer to a representation of the collection of $\{y_{i,t}\}$ as an "invertible representation" if $\{y_{i,t}\}$ is invertible with respect to all the underlying shocks in that representation.[4] Also, note that we use the term "shock process" to refer to a process that is uncorrelated over time.

Our second theoretical result of this subsection characterizes news and noise representations in terms of invertibility.

Proposition 9. *Any news representation of fundamentals and beliefs is invertible, but any noise representation is not invertible.*

This result is a generalization of the one that Blanchard, L'Huillier, and Lorenzoni (2013) prove in the context of a simple model of consumption determination, and the basic intuition is the same. If noise representations were invertible, then people would be able to distinguish the informative parts of their signals from the noise. By rationality, noise shocks could never affect people's beliefs. But then it would not be possible to recover those shocks from the current and past history of people's beliefs.

Any collection of observable processes is invertible with respect to infinitely many different collections of underlying shocks. However, these shocks have the important property that they are all related by an orthogonal transformation. We state this in the following proposition, which has a well-known finite-dimensional counterpart (e.g. Rozanov, 1967, p. 57):

Proposition 10. *If a collection of stochastic processes is invertible with respect to two different collections of shock processes, then the space spanned by those shocks at each date is the same.*

5.2 Using Structural VAR Analysis

For many researchers, Proposition (9) settles the question of whether structural VAR analysis can be used to recover shocks in news and noise representations. For example, the central methodological argument of Blanchard, L'Huillier, and Lorenzoni (2013) is that structural VAR analysis is not applicable for recovering noise shocks due to non-invertibility:

[4] What we call invertibility is sometimes called "fundamentalness" (cf. Rozanov, 1967, ch. 2).

> "[In situations with] a partially informative signal, the reduced-form VAR representation is non-invertible and a structural VAR approach cannot be used." (p. 3051)

However, in this subsection we argue that the applicability of structural VAR analysis should not be understood solely in terms of invertibility.

Structural VAR analysis has two steps: a VAR step and a structural step. The VAR step defines a "reduced-form" representation of the observables with residuals that come from a projection of observables on their past history. The structural step uniquely determines a collection of "economic shocks" from the reduced-form representation by using theoretical restrictions to pin down a single orthogonal transformation (an orthogonal matrix in the finite-dimensional case). In other words, we can say that structural VAR analysis is applicable whenever knowledge of the reduced-form representation is at most one orthogonal transformation away from knowledge of the economic shocks.

To be more precise, we can define a reduced-form representation of fundamentals and beliefs in the following way.

Definition 5. In a "reduced-form" representation of fundamentals and beliefs,

$$\hat{x}_{i,t} = \tilde{x}_{i,t-1} + \tilde{\epsilon}_{i,t} \quad \text{for all } i, t \in \mathbb{Z},$$

where $\tilde{x}_{i,t-1} \in \mathcal{H}_{t-1}(\hat{x})$ and $\tilde{\epsilon}_{i,t} \perp \mathcal{H}_{t-1}(\hat{x})$.

The first step of structural VAR analysis is to treat this representation as known. The second step is to determine whether the reduced-form shocks $\{\tilde{\epsilon}_{i,t}\}$ uniquely determine the space spanned by the underlying shocks in either a news or noise representation.

For a news representation, the answer is yes. Note that, by construction, $\{\hat{x}_{i,t}\}$ is invertible with respect to $\{\tilde{\epsilon}_{i,t}\}$. By Propositions (9) and (10), it follows that the space spanned by the reduced-form shocks is equal to the space spanned by the shocks in any news representation are the same at each date. Therefore, the reduced-form shocks uniquely determine the space spanned by the shocks in any news representation.

Interestingly, for a noise representation, the answer is also yes. To see why, note that Proposition (8) implies that the space spanned by the shocks in any noise representation is uniquely determined by $\{\hat{x}_{i,t}\}$. But then that space must also be also uniquely determined by the reduced-form residuals, because $\mathcal{H}_t(\tilde{\epsilon}) = \mathcal{H}_t(\hat{x})$ for all $t \in \mathbb{Z}$ by invertibility.

But isn't it true that in a news representation, the shocks themselves are uniquely determined (by Theorem (1)), while in a noise representation only the space spanned by the shocks is uniquely determined (by Theorem (1) and Proposition (8))? Yes, but that is only because the orthogonal transformation linking the reduced-form representation to the news representation is already embedded in the definition of a news representation. Of course, we could just have easily have appended one particular orthogonal transformation to the definition of a noise representation in the first place. Therefore, we can conclude that to recover the underlying shocks in both news and noise representations from the reduced-form representation, the same theoretical input is required: one orthogonal transformation.

One natural set of restrictions is that noise shocks are orthogonal, and that noise shock $i \in \mathcal{I}$ has a unit impact response on the forecast $\hat{x}_{i,t}$ but zero impact response on forecasts $\hat{x}_{j,t}$ for $j < i$. These restrictions impose a familiar lower-triangular structure on the shocks in a noise representation. They amount to a recursive causal ordering of the noise shocks in terms of the observable collection of forecasts $\{\hat{x}_{i,t}\}$.

Assumption 1. *In any noise representation of fundamentals and beliefs, the following conditions are satisfied:*

(a) $\epsilon^v_{i,t} \perp \epsilon^v_{j,t}$ *for all* $i \neq j \in \mathcal{I}$,

(b) $\langle \hat{x}_{i,t}, \epsilon^v_{i,t} \rangle / \|\epsilon^v_{i,t}\|^2 = 1$ *for all* $i \in \mathcal{I}$, *and*

(c) $\langle \hat{x}_{j,t}, \epsilon^v_{i,t} \rangle = 0$ *for all* $j < i \in \mathcal{I}$.

Proposition 11. *In any noise representation of fundamentals and beliefs that satisfies Assumption* (1), *the underlying shocks are uniquely determined.*

Of course, an immediate corollary of this proposition is that under Assumption (1), the impulse response function and variance decomposition of any endogenous process with respect to each shock in a noise representation are also uniquely determined.

5.3 Offsetting News Shocks

The orthogonal transformation implicit in Assumption (1) is also related to a popular thought experiment in the news-shock literature, which some researchers have used to try to isolate the effects of a change in beliefs that does not correspond to any change in fundamentals (e.g. Christiano et al. (2010) Section 4.2, Schmitt-Grohé and Uribe (2012) Section 4.2, Barsky, Basu, and Lee (2015) Section IV.A, or Sims (2016) Section 3.3). This experiment involves computing the impulse responses of endogenous variables in response to particular combinations of offsetting news and surprise shocks. The following description comes from Christiano et al. (2010), with slight modification to match the notation in this paper:

> "In the first period, a signal, $\epsilon^a_{n,t} > 0$ arrives, which creates the expectation that z_t n periods later will jump. However, that expectation is ultimately disappointed, because $\epsilon^a_{0,t} = -\epsilon^a_{n,t}$. Thus, in fact nothing real ever happens. The dynamics of the economy are completely driven by an optimistic expectation about future fundamentals, an expectation that is never realized." (p. 116)

It turns out that in models with i.i.d. news shocks, the noise shocks in any observationally equivalent representation generate exactly the sort of offsetting news shocks envisioned by this thought experiment. To see this, recall the simple news representation in system (1) from Section (2.1). In a noise representation of the type in system (2), which satisfies Assumption (1), observational equivalence implies that the news shocks are linear combinations of the Wold innovations:

$$a_{0,t} = w_{0,t} \quad \text{and} \quad a_{1,t} = \kappa w_{1,t},$$

where $\kappa = \sigma_x^2/(\sigma_x^2+\sigma_v^2)$ is a Kalman gain parameter controlling how much people trust the noisy signal. We can in turn express the Wold innovations in terms of observable variables, $w_{0,t} = x_t - \kappa s_t$ and $w_{1,t} = s_t$, so that the news shocks can be written in the form

$$a_{0,t} = x_t - \kappa s_{t-1} \quad \text{and} \quad a_{1,t} = \kappa s_t.$$

Therefore, the responses of the news shocks $a_{0,t}$ and $a_{1,t}$ in period h to a unit impulse in the noise shock v_t at date t are given by:

$$IR(a_{1,t+h}) = 1_{\{h=0\}}\kappa \quad \text{and} \quad IR(a_{0,t+h}) = -1_{\{h=1\}}\kappa.$$

This means that a positive noise shock is equivalent to a positive news shock today followed by an exactly offsetting surprise shock one period in the future.

While this discussion shows that it may be possible to find particular linear combinations of news shocks that mimic a noise shock, there are a number of advantages to working directly with noise shocks. First, we can think about how likely these situations arise, since we have an explicit probability distribution for the noise shocks: for example, how big is a "one standard deviation impulse?" Second, we can ask how important these types of situations are in the data sample overall; that is, we can do a proper variance decomposition. Third, the fact that noise representations are generally not unique helps us to remember that the dynamic response of the economy to noise shocks is also not unique. With news shocks at multiple different horizons, there are many ways people's expectations can be subsequently reversed, and the reversal may not occur only in the final period.

6 Conclusion

Models with news and noise are more intimately related than the literature has acknowledged. In fact, as we have argued here, there is a precise sense in which they are identical. The missing link is the observation that they are really just two different ways of describing the joint dynamics of exogenous economic fundamentals and people's beliefs about them. This observation is formalized by Theorem (1).

Far from being a purely negative result, the observational equivalence between news and noise representations serves an important positive purpose. Namely, it provides a way to determine the importance of beliefs as an independent cause of fluctuations. A number of prominent studies have constructed models to understand how beliefs can drive fluctuations. However, none of them has fully isolated the contribution of beliefs that is independent of the contribution due to fundamentals. This is because what these studies refer to as "news" shocks actually mix the fluctuations due fundamentals and beliefs.

In order to disentangle beliefs from fundamentals, it is necessary to first derive a noise representation of the model, and then perform variance decompositions in terms of noise shocks. These decompositions are always unique by Proposition (4). We also state a set of sufficient conditions for uniquely recovering the impulse response function of any endogenous process with respect to noise shocks. The uniqueness result that obtains under those conditions is presented in Proposition (11).

We then apply our results to three quantitative models of the U.S. economy, from Schmitt-Grohé and Uribe (2012), Barsky and Sims (2012), and Blanchard, L'Huillier, and Lorenzoni (2013). We show that these studies have all understated the importance of pure shocks to beliefs that is implied by their models. It turns that each of the models we considered attributes a very small importance of noise shocks for investment, but substantial disagreement remains regarding the importance of beliefs for other variables.

The observational equivalence between news and noise shocks also implies that structural VAR analysis is equally applicable to models with news or noise. The shocks in both news and noise representations can be recovered from reduced-form residuals with one orthogonal transformation. This means that, just as in the context of the dynamic general equilibrium models from Section (4), it is also possible to determine what estimated news-shock structural VARs of the type in Cochrane (1994) or Beaudry and Portier (2006) imply about the importance of noise shocks.

We conclude by pointing out that our results are also relevant for fields outside the realm of business-cycle macroeconomics. In the asset pricing literature, for example, a large body of research starting with Bansal and Yaron (2004) has argued that the following news representation of beliefs and fundamentals is useful for explaining many asset-pricing phenomena:

$$\begin{aligned} x_t &= a_{0,t} + a_{1,t-1} \\ a_{0,t} &= \epsilon^a_{0,t} \\ a_{1,t} &= \rho a_{1,t-1} + \epsilon^a_{1,t}, \end{aligned}$$

where

$$\begin{bmatrix} \epsilon^a_{0,t} \\ \epsilon^a_{1,t} \end{bmatrix} \overset{\text{iid}}{\sim} \mathcal{N}\left(0, \begin{bmatrix} \sigma^2_{a,0} & 0 \\ 0 & \sigma^2_{a,1} \end{bmatrix}\right).$$

The fundamental process $\{x_t\}$ captures exogenous variation in payoffs (e.g. dividend growth) and $\{a_{1,t}\}$ represents the "long-run risk" in those payoffs. How much of the implied variation in asset prices is attributable to independent variation in people's beliefs? To answer this question, we can follow the same approach as in Section (4). We leave this and other applications for future work.

References

Anderson, B. D. and J. B. Moore (1979). *Optimal Filtering*. Englewood Cliffs, NJ: Prentice-Hall.

Bansal, R. and A. Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *The Journal of Finance 59*(4), 1481–1509.

Barsky, R. B., S. Basu, and K. Lee (2015, July). Whither News Shocks? In *NBER Macroeconomics Annual*, Volume 29, pp. 225–264. Chicago: University of Chicago Press.

Barsky, R. B. and E. R. Sims (2012). Information, animal spirits, and the meaning of innovations in consumer confidence. *American Economic Review 102*(4), 1343–1377.

Beaudry, P. and F. Portier (2006). Stock prices, news, and economic fluctuations. *American Economic Review 96*(4), 1293–1307.

Beaudry, P. and F. Portier (2014, December). News-Driven Business Cycles: Insights and Challenges. *Journal of Economic Literature 52*(4), 993–1074.

Blanchard, O. J., J.-P. L'Huillier, and G. Lorenzoni (2013). News, noise, and fluctuations: An empirical exploration. *American Economic Review 103*(7), 3045–3070.

Christiano, L. J., C. Ilut, R. Motto, and M. Rostagno (2010). Monetary policy and stock market booms. *Proceedings of the Jackson Hole Economic Policy Symposium 34*(1), 85–145.

Cochrane, J. H. (1994, December). Shocks. *Carnegie-Rochester Conference Series on Public Policy 41*(1), 295–364.

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

Luenberger, D. G. (1969). *Optimization by Vector Space Methods*. New York, NY: John Wiley and Sons.

Pigou, A. C. (1927). *Industrial Fluctuations*. London: Macmillan.

Rozanov, Y. A. (1967). *Stationary Random Processes.* San Francisco, CA: Holden-Day.

Schmitt-Grohé, S. and M. Uribe (2012). What's news in business cycles. *Econometrica 80*(6), 2733–2764.

Sims, E. (2016, May). What's news in news? a cautionary note on using a variance decomposition to assess the quantitative importance of news shocks. Working paper, University of Notre Dame.

Whittle, P. (1983). *Prediction and Regulation by Linear Least-Square Methods.* Minneapolis, MN: University of Minnesota Press.

A Proofs

Proof of Proposition (1). Let $\hat{x}_t \equiv E_t[x_{t+1}]$ denote people's expectations of the fundamental at date $t+1$ given their information up through date t. The observable processes are $\{x_t\}$ and $\{\hat{x}_t\}$. Expectations at horizons greater than one are spanned by these two processes.

The two representations are observationally equivalent if and only if the covariance generating function (c.g.f.) of the data vector $d_t \equiv (x_t, \hat{x}_t)'$ is the same under either representation. Let $g_d(z)$ denote the c.g.f. of d_t, where z is a number in the complex plane. Then we can equate the c.g.f.'s implied by each representation:

$$g_d(z) = \underbrace{\begin{bmatrix} \sigma_{a,0}^2 + \sigma_{a,1}^2 & \sigma_{a,1}^2 z \\ \sigma_{a,1}^2 z^{-1} & \sigma_{a,1}^2 \end{bmatrix}}_{\text{news}} = \underbrace{\begin{bmatrix} \sigma_x^2 & \left(\frac{\sigma_x^4}{\sigma_x^2+\sigma_v^2}\right) z \\ \left(\frac{\sigma_x^4}{\sigma_x^2+\sigma_v^2}\right) z^{-1} & \left(\frac{\sigma_x^4}{\sigma_x^2+\sigma_v^2}\right) \end{bmatrix}}_{\text{noise}}.$$

This equality holds if and only if the relations in Proposition (1) are satisfied. □

Proof of Corollary (1). The proof is stated in the text. □

Proof of Theorem (1). To prove the first part of the theorem, note that because $\mathcal{H}_{t-1}(\hat{x}) \subset \mathcal{H}_t(\hat{x})$ for all $t \in \mathbb{Z}$, it is possible to decompose $\mathcal{H}_t(\hat{x})$ into an orthogonal family of subspaces

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

where $\mathcal{D}_t \equiv \mathcal{H}_t(\hat{x}) \ominus \mathcal{H}_{t-1}(\hat{x})$ (cf. Rozanov, 1967, ch. 2). This means that $x_t \in \mathcal{H}_t(\hat{x})$ has a unique representation of the form

$$x_t = \sum_{i=0}^{\infty} w_{i,t-i}, \tag{8}$$

where the random variable $w_{i,t-i}$ represents the projection of x_t onto $\mathcal{D}_{t-i}$ for any $i \in \mathbb{Z}_+$. By the orthogonality of the sequence of subspaces $\{\mathcal{D}_t\}$, the process $\{w_{i,t}\}$ is uncorrelated over time for each $i \in \mathbb{Z}_+$.

While equation (8) looks almost like a news representation, it does not satisfy Definition (1) because it may be that $w_{i,t} \not\perp w_{j,t}$ for some $i \neq j$. Therefore, we use a version of the Gram-Schmidt orthogonalization procedure to transform these into an orthogonal sequence of shocks (cf. Luenberger, 1969, Theorem 3.5.1). Specifically, let us define:

$$\epsilon^a_{0,t} = w_{0,t}$$
$$\epsilon^a_{i,t} = w_{i,t} - \sum_{j=0}^{i-1} \phi_{i,j}\epsilon^a_{j,t} \quad \text{for } i > 0,$$

where $\phi_{i,j} \equiv \langle w_{i,t}, \epsilon^a_{j,t}\rangle / \|\epsilon^a_{j,t}\|^2$ is a projection coefficient. Define the index set $\mathcal{I}$ to be the set of indices $i \in \mathbb{Z}_+$ such that $\|\epsilon^a_{i,t}\| > 0$. The collection of orthogonal shocks $\epsilon_{i,t}$ with $i \in \mathcal{I}$ is uniquely determined because the collection of input shocks $w_{i,t}$ with $i \in \mathbb{Z}_+$ is unique. Substituting the orthogonalized shocks into equation (8), it follows that x_t can be uniquely rewritten as:

$$x_t = \sum_{i=0}^{\infty}\sum_{j\leq i} \phi_{i,j}\epsilon^a_{j,t-i} = \sum_{j\in\mathcal{I}}\sum_{i=j}^{\infty} \phi_{i,j}\epsilon^a_{j,t-i} = \sum_{j\in\mathcal{I}} a_{j,t-j}.$$

The second equality rearranges the indexes on the double summation, and the third equality introduces the definition $a_{j,t-j} \equiv \sum_{i=j}^{\infty} \phi_{i,j}\epsilon^a_{j,t-i}$. The fact that the orthogonalized shocks are also uncorrelated over time implies that $a_{j,t} \perp a_{k,\tau}$ for all $j \neq k$ and $t, \tau \in \mathbb{Z}$. Therefore, this defines a news representation when people's date-t information set $\mathcal{H}_t(a)$.

What remains is to prove that the expectations implied by this news representation are in fact equal to $\{\hat{x}_{i,t}\}$ for any $i \in \mathbb{Z}$. Under rational expectations, the i-step ahead expectation of x_t at date t under the original noise representation is equal to the orthogonal projection of x_{t+i} onto $\mathcal{H}_t(\hat{x})$: $\hat{x}_{j,t} = E[x_{t+j}|\mathcal{H}_t(\hat{x})]$. By the uniqueness of orthogonal projections, it must be that

$$w_{i,t} = \hat{x}_{i,t} - \hat{x}_{i+1,t-1},$$

where $w_{i,t}$ was defined in equation (8). Therefore, $\mathcal{H}_t(w) = \mathcal{H}_t(\hat{x})$. But then because $\mathcal{H}_t(a) = \mathcal{H}_t(w)$ by construction, it follows that $\mathcal{H}_t(a) = \mathcal{H}_t(\hat{x})$. So expectations are indeed the same under both representations:

$$\hat{x}_{i,t} = E[x_{t+i}|\mathcal{H}_t(\hat{x})] = E[x_{t+i}|\mathcal{H}_t(a)],$$

which completes the proof of the first part of the theorem.

To prove the second part of the theorem, we start from the (unique) news representation and define

$$s_{i,t} \equiv a_{i,t} \quad \text{for all } i \in \mathcal{I}.$$

Because $\mathcal{H}(x) \subset \mathcal{H}(a)$, there exist unique elements $m_{i,t} \in \mathcal{H}(x)$ and $v_{i,t} \in \mathcal{H}(s) \ominus \mathcal{H}(x)$ such that

$$s_{i,t} = m_{i,t} + v_{i,t}.$$

This defines a noise representation when people's date-t information set is $\mathcal{H}_t(s)$. What remains is to prove that the expectations implied by this noise representation are the same as the ones implied by the original news representation. Because $\mathcal{H}_t(s) = \mathcal{H}_t(a)$ by construction, and $\mathcal{H}_t(a) = \mathcal{H}_t(\hat{x})$ by the definition of a news representation, it follows that $\mathcal{H}_t(s) = \mathcal{H}_t(\hat{x})$ and therefore expectations are the same:

$$\hat{x}_{i,t} = E[x_{t+i}|\mathcal{H}_t(\hat{x})] = E[x_{t+i}|\mathcal{H}_t(s)].$$

This completes the proof of the second part of the theorem. $\square$

Proof of Proposition (2). Observational equivalence requires that the dynamics of the forecasts $\{\hat{x}_{i,t}\}$ are the same for all $i, t \in \mathbb{Z}$. By the endogeneity of the process $\{c_t\}$, $c_t \in \mathcal{H}_t(\hat{x})$, so the dynamics of $\{c_t\}$ must also be the same. $\square$

Proof of Proposition (3). As in the proof of Proposition (1), we can equate the c.g.f. of $d_t \equiv (x_t, \hat{x}_t)'$ that is implied by each representation:

$$g_d(z) = \underbrace{\begin{bmatrix} \sigma_\eta^2 + \sigma_\lambda^2 & \left(\frac{\sigma_\lambda^4}{\sigma_\lambda^2+\sigma_\xi^2}\right) z \\ \left(\frac{\sigma_\lambda^4}{\sigma_\lambda^2+\sigma_\xi^2}\right) z^{-1} & \frac{\sigma_\lambda^4}{\sigma_\lambda^2+\sigma_\xi^2} \end{bmatrix}}_{\text{system (4)}} = \underbrace{\begin{bmatrix} \sigma_x^2 & \left(\frac{\sigma_x^4}{\sigma_x^2+\sigma_v^2}\right) z \\ \left(\frac{\sigma_x^4}{\sigma_x^2+\sigma_v^2}\right) z^{-1} & \frac{\sigma_x^4}{\sigma_x^2+\sigma_v^2} \end{bmatrix}}_{\text{noise}}.$$

This equality holds if and only if the relations in Proposition (3) are satisfied. $\square$

Proof of Proposition (4). Consider an arbitrary noise representation of fundamentals and beliefs and an arbitrary endogenous process $\{c_t\}$. Using the structure of signals in a noise representation, $\mathcal{H}(s) = \mathcal{H}(m) \oplus \mathcal{H}(v)$. Because $v_{i,t} \in \mathcal{H}(s) \ominus \mathcal{H}(x)$ for all $i \in \mathcal{I}$, the uniqueness of orthogonal decompositions implies that $\mathcal{H}(m) = \mathcal{H}(x)$. Therefore, $\mathcal{H}(s) = \mathcal{H}(x) \oplus \mathcal{H}(v)$. Furthermore, the definition of noise shocks implies that $\mathcal{H}(\epsilon^v) = \mathcal{H}(v)$, so

$$\mathcal{H}(s) = \mathcal{H}(x) \oplus \mathcal{H}(\epsilon^v). \tag{9}$$

By the endogeneity of $\{c_t\}$ and the rationality of expectations, $c_t \in \mathcal{H}(s)$ for all $t \in \mathbb{Z}$. Combining this with Equation (9), it follows that for each c_t, there exist two unique elements $a_t \in \mathcal{H}(x)$ and $b_t \in \mathcal{H}(\epsilon^v)$ such that

$$c_t = a_t + b_t.$$

To consider variance decompositions at different frequencies, let $f_y(\omega)$ denote the spectral density function of a stochastic process $\{y_t\}$. Then because $a_t \perp b_t$ for all $t \in \mathbb{Z}$, it follows that

$$f_c(\omega) = f_a(\omega) + f_b(\omega),$$

where the functions $f_a(\omega)$ and $f_b(\omega)$ are uniquely determined by the processes $\{a_t\}$ and $\{b_t\}$. These functions in turn uniquely determine the share of the variance of $\{c_t\}$ due to noise shocks in any frequency range $\underline{\omega} < \omega < \overline{\omega}$, which is equal to

$$\frac{\int_{\underline{\omega}}^{\overline{\omega}} f_b(\omega) d\omega}{\int_{\underline{\omega}}^{\overline{\omega}} f_c(\omega) d\omega}.$$

The share due to fundamentals is equal to one minus this expression. □

Proof of Corollary (2). Consider an arbitrary noise representation of fundamentals and beliefs, and an endogenous process $\{c_t\}$. By the rationality of expectations, people's best forecast of c_{t+h} as of date t is equal to

$$\hat{c}_{h,t} = E[c_{t+h}|\mathcal{H}_t(s)] = E[c_{t+h}|\mathcal{H}_t(\hat{x})].$$

Therefore, $\hat{c}_{h,t} \in \mathcal{H}_t(\hat{x})$. This means that the forecast error $w_t^h \equiv c_t - \hat{c}_{h,t-h}$ also satisfies $w_{h,t} \in \mathcal{H}_t(\hat{x})$. Therefore, $\{w_t^h\}$ is an endogenous process. By Proposition (4), the variance decomposition of this process in terms of noise and fundamentals is uniquely determined over any frequency range. Moreover, this result is true for any forecast horizon $h \in \mathbb{Z}$ because h was chosen arbitrarily. □

Proof of Proposition (5). The proof of this result is a straightforward generalization of the proof of Proposition (1). In a news representation with i.i.d. news processes, the cross-c.g.f. of any two forecast processes $\{\hat{x}_{i,t}\}$ and $\{\hat{x}_{j,t}\}$ for $i, j \in \mathbb{Z}_+$ is equal to

$$g_{i,j}(z) = \sum_{k \in \mathcal{K}} \sigma_{a,k}^2 z^{j-i}, \tag{10}$$

where $\mathcal{K}$ is defined as the set of indices $k \in \mathcal{I}$ such that $k \geq |j-i| + i$. In a news representation of the type described in the proposition, the cross-c.g.f. of any two forecast processes $\{\hat{x}_{i,t}\}$ and $\{\hat{x}_{j,t}\}$ for $i, j \in \mathbb{Z}_+$ is equal to

$$\begin{aligned} g_{0,0}(z) &= \sigma_x^2 \qquad (11) \\ g_{i,j}(z) &= \sigma_x^2 \left[1 + \frac{1/\sigma_x^2}{\sum_{k \in \mathcal{K}} 1/\sigma_{v,k}^2} \right]^{-1} z^{j-i} \quad \text{for } i, j > 0. \end{aligned}$$

Equating the c.g.f.'s in (10) with those in (11), and recursively solving for the parameters of the noise representation delivers the relations stated in the proposition. □

Proof of Corollary (3). Define the composite shock

$$\epsilon_t^x \equiv \epsilon_{0,t}^a + \epsilon_{4,t-4}^a + \epsilon_{8,t-8}^a. \tag{12}$$

The process $\{\epsilon_t^x\}$ is i.i.d. because $\{\epsilon_{i,t}^a\}$ is i.i.d. for each $i \in \mathcal{I} \equiv \{0, 4, 8\}$. People's date-$t$ information set in representation (5) is $\mathcal{H}_t(\epsilon^a)$. But based on this information set, equation (12) defines a news representation for $\{\epsilon_t^x\}$ with i.i.d. news processes. Therefore, we can apply Proposition (5) to the composite shock process, which gives the relations stated in the corollary. Finally, note that $\hat{x}_{i,t} \in \mathcal{H}_t(\hat{\epsilon}^x)$ for all $t \in \mathbb{Z}$, which means that each $\{\hat{x}_{i,t}\}$ is endogenous with respect to $\{\epsilon_t^x\}$. Therefore, observational equivalence in terms of $\{\epsilon_t^x\}$ implies observational equivalence in terms of $\{x_t\}$. □

Proof of Proposition (6). According to representation (6), the two signals observed by people in the economy are $s_{0,t} \equiv x_t$ and $s_{1,t} \equiv \lambda_t + \xi_t$. Because $\mathcal{H}(x) \subset \mathcal{H}(s)$, there exist two unique elements $m_t \in \mathcal{H}(x)$ and $v_t \perp \mathcal{H}(x)$ such that:

$$s_{1,t} = m_t + v_t \quad \text{for all } t \in \mathbb{Z}.$$

Using the finite-order ARMA restrictions in system (6), it follows that m_t is related to $\{x_t\}$ according to (cf. Whittle, 1983, ch. 5):

$$m_t = \alpha(L)x_t, \quad \alpha(z) \equiv \frac{\sigma_\lambda^2}{z[\sigma_\lambda^2 + \sigma_\eta^2(1-\rho z)(1-\rho z^{-1})]} \tag{13}$$

where L is the lag operator and $z \in \mathbb{C}$. We can factor this expression for $\alpha(z)$ as

$$\alpha(z) = \frac{\delta\sigma_\lambda^2}{\rho\sigma_\eta^2 z(1-\delta z)(1-\delta z^{-1})},$$

where δ is the root of the quadratic polynomial

$$\mathcal{P}(z) = \rho z^2 - \left(1 + \rho^2 + \frac{\sigma_\lambda^2}{\sigma_\eta^2}\right) z + \rho$$

that lies inside the unit circle. Plugging the law of motion for $\{x_t\}$ into Equation (13), we find that the c.g.f. of $\{m_t\}$ is

$$g_m(z) = \frac{\delta\sigma_\lambda^4}{\rho\sigma_\eta^2(1-\rho z)(1-\rho z^{-1})(1-\delta z)(1-\delta z^{-1})},$$

which is in canonical form (cf. Whittle, 1983, ch. 2). This means that $\{m_t\}$ has an ARMA(2,0) representation of the form:

$$m_t = \phi_1 m_{t-1} + \phi_2 m_{t-2} + \epsilon_t^x, \quad \epsilon_t^x \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_x^2),$$

where $\phi_1 \equiv \rho + \delta$, $\phi_2 \equiv -\rho\delta$, and $\sigma_x^2 \equiv \delta\sigma_\lambda^4/(\rho\sigma_\eta^2)$. Inverting the relation in Equation (13), we have that $x_t = \alpha(L)^{-1} m_t$. Expanding the lag polynomial, it follows that $\{x_t\}$ has a representation of the form:

$$x_t = \gamma_0 m_t + \gamma_1 m_{t-1} + \gamma_0 m_{t-2},$$

where $\gamma_0 \equiv -\rho\sigma_\eta^2/\sigma_\lambda^2$ and $\gamma_1 \equiv -\gamma_0(1+\delta^2)/\delta$. Finally, using the definition of v_t,

$$g_v(z) = \frac{\sigma_\lambda^2 \sigma_\eta^2}{\sigma_\lambda^2 + \sigma_\eta^2(1-\rho z)(1-\rho z^{-1})} + \sigma_\xi^2 = \sigma_\xi^2 \frac{\delta}{\beta} \frac{(1-\beta z)(1-\beta z^{-1})}{(1-\delta z)(1-\delta z^{-1})},$$

where β is the root of the polynomial

$$\mathcal{P}(z) = \rho z^2 - \left(1 + \rho^2 + \frac{\sigma_\lambda^2(\sigma_\lambda^2 + \sigma_\xi^2)}{\sigma_\eta^2 \sigma_\xi^2}\right) z + \rho$$

that lies inside the unit circle. This is the canonical form for $g_v(z)$, which means that $\{v_t\}$ has an ARMA(1,1) representation of the form:

$$v_t = \delta v_{t-1} + \epsilon_t^v - \beta \epsilon_{t-1}^v, \quad \epsilon_t^v \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_v^2),$$

where $\sigma_v^2 \equiv \sigma_\xi^2 \delta/\beta$. Because $\mathcal{H}_t(s)$ is unchanged from representation (6) for all $t \in \mathbb{Z}$, it follows that $\hat{x}_{i,t} \equiv E[x_{t+i}|\mathcal{H}_t(s)]$ is also unchanged for any $i \in \mathbb{Z}$. Therefore these two representations are observationally equivalent. Uniqueness follows from the uniqueness of the processes $\{m_t\}$ and $\{v_t\}$ in the orthogonal projection of $s_{1,t}$ on $\mathcal{H}(x)$ at each date, and the fact that the polynomials defining δ and β each have only one root inside the unit circle. $\square$

Proof of Proposition (7). According to representation (7), the two signals observed by people in the economy are $s_{0,t} \equiv x_t$ and $s_{1,t} \equiv -\eta_t + \xi_t$. Because $\mathcal{H}(x) \subset \mathcal{H}(s)$, there exist two unique elements $\tilde{m}_t \in \mathcal{H}(x)$ and $\tilde{v}_t \perp \mathcal{H}(x)$ such that:

$$s_{1,t} = \tilde{m}_t + \tilde{v}_t \quad \text{for all } t \in \mathbb{Z}.$$

Using the finite-order ARMA restrictions in system (7), it follows that $\tilde{m}_t$ is related to $\{x_t\}$ according to:

$$\tilde{m}_t = \alpha(L)x_t, \quad \alpha(z) \equiv -\frac{(1-z^{-1})\sigma_\eta^2}{\sigma_\lambda^2 + (1-z)(1-z^{-1})\sigma_\eta^2},$$

where L is the lag operator and $z \in \mathbb{C}$. Using the parametric restriction from Blanchard, L'Huillier, and Lorenzoni (2013) that $\rho\sigma_\lambda^2 = (1-\rho)^2\sigma_\eta^2$, we can factor this expression for $\alpha(z)$ as

$$\alpha(z) = -\frac{(1-z^{-1})\rho}{(1-\rho z)(1-\rho z^{-1})}.$$

Now define the process $m_t \equiv (1-\rho L^{-1})^{-1}x_t$, so we can write $\tilde{m}_t = \rho(1-L^{-1})/(1-\rho L)m_t$. Using the law of motion for $\{x_t\}$, it follows that

$$m_t = \rho m_{t-1} + \epsilon_t^x, \quad \epsilon_t^x \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_x^2),$$

where $\sigma_x^2 \equiv \sigma_\eta^2/\rho$. Defining the new signal variable $s_t \equiv (1-\rho L)s_{1,t}$ and the new noise variable $v_t \equiv (1-\rho L)\tilde{v}_t$, we have that

$$s_t = \rho m_{t+1} - \rho m_t + v_t.$$

Note that because $|\rho| < 1$, the signal process $\{s_t\}$ spans the same closed subspace as $\{s_{1,t}\}$. Finally, using the definition of $\{v_t\}$,

$$\begin{aligned} g_v(z) &= \frac{\rho\sigma_\lambda^2 + \sigma_\xi^2[(1-\rho z)(1-\rho z^{-1}]^2}{(1-\rho z)(1-\rho z^{-1})} \\ &= \sigma_v^2 \frac{\rho^2}{\beta_1\beta_2} \frac{(1-\beta_1 z)(1-\beta_1 z^{-1})(1-\beta_2 z)(1-\beta_2 z^{-1})}{(1-\rho z)(1-\rho z^{-1})}, \end{aligned}$$

where β_1 and β_2 are the two roots of the polynomial

$$\mathcal{P}(z) = \rho^2 z^4 - 2\rho(1+\rho^2)z^3 + \left(1 + \rho^4 + 4\rho^2 + \rho\frac{\sigma_\lambda^2}{\sigma_\xi^2}\right) z - 2\rho(1+\rho^2)z + \rho^2$$

that lie inside the unit circle. This is the canonical form for $g_v(z)$, which means that $\{v_t\}$ has an ARMA(1,2) representation of the form

$$v_t = \phi v_{t-1} + \epsilon_t^v - (\beta_1 + \beta_2)\epsilon_{t-1}^v + \beta_1\beta_2\epsilon_{t-2}^v, \quad \epsilon_t^v \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_v^2),$$

where $\sigma_v^2 \equiv \rho^2\sigma_\xi^2/(\beta_1\beta_2)$. Because $\mathcal{H}_t(s)$ is the same as in representation (7) for all $t \in \mathbb{Z}$, it follows that $\hat{x}_{i,t} \equiv E[x_{t+i}|\mathcal{H}_t(s)]$ is also unchanged for any $i \in \mathbb{Z}$. Therefore, these two representations are observationally equivalent. Uniqueness follows from the uniqueness of the processes $\{m_t\}$ and $\{v_t\}$ in terms of $\{\tilde{m}_t\}$ and $\{\tilde{v}_t\}$ in the orthogonal projection of $s_{1,t}$ on $\mathcal{H}(x)$ at each date, and the fact that the polynomial defining β_1 and β_2 only has two roots inside the unit circle. $\square$

Proof of Proposition (8). Consider an arbitrary noise representation of fundamentals and beliefs. Because $\hat{x}_{i,t} \in \mathcal{H}(s)$ and $\mathcal{H}(x) \oplus \mathcal{H}(v)$ (cf. the proof of Proposition (4)), it follows that $\hat{x}_{i,t}$ has a unique decomposition of the form

$$\hat{x}_{i,t} = \hat{m}_{i,t} + \hat{v}_{i,t},$$

where $\hat{m}_{i,t} \in \mathcal{H}(x)$ and $\hat{v}_{i,t} \perp \mathcal{H}(x)$ for any $i \in \mathbb{Z}$. To prove that $\mathcal{D}_t \equiv \mathcal{H}_t(v) \ominus \mathcal{H}_{t-1}(v)$ is uniquely determined for all $t \in \mathbb{Z}$, it is sufficient to prove that $\mathcal{H}_t(v) = \mathcal{H}_t(\hat{v})$. This is because the uniqueness of orthogonal decompositions ensures that if $\mathcal{H}_t(v) = \mathcal{H}_t(\hat{v})$ then $\mathcal{H}_t(v) \ominus \mathcal{H}_{t-1}(v) = \mathcal{H}_t(\hat{v}) \ominus \mathcal{H}_{t-1}(\hat{v})$.

Because $\mathcal{H}_t(s) = \mathcal{H}_t(\hat{x})$, $s_{i,t}$ has a unique representation in terms of current and past expectations:

$$s_{i,t} = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \psi_{i,j,k} \hat{x}_{k,t-j},$$

where $\psi_{i,j,k} \equiv \langle s_{i,t}, \hat{x}_{k,t-j} \rangle / \|\hat{x}_{k,t}\|^2$. Projecting both sides onto the subspace $\mathcal{H}(s) \ominus \mathcal{H}(x)$, and using the linearity of orthogonal projections, we can write

$$v_{i,t} = \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \psi_{i,j,k} \hat{v}_{k,t-j},$$

so $\mathcal{H}_t(v) \subseteq \mathcal{H}_t(\hat{v})$. Conversely, $\mathcal{H}_t(\hat{x}) = \mathcal{H}_t(s)$ implies that $\mathcal{H}_t(\hat{v}) \subseteq \mathcal{H}_t(v)$ because

$$\hat{v}_{i,t} = \sum_{j=0}^{\infty} \sum_{k \in \mathcal{I}} \theta_{i,j,k} v_{k,t-j},$$

where $\theta_{i,j,k} \equiv \langle \hat{x}_{i,t}, s_{k,t-j} \rangle / \|s_{k,t}\|^2$. Therefore, $\mathcal{H}_t(v) = \mathcal{H}_t(\hat{v})$. Finally, because $\hat{x}_{0,t} = x_t$ by $\mathcal{H}_t(x) \subset \mathcal{H}_t(s)$ and the rationality of expectations, it follows that the space spanned by fundamental shocks is uniquely determined as well. $\square$

Proof of Proposition (9). Begin with an arbitrary news representation. By definition of shocks, $\mathcal{H}_t(\epsilon^a) = \mathcal{H}_t(a)$. By the rationality of expectations, $\mathcal{H}_t(\hat{x}) = \mathcal{H}_t(a)$. Therefore, $\mathcal{H}_t(\epsilon^a) = \mathcal{H}_t(\hat{x})$, so this is an invertible representation.

Now consider an arbitrary noise representation. Suppose that it is invertible. Then $\epsilon^v_{i,t} \in \mathcal{H}_t(\hat{x})$ for any $i \in \mathcal{I} \subseteq \mathbb{Z}_+$ and $t \in \mathbb{Z}$. This implies that the noise shocks are contained in the information set of agents. But by rationality, if $\epsilon^v_{i,t}$ is contained in the information set of agents at date t, because it is uncorrelated with fundamentals, $\epsilon^v_{i,t} \notin \mathcal{H}_t(\hat{x})$. This is a contradiction. Therefore, the representation is not invertible. □

Proof of Proposition (10). By invertibility, $\mathcal{H}_t(\epsilon) = \mathcal{H}_t(y) = \mathcal{H}_t(\tilde{\epsilon})$ for all $t \in \mathbb{Z}$. Then by the uniqueness of orthogonal projections:

$$\mathcal{H}_t(\epsilon) \ominus \mathcal{H}_{t-1}(\epsilon) = \mathcal{H}_t(\tilde{\epsilon}) \ominus \mathcal{H}_{t-1}(\tilde{\epsilon}) \quad \text{for all } t \in \mathbb{Z}.$$

□

Proof of Proposition (11). By definition, the collection of noise shocks $\{\epsilon^v_{i,t}\}$ for $i \in \mathcal{I}$ and fixed $t \in \mathbb{Z}$ generates the subspace $\mathcal{D}_t \equiv \mathcal{H}_t(v) \ominus \mathcal{H}_{t-1}(v)$. By Proposition (8), $\mathcal{H}_t(v) = \mathcal{H}_t(\hat{v})$, so $\mathcal{D}_t = \mathcal{H}_t(\hat{v}) \ominus \mathcal{H}_{t-1}(\hat{v})$. Therefore, the shock $\hat{\epsilon}^v_{i,t} \equiv \hat{v}_{i,t} - E[\hat{v}_{i,t}|\mathcal{H}_{t-1}(\hat{v})]$ can be represented in the form:

$$\hat{\epsilon}^v_{i,t} = \sum_{j \in \mathcal{I}} \gamma_{i,j} \epsilon^v_{j,t}, \tag{14}$$

where $\gamma_{i,j} \equiv \langle \hat{\epsilon}^v_{i,t}, \epsilon^v_{j,t} \rangle / \|\epsilon^v_{j,t}\|^2$. By Assumption (1), the collection $\{\epsilon^v_{i,t}\}$ for fixed t is an orthogonal basis for $\mathcal{D}_t$ with $\gamma_{i,i} = 1$ and $\gamma_{i,j} = 0$ for all $i < j \in \mathcal{I}$. This implies that the noise shocks can be recovered by recursively solving system (14) at each date $t \in \mathbb{Z}$ to obtain $\{\epsilon^v_{i,t}\}$ in terms of the shocks $\{\hat{\epsilon}^v_{i,t}\}$ for all $i \in \mathcal{I}$:

$$\epsilon^v_{i,t} = \hat{\epsilon}^v_{i,t} - \sum_{j<i} \gamma_{i,j} \epsilon^v_{j,t}, \quad \text{for all } i \in \mathcal{I}.$$

Finally, the fundamentals shocks $\{\epsilon^x_t\}$ can be recovered from the history of the process $\{x_t\}$, which is observable because $\mathcal{H}_t(x) \subset \mathcal{H}_t(s)$ for all $t \in \mathbb{Z}$. □

B Quantitative Models

The following subsections provide a sketch of each of the three quantitative models considered in this paper. For more details, we refer the reader to the original articles and their supplementary material.

B.1 Model of Schmitt-Grohé and Uribe (2012)

A representative household chooses consumption $\{C_t\}$, labor supply $\{h_t\}$, investment $\{I_t\}$, and the utilization rate of existing capital $\{u_t\}$ to maximizes its lifetime utility subject to a standard budget constraint:

$$\max \quad E\sum_{t=0}^{\infty}\beta^t\zeta_t\frac{(C_t - bC_{t-1} - \psi h_t^{\theta} S_t)^{1-\sigma}}{1-\sigma} \quad \text{subject to}$$

$$\begin{aligned} S_t &= (C_t - bC_{t-1})^{\gamma} S_{t-1}^{1-\gamma} \\ C_t + A_t I_t + G_t &= \frac{W_t}{\mu_t} h_t + r_t u_t K_t + P_t \\ K_{t+1} &= (1-\delta(u_t))K_t + z_t^I I_t \left[1 - \Phi\left(\frac{I_t}{I_{t-1}}\right)\right] \end{aligned}$$

Relative to the standard real business cycle model, this model features investment adjustment costs $\Phi(I_t/I_{t-1})$; variable capacity utilization, which increases the return on capital $r_t u_t$ at the cost of increasing its rate of depreciation through $\delta(u_t)$; one period internal habit formation in consumption, controlled by $0 < b < 1$; a potentially low wealth effect on labor supply, when $0 < \gamma < 1$ approaches its lower limit; and monopolistic labor unions, which effectively reduce the wage rate by an amount μ_t each period but rebate profits lump sum to the household through P_t.

Output is produced by a representative firm, which combines capital K_t, labor h_t, and a fixed factor of production L using a (potentially) decreasing returns to scale production function:

$$Y_t = z_t(u_t K_t)^{\alpha_k}(X_t h_t)^{\alpha_h}(X_t L)^{1-\alpha_k-\alpha_h}.$$

Market clearing requires that the goods and labor markets clear so that the aggregate resource constraint is satisfied: $C_t + A_t I_t + G_t = Y_t$. The seven fundamental processes capture exogenous variation in permanent and transitory neutral productivity $\{X_t, z_t\}$, permanent and transitory investment-specific productivity $\{A_t, z_t^I\}$, government spending $\{G_t\}$, wage markups $\{\mu_t\}$, and preferences $\{\zeta_t\}$.

B.2 Model of Barsky and Sims (2012)

A representative household chooses consumption $\{C_t\}$, labor supply $\{N_t\}$, and real holdings of riskless one-period bonds $\{B_t\}$ to maximize its lifetime utility subject to a standard budget constraint:

$$\max \quad E\sum_{t=0}^{\infty}\beta^t\left[\ln(C_t-\kappa C_{t-1})-\frac{N_t^{1+1/\eta}}{1+1/\eta}\right] \quad \text{subject to}$$

$$C_t+B_t=w_tN_t-T_t+(1+r_{t-1})B_{t-1}+\Pi_t,$$

where r_t is the net nominally risk-free interest rate, w_t is the wage, T_t denotes lump-sum taxes, and Π_t is aggregate profits.

Final goods producers are competitive and take the price of intermediate goods, $P_t(j)$, as given and each have a production function of the form:

$$Y_t=\left[\int_0^1 Y_t(j)^{\frac{\xi-1}{\xi}}\right]^{\frac{\xi}{\xi-1}}$$

Intermediate goods firms are monopolistically competitive and take the demands of final goods firms as given. They each have a production function of the form $Y_t(j)=A_tK_t(j)^{\alpha}N_t(j)^{1-\alpha}$. Each intermediate firm chooses a price for its own good, subject to the constraint that it will only be able to re-optimize its price each period with constant probability $1-\theta$.

A continuum of capital producers produce new capital (to sell to intermediate firms) according to the production function

$$Y_t^k(\nu)=\phi\left(\frac{I_t(\nu)}{K_t(\nu)}\right)K_t(\nu),$$

where ϕ is an increasing and concave function. The aggregate capital stock evolves according to $K_t=\phi(I_t/K_t)K_{t-1}+(1-\delta)K_{t-1}$, where $0<\delta<1$ is the depreciation rate. The aggregate resource constraint is $Y_t=C_t+I_t+G_t$ (ignoring resources lost due to inefficient price dispersion). The monetary authority sets the one-period nominally risk-free rate of return according to a feedback rule of the (log-linear approximate) form:

$$i_t=\rho_i i_{t-1}+(1-\rho_i)\phi_\pi(\pi_t-\pi^*)+(1-\rho_i)\phi_y(\Delta Y_t-\Delta Y^*)+\varepsilon_{i,t}.$$

The three fundamental processes capture exogenous variation in permanent neutral productivity $\{A_t\}$, government spending $\{G_t\}$, and monetary policy $\{\varepsilon_{i,t}\}$

B.3 Model of Blanchard, L'Huillier, and Lorenzoni (2013)

A representative household chooses consumption $\{C_t\}$, investment $\{I_t\}$, nominally risk-free bond holdings $\{B_t\}$, and the rate of capital utilization $\{U_t\}$ to maximize its lifetime utility subject to a standard budget constraint:

$$\max \quad E\sum_{t=0}^{\infty}\beta^t\left[\ln(C_t - hC_{t-1}) - \frac{1}{1+\zeta}\int_0^t N_{j,t}^{1+\zeta}dj\right] \quad \text{subject to}$$

$$P_tC_t + P_tI_t + T_t + P_t\mathcal{C}(U_t)\bar{K}_{t-1} = R_{t-1}B_{t-1} + \Upsilon_t + \int_0^1 W_{jt}N_{jt}dj + R_t^kU_t\bar{K}_{t-1},$$
$$\bar{K}_t = (1-\delta)\bar{K}_{t-1} + D_t[1-\mathcal{G}(I_t/I_{t-1})]I_t$$

where P_t is the price level, T_t is a lump sum tax, R_t is the gross nominally risk-free rate, Υ_t is aggregate profits, W_{jt} is the wage of labor type j, R_t^k is the capital rental rate, $0 < \delta < 1$ is the rate of depreciation, $\mathcal{G}(I_t/I_{t-1})$ represents investment adjustment costs, $\mathcal{C}(U_t)$ represents the marginal cost of increasing capacity utilization. It also chooses the wage $\{W_{jt}\}$ for each type of labor subject to the constraint that it will only be able to re-optimize its wage each period with constant probability $1-\theta_w$.

Final goods producers are competitive and take the price of intermediate goods as given, P_{jt}, and each have a production function of the form

$$Y_t = \left[\int_0^1 Y_{jt}^{\frac{1}{1+\mu_{pt}}}dj\right]^{1+\mu_{pt}}.$$

Intermediate goods firms are monopolistically competitive, each with a production function of the form $Y_{jt} = (K_{jt})^{\alpha}(A_tL_{jt})^{1-\alpha}$. Each intermediate firm chooses a price for its own good, subject to a $1-\theta_p$ probability of re-optimization each period.

Labor services are supplied to intermediate goods producers by competitive labor agencies that take wages as given, W_{jt}, and have a production function of the form

$$N_t = \left[\int_0^1 N_{jt}^{\frac{1}{1+\mu_{wt}}}dj\right]^{1+\mu_{wt}}.$$

Market clearing in the final goods market requires that $C_t + I_t + \mathcal{C}(U_t)\bar{K}_{t-1} + G_t = Y_t$, and in the labor market that $\int_0^1 L_{jt}dj = N_t$. Monetary policy follows the rule:

$$r_t = \rho_r r_{t-1} + (1-\rho_r)(\gamma_\pi\pi_t + \gamma_y\hat{y}_t) + q_t.$$

The six fundamental processes capture exogenous variation in permanent neutral productivity $\{A_t\}$, transitory investment-specific productivity $\{D_t\}$, price markups $\{\mu_{pt}\}$, wage markups $\{\mu_{wt}\}$, government spending $\{G_t\}$, and monetary policy $\{q_t\}$.