News or Noise? The Missing Link

Ryan Chahrour
Boston College

Kyle Jurado
Duke University

November 2, 2017

Abstract

The literature on belief-driven business cycles treats news and noise as distinct representations of agents' beliefs. We prove they are empirically the same. Our result lets us isolate the importance of purely belief-driven fluctuations. Using three prominent estimated models, we show that existing research understates the importance of pure beliefs. We also explain how differences in both economic environment and information structure affect the estimated importance of pure beliefs.

JEL classification: D84, E32, C31
Keywords: News, noise, business cycles

Chahrour: Department of Economics, Boston College, 140 Commonwealth Avenue, Chestnut Hill, MA 02467 (e-mail: ryan.chahrour@bc.edu); Jurado: Department of Economics, Duke University, 419 Chapel Drive, Durham, NC 27708 (e-mail: kyle.jurado@duke.edu).

1 Introduction

A large literature in macroeconomics has argued that changes in agents' beliefs about the future can be an important cause of economic fluctuations.[1] This idea, which dates at least to Pigou (1927), has been formalized in two ways. In the first way, which we call a "news representation," agents perfectly observe part of an exogenous fundamental in advance. As an analogy, this is like learning today that in next week's big game your favorite team will certainly 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. In the second way, which we call a "noise representation," agents imperfectly observe 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 more than you do, and is often right, but sometimes he gets it wrong.

Much of the literature emphasizes the differences between these two ways of representing agents' beliefs.[2]

agents' beliefs about them always have both a news representation and a noise representation. This implies that associated with every noise representation is an observationally equivalent news representation and vice versa. We present a constructive proof of the theorem using Hilbert space methods. Because it is constructive, our proof also provides a method for explicitly deriving the mapping from one representation to another. We compute this mapping in closed form for several models of interest from the literature.

The main step in moving from noise to news amounts to nding the Wold representation of the noise model. This is because the shocks in the news representation are static rotations of the Wold innovations implied by the noise 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" agents have perfect information.[3] To move in the opposite direction, from news to noise, the idea is to reverse engineer the signal extraction problem that generates a given Wold representation. The challenge is to ensure that the noise shocks in that signal extraction problem are independent of fundamentals at all leads and lags, and that they capture all the non-fundamental variation in beliefs.

Beyond clarifying the link between news and noise, our representation theorem sheds new light on the importance of purely belief-driven uctuations. Existing studies that either use models with only news shocks or some combination of news and noise shocks do not isolate the pure contribution of beliefs above and beyond fundamentals. The reason is that news shocks mix the uctuations due purely to beliefs with the those due to fundamentals. News shocks can change beliefs on impact without any change in current fundamentals, but they are tied by construction to changes in future fundamentals. Beliefs change today, and on average fundamentals change tomorrow. But which is more important, the change in beliefs or the subsequent change in fundamentals?

To isolate the contribution of pure beliefs, it is necessary to disentangle the e ects due only to expected changes in fundamentals fromidef9(con)equen ifheheheange7 -3.61(ts')-3238(b)4(

role compared to current and past fundamental shocks. For example, in the model of Barsky and Sims (2012), future fundamentals are responsible for less than 0.5% of consumption fluctuations, while current and past fundamentals are responsible for over 80%. Future fundamentals matter the most in the model of Blanchard et al. (2013), but even in that model they are responsible for less than 7% of consumption fluctuations.

We conclude our paper by investigating the sources of disagreement across models regarding the overall importance of noise shocks. We show that the disagreement is due to differences both in the models' economic environments and information structures. For noise shocks to play a large role, agents' actions need to depend heavily on their forecasts of future fundamentals (economic environment), and their forecasts in turn need to depend heavily on noise-ridden signals (information structure). The model of Blanchard et al. (2013) has both of these features, which is why they find a large role for noise shocks. In their model, productivity is a random walk, so agents rely heavily on their noisy signal to forecast future productivity. Nominal price and wage rigidity and an accommodative monetary policy rule work together to make agents' consumption decisions highly forward-looking, and allow the model to generate empirically realistic patterns of co-movement in response to a noise shock.

To quantify the relative contribution of economic environment and information structure on the estimated importance of noise shocks, we re-estimate the models of Barsky and Sims (2012) and Blanchard et al. (2013) using the same data, exogenous shocks, and estimation procedure (maximum likelihood) across both models. Consistent with the authors' original estimates, we find that noise shocks play a small role in the model of Barsky and Sims (2012) and a much larger role in the model of Blanchard et al. (2013). This suggests that differences in data, shocks, and estimation procedure are not the primary reasons these models deliver different estimates of the importance of noise shocks.

We then swap information structures and re-estimate both models. Substituting the information structure of Blanchard et al. (2013) into the economic environment of Barsky structure play an important role in generating a large role for noise shocks, di er-ences in environment turn out to be quantitatively more important in explaining the disagreement between these two models.

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:

$$x_t = a_{0;t} + a_{1;t-1}$$

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 that often encountered in time series analysis. To use a familiar example (cf. Hamilton, 1994, pp. 64-67), it is well-known that

$$y_t = \varepsilon_t - \theta\varepsilon_{t-1} \quad \varepsilon_t \overset{iid}{\sim} N(0;\sigma^2) \quad \text{and} \quad y_t = \tilde{\varepsilon}_t - \tilde{\theta}\tilde{\varepsilon}_{t-1} \quad \tilde{\varepsilon}_t \overset{iid}{\sim} N(0;\tilde{\sigma}^2) \tag{3}$$

are two observationally equivalent representations of the stationary MA(1) process $\{y_t\}$ when $\tilde{\theta} = 1/\theta$ and $\sigma^2 = \theta^2\tilde{\sigma}^2$. However, this applies only when $\{y_t\}$ is the sole observable. If (rational) expectations of future values of $\{y_t\}$ are also observable, then the two representations in (3) are no longer the same. To see why, note that the variance of the one-step-ahead rational forecast $\hat{y}_t \equiv E_t[y_{t+1}]$ is equal to $\theta^2\sigma^2$ under the first representation, but σ^2 under the second. Therefore, an econometrician observing $\{\hat{y}_t\}$ *and* $\{y_t\}$ (or independent functions of these objects) could discriminate between these two representations.

The following proposition states the equivalence result for the simple example of this subsection, and provides the parametric mapping from one representation to the other. Its proof is in the Appendix.

Proposition 1. *The news representation* (1) *is observationally equivalent to the noise representation* (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:

$$\begin{aligned} x_t &= \hat{x}_{t-1} + w_{0,t} \\ \hat{x}_t &= \quad w_{1,t} \end{aligned} \qquad \begin{bmatrix} w_{0,t} \\ w_{1,t} \end{bmatrix} \overset{iid}{\ } N\left(0, \begin{bmatrix} {}^2_v & 0 \\ 0 & {}^2_x + {}^2_v \end{bmatrix}\right), \qquad (4)$$

where $= {}^2_x / ({}^2_x + {}^2_v)$ is a Kalman gain parameter controlling how much agents trust the noisy signal, and $w_t \quad (w_{0,t}, w_{1,t})'$ is the vector of Wold innovations. But system (4) is the same as the news representation in system (1) when $a_{0,t} = w_{0,t}$ and $a_{1,t} = \quad w_{1,t}$. The situations arise, since we have an explicit probability distribution for the noise shocks: for example, how big is a "one standard deviation impulse" of a news reversal? Second, we can ask how important these types of news reversals are in the data overall; that is, we can do a proper variance decomposition. Third, in models with news shocks that are not i.i.d., it is not as straightforward to determine the configuration of news shocks that correspond to a noise shock. Therefore, it is desirable to have a more general characterization of the link between news and noise shocks. We turn to this more general characterization next.

2.2 Representation Theorem

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 L^2 to denote the space of (equivalence classes of) complex random variables with finite second moments, which is a Hilbert space when equipped with the inner product $(a, b) = E[a\bar{b}]$ for any $a, b \in L^2$. Completeness of this space is with respect to the norm $\|a\| \equiv (a, a)^{1/2}$. For any collection of random variables in L^2,

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

we let $H(y)$ denote the closed subspace spanned by the variables $y_{i,t}$ for all $i \in I_y$ and $t \in \mathbb{Z}$. Similarly, $H_t(y)$ denotes the closed subspace spanned by these variables over all i but only up through date t.

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, Gaussian, and purely non-deterministic.[5] The fact that $\{x_t\}$ is 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.

Agents' 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 rational expectations, $\hat{x}_{i,t}$ is equal to the mathematical expectation of x_{t+i} with respect to a particular date-t information set. We

5

assume that $\{x_t\}$ and $\{\hat{x}_{i,t}\}$ jointly form a Gaussian system; that is, the vector formed by any finite subset of these random variables is Gaussian. This allows us to summa-

The idea behind this representation is that agents may receive signals about the fundamental realization x_t prior to date t, but those signals are contaminated with noise. The variable $\eta^v_{i,t} \equiv v_{i,t} - E[v_{i,t}|H_{t-1}(v)]$ is called the "noise shock" associated with signal i. The variable $\eta^x_t \equiv x_t - E[x_t|H_{t-1}(x)]$ 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 agents cannot separately observe $m_{i,t}$ and $v_{i,t}$ at date t, their beliefs are still affected by noise. The condition that $H_t(s) = H_t(\hat{x})$ simply rules out redundant or totally uninformative signals.

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

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. The proof 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. We explain in the subsequent sections, however, that this multiplicity of noise representations does not pose much of a problem.

An implication of Theorem (1) is that any model *economy* with a news representation of fundamentals and beliefs has an observationally equivalent version with a noise representation of fundamentals and beliefs, and vice versa. This is because the equivalence of fundamentals and beliefs implies the equivalence of any endogenous processes that are functions of them. To make this statement more precise, we first define here what we mean by an endogenous process, and then present this statement as a proposition. The proof of the proposition, together with all remaining proofs, are contained in the Online Appendix.

Definition 3. Given a fundamental process $\{x_t\}$ and a collection of forecasts $\{\hat{x}_{i,t}\}$ satisfying $H_t($

Theorem (1) provides a way to determine the importance of pure beliefs as a driver of uctuations.

The rst subsection explains the problem with using news shocks to determine

$a_{0,t} + a_{1,t-1}$. Therefore, the fraction of the variation in $\{x_t\}$ due to news shocks, $\{a_{1,t}\}$ is given by:

$$\frac{\text{var}[x_t | a_{0,t} = 0]}{\text{var}[x_t]} = \frac{\text{var}[a_{1,t}]}{\text{var}[x_t]} = \frac{\sigma_{a,1}^2}{\sigma_{a,0}^2 + \sigma_{a,1}^2}. \tag{5}$$

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 pure beliefs from fundamentals in models with news shocks, we can 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 independent of fundamentals at all horizons, they capture precisely those changes in beliefs that cannot be explained by fundamentals. That is, noise shocks are pure belief shocks.

Returning to the example from Section (2.1), we have already shown that an observationally equivalent noise representation involves $x_t \overset{iid}{\sim} N(0, \sigma_x^2)$ with $\sigma_x^2 = \sigma_{a,0}^2 + \sigma_{a,1}^2$. Therefore, 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 pure 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 pure beliefs has lead some researchers to conclude that there is a fundamental problem with using variance decompositions for that purpose.[6] We would like to suggest that the problem is not with 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 decompositions can be performed as usual.

[6] For example, Sims (2016) p.42 describes the problem of identifying the importance of pure beliefs (which both he and Barsky et al. (2015) call “pure news”) as a fundamental limitation of the traditional variance decomposition.

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 et al., 2013; Barsky and Sims, 2012). A simple example is:

$$\begin{aligned} x_t &= \eta_{t-1} + \varepsilon_t \\ s_t &= \eta_t + \nu_t \end{aligned} \qquad \begin{bmatrix} \eta_t \\ \varepsilon_t \\ \nu_t \end{bmatrix} \stackrel{iid}{\sim} N\left(0, \begin{bmatrix} \sigma_\eta^2 & 0 & 0 \\ 0 & \sigma_\varepsilon^2 & 0 \\ 0 & 0 & \sigma_\nu^2 \end{bmatrix}\right). \tag{6}$$

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

The problem with this type of representation, at least from the perspective of isolating the importance of pure beliefs, is that while ν_t can generate non-fundamental fluctuations in beliefs, so can certain combinations of η_t and ε_t. To see this, notice that in the limit case $\nu_t = 0$, we have that $s_t = \eta_t$ and this representation collapses to a news representation with $a_{0,t} \equiv \varepsilon_t$ and $a_{1,t} \equiv \eta_t$. We have already seen in Proposition (1) that such a news representation has an observationally equivalent noise representation with (non-zero) noise shocks. Therefore $\nu_t = 0$ does not mean that beliefs do not have an independent role to play as a driver of fluctuations.[7]

Of course, Theorem (1) implies that the representation in (6), which is neither news or noise representation, still has an observationally equivalent noise representation. The following proposition presents the mapping from one representation to the other.

Proposition 3. *The representation of fundamentals and beliefs in* (6) *is observationally equivalent to the noise representation in*

To see how the process $\{\ _t\}$ understates the importance of pure beliefs, consider the endogenous variable $\hat{x}_t = E_t[x_{t+1}]$. Under representation (6), $\hat{x}_t = \frac{\ ^2}{\ ^2 + \ ^2}(\ _t + \ _t)$, so the contribution of the process $\{\ _t\}$ is

$$\frac{\mathrm{var}[\hat{x}_t | \ _t = \ _t = 0]}{\mathrm{var}[\hat{x}_t]} = \frac{\ ^2}{\ ^2 + \ ^2}.$$

On the other hand, in the observationally equivalent noise representation implied by Proposition (3), $\hat{x}_t = \frac{\ ^2_x}{\ ^2_x + \ ^2_v}(x_{t+1} + v_t)$. Therefore, the contribution of $\{v_t\}$ is

$$\frac{\mathrm{var}[\hat{x}_t | x_t = 0]}{\mathrm{var}[\hat{x}_t]} = \frac{\ ^2_v}{\ ^2_x + \ ^2_v} = \frac{\ ^2\ ^2}{(\ ^2 + \ ^2)(\ ^2 + \ ^2)} + \frac{\ ^2}{\ ^2 + \ ^2},$$

where the second equality uses the parametric restrictions from Proposition (3). Because the first term in this expression is positive, it follows that $\{\ _t\}$ understates the importance of pure beliefs for explaining variations in $\{\hat{x}_t\}$. It is also easy to see how the importance of pure beliefs can be strictly positive even as $\ ^2 \to 0$.

3.3 Different Noise Representations

So far we have argued that it is possible to use a noise representation to separate fluctuations that are due to actual changes in fundamentals versus those that are due purely to changes in beliefs. 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

An immediate corollary of this proposition is that the variance decomposition of agents' 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 1. *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.*

3.4 Past, Present, and Future Fundamentals

Our discussion in this section has focused on the distinction between the relative contributions of fundamental shocks and non-fundamental noise shocks. However, it is also possible to further decompose the contribution of fundamental shocks into parts separately due to past, present, and future fundamental shocks. Even if news shocks don't capture the contribution of noise shocks, maybe they capture something like the sum of the contribution of noise shocks and future fundamental shocks.

where the sum of the rst three parts equals the total contribution of fundamentals. Notice from this equation that even when there are no noise shocks ($_v$ = 0), the contribution of future fundamentals is not necessarily equal to zero. In that case, = 1, so the share of the variance of c_t

corresponding parameters from the observationally equivalent news representation:

$$ {}_1^2 \; {}^2 \; {}_x^2 + {}_1^2 \; {}^2 \; {}_v^2 = {}_1^2 \frac{{}_{a;1}^4}{{}_{a;0}^2 + {}_{a;1}^2} + {}_1^2 \frac{{}_{a;0}^2 \; {}_{a;1}^2}{{}_{a;0}^2 + {}_{a;1}^2} : $$

As the variance of news shocks, ${}_{a;1}^2$, approaches zero, this expression also approaches zero (term by term). From this we can conclude that a large contribution of news shocks is necessary but not su cient for there to be a large contribution of either future fundamental shocks or noise shocks.

One di erence relative to Proposition (4) is that Proposition (5) does not apply \over any frequency range." It only applies to *unconditional* variance decompositions; that is, to decompositions across all frequencies $2 [\ ;\ ]$. The distinction between past, present, and future makes sense in the time domain, but not in the frequency domain. Either we can look at the contribution of fundamentals over di erent time ranges or frequency ranges, but not both at the same time.

Finally, it is worth noting that the extent to which an endogenous process depends on future fundamental shocks depends on both the physical economic environment and agents' information structure. In equation (8), the weight of c_t on x_{t+1} depends both on $_1$ and . If the economic model is not su ciently \forward-looking," so $_1$! 0, then the share of future fundamentals will be small. Perhaps less intuitively, if ! 0 then the share of future fundamentals will also be small. Even if the model is forward-looking, so $_1 > 0$, future fundamental shocks can still be unimportant for current actions if the only information agents have about future fundamentals is completely contained in current and past fundamentals. Note that this is true even if the model is *purely*

Schmitt-Grohe and Uribe (2012), the model of news and animal spirits from Barsky and Sims (2012), and the model of noise shocks from Blanchard et al. (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-Grohe and Uribe (2012) take a pure news perspective while the Barsky and Sims (2012) and Blanchard et al. (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

4.1 Schmitt-Grohe and Uribe (2012)

The first model comes from Schmitt-Grohe 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 fluctuations due to beliefs and fundamentals. As a result, exactly what this model implies about the importance of pure 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 Online Appendix (B.1).

Each of the seven exogenous fundamentals follows a law of motion:

$$x_t = \rho_x x_{t-1} + \epsilon^a_{0,t} + \epsilon^a_{4,t-4} + \epsilon^a_{8,t-8}, \quad \begin{bmatrix} \epsilon^a_{0,t} \\ \epsilon^a_{4,t} \\ \epsilon^a_{8,t} \end{bmatrix} \overset{iid}{\sim} N\left(0, \begin{bmatrix} \sigma^2_{a,0} & 0 & 0 \\ 0 & \sigma^2_{a,4} & 0 \\ 0 & 0 & \sigma^2_{a,8} \end{bmatrix}\right). \tag{11}$$

where $0 < \rho_x < 1$. The model is estimated by likelihood-based methods on a sample of quarterly U.S. data from 1955:Q2 to 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, we would like to construct a noise representation that is observationally equivalent to representation (11). One such noise representation is in the following proposition.

Proposition 6. *The representation of fundamentals and beliefs in system* (11) *is observationally equivalent to the noise representation*

$$
\begin{aligned}
x_t &= \rho_x x_{t-1} + \varepsilon^x_t \\
s_{4,t} &= \varepsilon^x_{t+4} + v_{4,t} \\
s_{8,t} &= \varepsilon^x_{t+8} + v_{8,t},
\end{aligned}
\qquad
\begin{bmatrix} \varepsilon^x_t \\ v_{4,t} \\ v_{8,t} \end{bmatrix} \overset{iid}{\sim} N\left(0, \begin{bmatrix} \sigma^2_x & 0 & 0 \\ 0 & \sigma^2_{v,4} & 0 \\ 0 & 0 & \sigma^2_{v,8} \end{bmatrix}\right)
$$

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

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

We can use the noise representation in Proposition (6) 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 $\{\varepsilon^x_t\}$, and the fourth column shows the share attributable to the noise shocks $\{v_{4,t}\}$ and $\{v_{8,t}\}$ combined.

The main result is that nearly all of the variation in output, consumption, investment, and hours is due to 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 89% 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.

Variable	Surprise	News	Fundamental	Noise
Output	57	43	94	6
Consumption	50	50	95	5
Investment	55	45	89	11
Hours	16	84	97	3

Table 1: Variance decomposition (%) in the model of Schmitt-Grohe 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.

4.2 Barsky and Sims (2012)

The second model comes from Barsky and Sims (2012). It was constructed to determine whether measures of consumer con dence 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 con dence are mostly driven by news and not noise. Noise shocks account for negligible shares of the variation in forecast errors of consumption and output, while news shocks account for over half the process $\{x_t\}$ is assumed to follow a law of motion of the form:

$$\begin{aligned} x_t &= \mu_{t-1} + \varepsilon_t \\ \mu_t &= \rho\mu_{t-1} + \eta_t \\ s_t &= \mu_t + \nu_t \end{aligned} \qquad \begin{bmatrix} \varepsilon_t \\ \eta_t \\ \nu_t \end{bmatrix} \overset{iid}{\sim} N\left(0, \begin{bmatrix} \sigma_\varepsilon^2 & 0 & 0 \\ 0 & \sigma_\eta^2 & 0 \\ 0 & 0 & \sigma_\nu^2 \end{bmatrix}\right), \tag{12}$$

where $0 < \rho < 1$. Barsky and Sims (2012) refer to η_t as a news shock, ε_t as a surprise shock, and ν_t as a noise (animal spirits) shock.[8] However, these definitions 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 we refer to η_t as a news* shock, ε_t as a surprise* shock, and ν_t as 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 autoregressions. The vector autoregressions are estimated on quarterly U.S. data from 1960:Q1 to 2008:Q4. The time series used to estimate the vector autoregression 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 $\{\varepsilon_t\}$, the second shows the share attributable to news* shocks $\{\eta_t\}$, and the third shows the share attributable to noise* shocks $\{\nu_t\}$. Due to the presence of exogenous 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).

To properly isolate the independent contributions of beliefs, we would again like to construct a noise representation that is observationally equivalent to representation (12). The following proposition presents one such noise representation.

Proposition 7. *The representation of fundamentals and beliefs in system* (12) *is*

[8] While these authors refer to signal noise as "animal spirits," they also use the term "pure noise" to refer to statistical measurement error. We are only concerned with noise in the first sense.

observationally equivalent to the noise representation

$$
\begin{aligned}
x_t &= \frac{\ ^2}{\ ^2} m_t \quad \frac{1+\ ^2}{\ } m_{t-1} + m_{t-2} \\
m_t &= (\ +\ ) m_{t-1} \quad m_{t-2} + \ ^m_t \\
s_t &= m_t + v_t \\
v_t &= \ v_{t-1} + \ ^v_t \quad \ ^v_{t-1} \\
\begin{bmatrix} \ ^m_t \\ \ ^v_t \end{bmatrix} &\overset{iid}{\sim} N\left(0; \begin{bmatrix} \ ^4/(\ ^2) & 0 \\ 0 & \ ^2/ \end{bmatrix}\right);
\end{aligned}
$$

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

$$
= \frac{1}{2}\left(1 + \ ^2 + \frac{\ ^2}{\ ^2} \quad \left[\left(1 + \ ^2 + \frac{\ ^2}{\ ^2}\right)^2 \quad 4\ ^2\right]^{1/2}\right)
$$

$$
= \frac{1}{2}\left(1 + \ ^2 + \frac{\ ^2(\ ^2 + \ ^2)}{\ ^2\ ^2} \quad \left[4\left(1 + \ ^2 + \frac{\ ^2(\ ^2 + \ ^2)}{\ ^2\ ^2}\right)^2 \quad 4\ ^2\right]^{1/2}\right).
$$

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, and the fifth column shows the share attributable to productivity noise shocks. 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.

As in the model of Schmitt-Grohe and Uribe (2012), nearly all of the variation in output, consumption, investment, and hours is due to 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.

To further highlight the difference between noise and noise* shocks, we plot in Figure (1) both the noise and noise* shares of consumption for different values of the standard deviation of noise shocks, $\ $. The striking result is that the noise share

Variable	Surprise*	News*	Noise*	Fundamental	Noise
Output	53	37	0	89	1
Consumption	61	34	1	89	9
Investment	40	43	1	80	4
Hours	62	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.

of consumption is monotonically decreasing in . This means that removing noise* shocks altogether, by taking ! 0, actually leads to a *larger* noise share.

The intuition for this result is that the noise share of agents' forecasts (and their actions) is a hump-shaped function of the relative size of noise shocks. When noise shocks are very small, agents' signal is very precise, and noise shocks do not a ect their forecasts very much. At the other extreme, when noise shocks are very large, agents' signal is very imprecise, so they rationally ignore it. The maximum contribution of noise shocks occurs is achieved for an intermediate size of these shocks.

In this model, noise is generated explicitly by the noise* shocks f $_t g$, but also implicitly by the two shocks f $_t g$ and f $_t g$. The left panel of Figure (1) indicates

8e67 Td age of other non-technology fundamentalnon-tdamental52 Tfn-tpaTf 342.fn-tl400(q(tal)ttnd)-

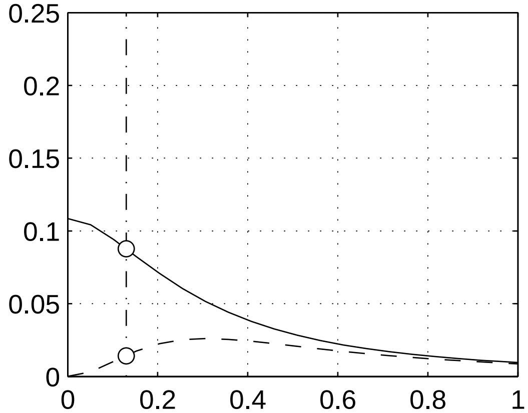

Figure 1: Noise versus noise*. This gure plots the noise and noise* shares of consumption over business cycle frequencies of 6 to 32 quarters, for di erent values of the variance of noise* shocks. The vertical dash-dotted line marks the estimated value of this parameter; the white circles correspond to the consumption noise shares reported in Tables (2) and (3). (The asterisk denotes the authors' original terminology.)

their paper is that noise shocks explain a sizable fraction of short-run consumption uctuations. However, it turns out that what the authors call \noise" shocks do not fully isolate uctuations due to temporary errors in agents' estimates. So we can investigate what this model implies about the importance of pure beliefs.

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

Agents only receive advance information about productivity, and not about the other ve fundamentals. So it is only pure beliefs about productivity that can play an independent role in driving uctuations. Let x_t denote the level of productivity, which is observed by agents in the economy, and let s_t denote the additional informative signal that agents receive. Then the processes fs_tg and fx_tg are assumed to evolve according to a system of the form

$$x_t = \;_t + \;_t$$

$$s_t = \;_t + \;_t$$

$$_t = \;\;\;_t$$

observationally equivalent to the noise representation

$$
\begin{aligned}
x_t &= \frac{-\rho}{(1-\rho)^2} m_{t+1} + \frac{(1+\rho^2)}{(1-\rho)^2} m_t - \frac{\rho}{(1-\rho)^2} m_{t-1} \\
s_t &= m_t + v_t \\
m_t &= (1+2\rho) m_{t-1} - \rho(2+\rho) m_{t-2} + \rho^2 m_{t-3} + \varepsilon^m_t \\
v_t &= 2\lambda v_{t-1} - \lambda^2 v_{t-2} + \varepsilon^v_t - (\lambda+\bar\lambda)\varepsilon^v_{t-1} + \lambda\bar\lambda\,\varepsilon^v_{t-2} \\
\begin{bmatrix} \varepsilon^m_t \\ \varepsilon^v_t \end{bmatrix} &\overset{iid}{\sim} N\left(0;\ \begin{bmatrix} (1-\rho)^2\sigma^2 & 0 \\ 0 & \rho^2\sigma^2/(\lambda\bar\lambda) \end{bmatrix}\right);
\end{aligned}
$$

with the convention that $s_{0,t} = x_t$, *and where*[11]

$$
\lambda = \frac{1}{2}\left(1 + \rho^2 + \rho^{1/2}\sqrt{\ }\, i - \left[\left(1 + \rho^2 + \rho^{1/2}\sqrt{\ }\, i\right)^2 - 4\rho^2\right]^{1/2}\right).
$$

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 and the fth column shows the share attributable to productivity noise shocks. 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 et al. (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-Grohe and Uribe (2012) and Barsky and Sims (2012) models, we nd that a sizable fraction of the variation in output, consumption, and hours worked can be attributed to noise shocks. For example, nearly 60%

[11] In the de nition of λ, $i \equiv \sqrt{-1}$ is the imaginary unit, and $\bar\lambda$ denotes the complex conjugate of λ. Both $\lambda + \bar\lambda$ and $\lambda\bar\lambda$ are real numbers.

of the variation in consumption is due to noise shocks. This is more than 10% larger than the share Blanchard et al. (2013) originally attributed to independent fluctuations in beliefs. A result of similar magnitude is true for output and hours worked. It is interesting that for all variables in the table, noise about productivity is in fact more important than productivity itself. This cannot be seen from the original decomposition.

Moreover, the right panel of Figure (1) indicates that, as in the model of Barsky and Sims (2012), the noise share of consumption is maximized when the size of noise* shocks is zero. This emphasizes the fact that variance decompositions in terms of noise* shocks can be a misleading measure of the importance of pure beliefs.

4.4 Future Fundamentals

Across all three of the models we consider, fundamental shocks appear to play a relatively large role. This is especially true in the models of Schmitt-Grohe and Uribe (2012) and Barsky and Sims (2012). Are fundamentals important because agents are correctly anticipating future fundamental changes before they occur, or because they are merely reacting to past fundamental changes? To answer this question, we can use the decomposition in equation (10) to compare the importance of current and past fundamental shocks relative to future fundamental shocks.

As we described in Section (3.4), it is only possible to consider decompositions in terms of past, present, and future fundamental shocks if the endogenous process under consideration is stationary. Each of the three models in this section exhibits trend growth in output, consumption, and investment. One option would be to first de-trend these processes using a frequency-domain filter (e.g. band-pass filter) and then perform the past versus future decomposition. However, this would not be a good idea, because frequency filters of this type scramble up the dependence across time periods. As a result, they can introduce spurious dynamic relationships that are not part of the underlying economic model.

Therefore, we propose to use a flexible exponential de-trending procedure that preserves the distinction between past and future shocks. For a difference-stationary process $\{y_t\}$, we define the stochastic trend $y_t(\ )$ to be an exponential moving average of past values,

$$y_t(\ ) = (1 \quad)y_{t-1} + \ y_{t-1}(\ );$$

where $\in [0,1)$. We then define the de-trended process $\{\tilde{y}_t(\ )\}$ as $\tilde{y}_t(\ )\ \ y_t\ \ y_t(\ )$.

The parameter controls the extent to which the trend depends on past values. When $= 0$, $\tilde{y}_t(\ ) = \ y_t$, so the de-trended process is the first-differenced version of the original process. As $\to 1$, $\tilde{y}_t(\ ) \to y_t$. By varying , we can therefore consider a range of different hypotheses regarding the stochastic trend. Because the filter is one-sided for any (unlike most frequency-domain filters), it preserves the notions of past, present, and future defined by the original process $\{y_t\}$.[12]

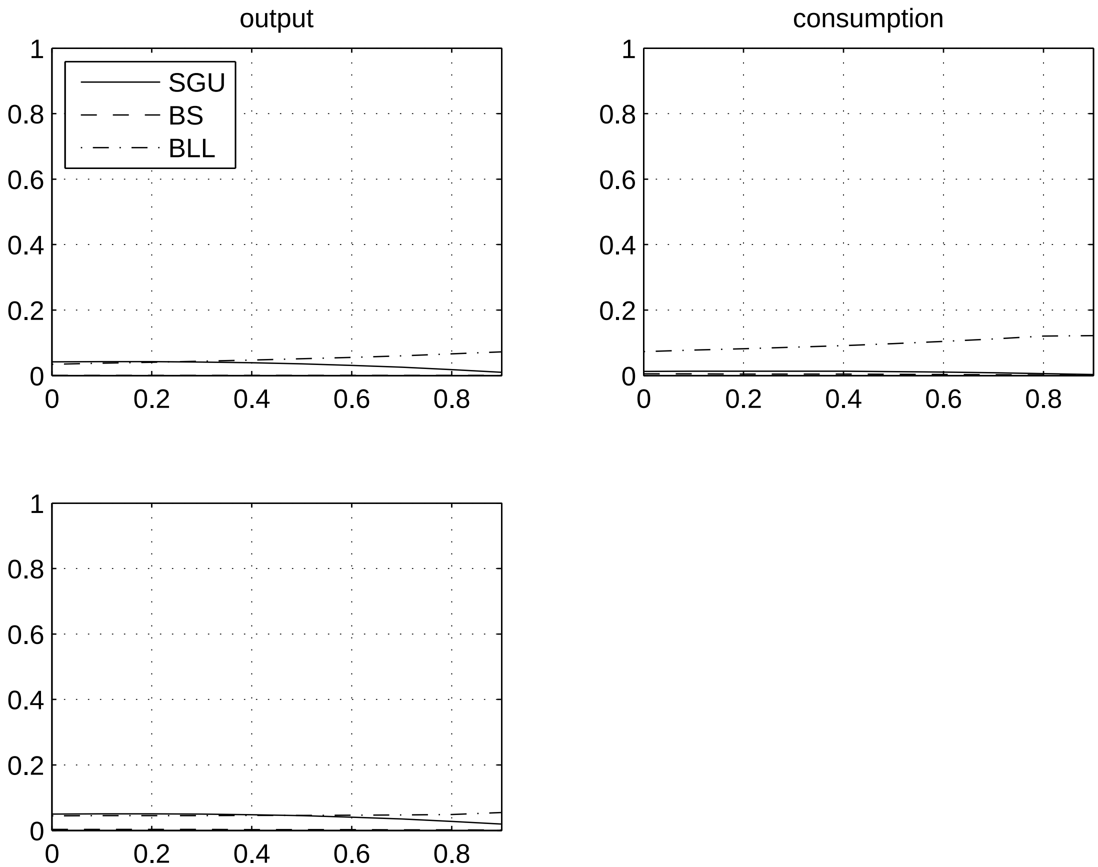

Figure 2: Fraction of the fundamental share due to future fundamental shocks, as a function of the de-trending parameter $\in [0,1)$. $= 0$ corresponds to a decomposition in (log) first differences, and $\to 1$ corresponds to a decomposition in (log) levels.

Figure (2) plots the fraction of the fundamental share due to future fundamental shocks, for each of the three models considered in this section. We plot this fraction

[12] In this respect, our proposal is similar to the procedure recently suggested by Hamilton (2017).

for a range of different de-trended versions of the endogenous variables, corresponding to a different values of . As in the previous decompositions in this section, we focus only on fundamentals about which agents receive some advance information. That means that for the models of Barsky and Sims (2012) and Blanchard et al. (2013), we focus only on productivity, while in the model of Schmitt-Grohe and Uribe (2012) we include all seven fundamentals.[13]

The consistent result across all three models is that the bulk of the contribution of fundamentals comes from current and past — not future — fundamental shocks. In some cases, it is difficult to see that there are actually three lines in each subplot. This is because one of the lines is visually indistinguishable from zero. In the model of Barsky and Sims (2012), endogenous variables are the least sensitive to future shocks (on average across), followed by the model of Schmitt-Grohe and Uribe (2012) and then Blanchard et al. (2013).

This result may seem surprising considering that news* shocks are fairly important in all three models. How can it be that news* shocks are so important, but future fundamental shocks are not? As discussed in Section (3.4), two conditions must be satisfied for future fundamental shocks to be an important driver of current actions. First, agents' actions must depend to a sufficient degree on their expectations of future fundamentals. Second, agents must receive signals that provide substantial information about future fundamentals, above and beyond what they can infer from current and past fundamentals.

While the different models deliver the same conclusion regarding the importance of future fundamentals, they do so for very different reasons. The model of Schmitt-Grohe and Uribe (2012) is not very forward-looking, so the first condition is not met. This can be seen in the forecast error variance decompositions from Figure (3), which report the share of news shocks in explaining the variance of forecast errors in various endogenous variables as a function of the forecast horizon. Most of the contribution of news shocks occurs only *after* the 4-quarter-ahead and 8-quarter-ahead news shocks actually materialize. This is the reason the news shares look like step functions with jumps just after 4 and 8 quarters.[14]

The models of Barsky and Sims (2012) and Blanchard et al. (2013) are more

[13] Agents only receive advance information about productivity in the first two models, so including other non-productivity fundamentals would only reduce the future fundamental share.

[14] This same observation is made by Sims (2016).

forward-looking, but as we will discuss in more detail in Section (4.5) below, agents' signals do not provide substantial information about future fundamentals above and

4.5 Understanding the Di erences

How is it that the three models we consider in this section, especially the rather similar models of Barsky and Sims (2012) and Blanchard et al. (2013), deliver such di erent results regarding the importance of noise shocks? The existing literature has o ered two separate explanations, one that emphasizes di erences in information structures and another that emphasizes di erences in physical economic environments. Beaudry and Portier (2014) argue that the key di erence is that agents in the model of Blanchard et al. (2013) face a more di cult inference problem, which leads them to make larger and more persistent forecast errors. By contrast, Barsky and Sims (2012) argue that the key di erence is that Blanchard et al. (2013) estimate a very accommodative monetary policy rule and a high degree of price rigidity, which work together to allow expectational shocks to propagate to the real side of the economy.

In this section we perform several exercises to better understand the reasons why these models disagree about the importance of noise shocks. We focus exclusively on the models of Barsky and Sims (2012) and Blanchard et al. (2013), since those are the most similar. We will argue that, at least with respect to these models, both the \right" information structure and the \right" physical environment are needed. Neither one alone is su cient to generate an large role for noise shocks.

First, we present in Figure (4) some prima facie evidence that the disagreement is not just due to di erences in information structure. If we replace the information structure in the Barsky and Sims (2012) model with the information structure from Blanchard et al. (2013), keeping all parameters at their original estimated values, the noise share of consumption does not change by much. This suggests that having the right information structure alone is not enough. However, having the right information structure is still important. If we replace the information structure in the Blanchard et al. (2013) model with the information structure from Barsky and Sims (2012), the noise share of consumption falls dramatically.

What is it about the information structure of Blanchard et al. (2013) that makes it amenable to a high consumption noise share? With this information structure, agents have to rely a good deal on their noisy signal in order to forecast future productivity. With the Barsky and Sims (2012) information structure, on the other hand, agents can forecast future productivity fairly well from the past history of productivity alone. As a result, they rely less on the noisy signal.

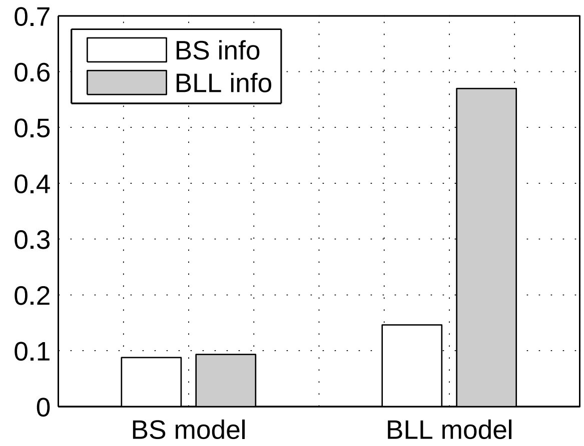

Figure 4: Swapping information structures. This gure plots the noise share of con-

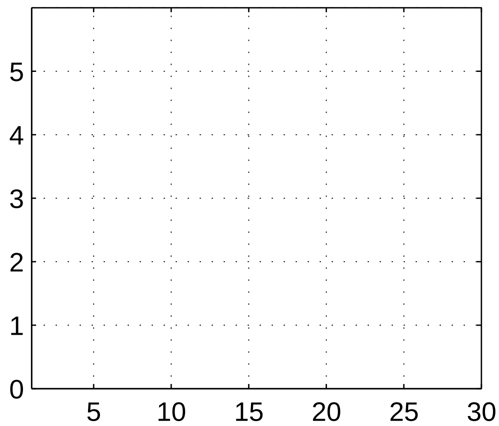

Figure 5: Comparing information structures. *Left*: standard deviation of j-quarter ahead productivity forecast errors, $a_{t+j} \quad E_t[a_{t+j}]$, with and without noisy signals. *Right*: fraction of j

consumption is 17%, and the bottom entry of the second column reports that in our version of the Blanchard et al. (2013) model, the noise share of consumption is 51%. This is close to what we found under the authors' original estimates.

both. When wages are flexible, the first effect dominates; households increase their wages enough that in equilibrium hours begin to fall. Since labor and capital are complementary, and there are investment adjustment costs, equilibrium investment falls on impact. When wages are sticky, however, the second effect dominates; households expect to be working more and therefore increase investment on impact. In either case, as time passes agents begin to learn that the shock was noise, and eventually reverse their actions and return back to the original steady state.

Lastly, we also report in Table (4) the Bayesian information criterion (BIC) values reverse

How can news and noise representations be observationally equivalent if it is only possible to use semi-structural methods to analyze models with news shocks and not models with noise shocks? The answer, as it turns out, is that \invertibility" is not a necessary condition for using these methods. What matters is not whether shocks can be recovered from the current and past history of observables, but simply whether shocks can be recovered from the observables. This weaker condition, which we refer

Beaudry, Paul and Franck Portier (2006) \Stock Prices, News, and Economic Fluctuations," *American Economic Review*, 96(4):1293{1307.

Beaudry, Paul and Franck Portier (2014) \News-Driven Business Cycles: Insights and Challenges," *Journal of Economic Literature*, 52(4):993{1074.

Benhabib, Jess, Pengfei Wang, and Yi Wen (2015) \Sentiments and Aggregate Demand Fluctuations," *Econometrica*, 83(2):549{585.

Blanchard, Olivier J., Jean-Paul L'Huillier, and Guido Lorenzoni (2013) \News, Noise, and Fluctuations: An Empirical Exploration," *American Economic Review*, 103(7):3045{3070.

Born, Benjamin, Alexandra Peter, and Johannes Pfeifer (2013) \Fiscal News and Macroeconomic Volatility," *Journal of Economic Dynamics and Control*, 37(12):2582{2601.

Chahrour, Ryan and Kyle Jurado (2017) \Recoverability," Working Paper, Duke University.

Christiano, Lawrence J., Cosmin Ilut, Roberto Motto, and Massimo Rostagno (2010) \Monetary policy and stock market booms," *Proceedings of the Jackson Hole Economic Policy Symposium*, 34(1):85{145.

Christiano, Lawrence J, Roberto Motto, and Massimo Rostagno (2014) \Risk shocks," *The American Economic Review*, 104(1):27{65.

Cochrane, John H. (1994) \Shocks," *Carnegie-Rochester Conference Series on Public Policy*, 41(1):295{364.

Forni, Mario, Luca Gambetti, Marco Lippi, and Luca Sala (2017) \Noisy News in Business Cycles," *American Economic Journal: Macroeconomics*, 9(4):122{52.

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

Hamilton, James D. (2017) \Why You Should Never Use the Hodrick-Prescott Filter," Forthcoming, *The Review of Economics and Statistics*.

Jaimovich, Nir and Sergio Rebelo (2009) \Can News about the Future Drive the Business Cycle?," *American Economic Review*, 99(4):1097{1118.

Jinnai, Ryo (2014) \R&D Shocks and News Shocks," *Journal of Money, Credit and Banking*, 46(7):1457{1478.

Kurmann, Andre and Christopher Otrok (2013) \News shocks and the slope of the term structure of interest rates," *The American Economic Review*, 103(6):2612{2632.

Leeper, Eric M, Todd B Walker, and Shu-Chun Susan Yang (2013) \Fiscal Foresight and Information Flows," *Econometrica*, 81(3):1115{1145.

Lorenzoni, Guido (2009) \A Theory of Demand Shocks," *American Economic Review*, 99(5):2050{84.

Lorenzoni, Guido (2011) \News and Aggregate Demand Shocks," *Annual Review of Economics*, 3(1):537{557.

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

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

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

Schmitt-Grohe, Stephanie and Mart n Uribe (2012) \What's News in Business Cycles," *Econometrica*, 80(6):2733{2764.

Sims, Eric (2016) \What's news in News? A cautionary note on using a variance decomposition to assess the quantitative importance of news shocks," *Journal of Economic Dynamics and Control*, 73(1):41{60.

Walker, Todd B. and Eric M. Leeper (2011) \Information ows and news driven business cycles," *Review of Economic Dynamics*, 14(1):55{71 Special issue: Sources of Business Cycles.

Appendix

Proof of Proposition

where $\psi_{i,j} \equiv \langle w_{i,t}, \varepsilon^a_{j,t} \rangle / \| \varepsilon^a_{j,t} \|^2$ is the projection coefficient. Define the index set I_a to be the set of indices $i \in \mathbb{Z}_+$ such that $\| \varepsilon^a_{i,t} \| > 0$. The collection of orthogonal shocks $\varepsilon^a_{i,t}$ with $i \in I_a$ 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 (14), x_t can be uniquely rewritten as:

$$x_t = \sum_{i=0}^{\infty} \sum_{j \leq i} \psi_{i,j} \varepsilon^a_{j,t-i} = \sum_{j \in I_a} \sum_{i=j}^{\infty} \psi_{i,j} \varepsilon^a_{j,t-i} = \sum_{j \in I_a} a_{j,t-j}.$$

The second equality rearranges the indexes on the double summation, and the third equality introduces the definition $a_{j,t-j} \equiv \sum_{i=j}^{\infty} \psi_{i,j} \varepsilon^a_{j,t-i}$. 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 the unique news representation when agents' date-t information set is $H_t(a)$.

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 $H_t(\hat{x})$: $\hat{x}_{i,t} = E[x_{t+i} | H_t(\hat{x})]$. By the uniqueness of orthogonal projections,

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

where $w_{i,t}$ was defined in equation (14). Therefore, $H_t(w) = H_t(\hat{x})$. But then because $H_t(a) = H_t(w)$ by construction, it follows that $H_t(a) = H_t(\hat{x})$. So expectations are indeed the same under both representations, $\hat{x}_{i,t} = E[x_{t+i} | H_t(\hat{x})] = E[x_{t+i} | H_t(a)]$, which completes the proof of the first part of the theorem.

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

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

Because $H(x) \subseteq H(a)$, there exist unique elements $m_{i,t} \in H(x)$ and $v_{i,t} \in H(s) \ominus H(x)$ such that $s_{i,t} = m_{i,t} + v_{i,t}$. This defines a noise representation when agents' date-t information set is $H_t(s)$. 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 $H_t(s) = H_t(a)$ by construction, and $H_t(a) = H_t(\hat{x})$ by the definition of a news representation, it follows that $H_t(s) = H_t(\hat{x})$ and therefore expectations are the same, $\hat{x}_{i,t} = E[x_{t+i} | H_t(\hat{x})] = E[x_{t+i} | H_t(s)]$. This completes the proof of the second part of the theorem. □

Appendix for Online Publication

A Proofs

Proof of Proposition (2). By rational expectations, H [illegible] that $f\hat{x}_{i;t}g$ forms a Gaussian system, it follows that ag[illegible] fully summarized by the random variables $\hat{x}_{i;}$ across all [illegible] t.

We can let $F_t(\hat{x})$ denote the smallest -algebra genera[illegible]se variables. That is, $F_t(\hat{x})$ is generated by cylinder sets of the form

$$A_t \quad f ! \ 2 \quad : \hat{x}_{i_1;t_1} \ 2 \ G_1; : : : ; \hat{x}_{i_n;t_n}$$ [illegible]

where denotes the space of elementary events, $G_1; : : :$ [illegible]bitrary Borel sets in $\mathbb{R}$, the indices $t_1; : : : ; t_n$ assume values in the set $f \quad 2$ [illegible] and the indices $i_1; : : : ; i_n$ assume values in $\mathbb{Z}$. By construction, the seque[illegible] uniquely determined by the forecasts $f\hat{x}_{i;t}g$. If two repr[illegible] and beliefs imply the same dynamics for $f\hat{x}_{i;t}g$, they i[illegible] same information structure $fF_t(\hat{x})g$. Therefore, the conditional distributio[illegible] of any stochastic process fc_tg, such that c_t is measurable with respect to [illegible]each $t \ 2 \ \mathbb{Z}$, is also the same. □

Proof of Proposition (3). As in the proof of Proposi[illegible]ve can equate the spectral density of fd_tg with $d_t \quad (x_t; \hat{x}_t)^0$ under each rep[illegible]n. In this case,

$$f_d(\ ) = \frac{1}{2} \begin{bmatrix} {}^2 + {}^2 & \frac{4}{2+2} e^{i} \\ \frac{4}{2+2} e^{i} & \frac{4}{2+2} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} [illegible] \end{bmatrix}$$

=2.958 0 Td J/F48/F17 1tion t1q 1 0 0 1 450 0 0J/F49 [illegible]
t Borel22 Tf 4]TJ/F48 [2.9701 Tf 6.052 –]1.795 Td [(t)] [illegible] [(...)]T0/F17 11

system (6) noise

This equality holds if and only if the relations +

By the endogeneity of $\{c_t\}$ and the rationality of expectations, $c_t \in H(s)$ for all $t \in \mathbb{Z}$. Combining this with Equation (15), it follows that for each c_t, there exist two unique elements $a_t \in H(x)$ and $b_t \in H(\nu)$ such that

$$c_t = a_t + b_t. \tag{16}$$

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 $\underline{\omega} < \omega < \bar{\omega}$, which is equal to

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

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

Proof of Proposition (5). Beginning with the decomposition of $H(s)$ in equation (15), we can further decompose $H(x)$ uniquely into the sum of subspaces $D_t(x) \equiv H_t(x) \ominus H_{t-1}(x)$,

$$H(s) = \left(\bigoplus_{j=-\infty}^{\infty} D_{t-j}(x) \right) \oplus H(\nu).$$

By definition, each fundamental shock $\varepsilon_t^x \equiv x_t - E[x_t|H_{t-1}(x)]$ forms a basis in the space $D_t(x)$. Since $c_t \in H(s)$ for all $t \in \mathbb{Z}$, it follows that for each c_t, there exists a unique sequence of projection coefficients $\{\psi_j\}$ such that

$$c_t = \sum_{j=-\infty}^{\infty} \psi_j \varepsilon_{t-j}^x + b_t,$$

where $\psi_j \equiv E[c_t \varepsilon_{t-j}^x]/\mathrm{var}[\varepsilon_t^x]$ and $b_t \perp H(x)$. The shares of the variance of $\{c_t\}$ due to past, present, and future fundamental shocks are therefore uniquely determined, and are given by

$$\underbrace{\frac{\sum_{j=1}^{\infty} \psi_j^2 \mathrm{var}[\varepsilon_t^x]}{\mathrm{var}[c_t]}}_{\text{past}}, \quad \underbrace{\frac{\psi_0^2 \mathrm{var}[\varepsilon_t^x]}{\mathrm{var}[c_t]}}_{\text{present}}, \quad \text{and} \quad \underbrace{\frac{\sum_{j=-\infty}^{-1} \psi_j^2 \mathrm{var}[\varepsilon_t^x]}{\mathrm{var}[c_t]}}_{\text{future}}.$$

□

Proof of Corollary (1). Consider an arbitrary noise representation of fundamentals and beliefs, and an endogenous process $\{c_t\}$. By the rationality of expectations, agents' best forecast of c_{t+h} as of date t is equal to

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

Therefore, $\hat{c}_{h,t} \in H_t(\hat{x})$. This means that the forecast error $w_t^h \equiv c_t - \hat{c}_{h,t-h}$ also satisfies $w_{h,t} \in H_t(\hat{x})$. 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. □

Lemma 1. *Any news representation in which each process $\{a_{i,t}\}$ is i.i.d. over time is observationally equivalent to a noise representation with $x_t \overset{iid}{\sim} N(0, \sigma_x^2)$ and*

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

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

$$\sigma_x^2 = \sum_{i \in I_s} \sigma_{a,i}^2 \quad \text{and} \quad \sigma_{v,i}^2 = \frac{1}{\sigma_{a,i}^2}\left(\sum_{j<i} \sigma_{a,j}^2\right)\left(\sum_{j \geq i} \sigma_{a,j}^2\right) \quad \text{for all } i \in I_s.$$

Proof of Lemma (1). 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 joint spectral density of any two forecast processes $\{\hat{x}_{j,t}\}$ and $\{\hat{x}_{k,t}\}$ for $j,k \in \mathbb{Z}_+$ is equal to

$$f_{j,k}(\omega) = \frac{1}{2\pi} \sum_{m \in M} \sigma_{a,m}^2 e^{-i\omega(k-j)}, \tag{17}$$

where M is defined as the set of indices $m \in I_a$ such that $m \geq |k-j| + j$. In a noise representation of the type described in the proposition, the joint spectral density of any two forecast processes $\{\hat{x}_{j,t}\}$ and $\{\hat{x}_{k,t}\}$ for $j,k \in \mathbb{Z}_+$ is equal to

$$f_{0,0}(\omega) = \frac{1}{2\pi} \sigma_x^2 \tag{18}$$

586 0 Td [(i)]T

Proof of Proposition (6). Define the composite shock

$$\varepsilon^x_t \equiv \varepsilon^a_{0,t} + \varepsilon^a_{4,t-4} + \varepsilon^a_{8,t-8}. \tag{19}$$

The process $\{\varepsilon^x_t\}$ is i.i.d. because $\{\varepsilon^a_{i,t}\}$ is i.i.d. for each $i \in I_a \equiv \{0, 4, 8\}$. agents' date-$t$ information set in representation (11) is $H_t(\varepsilon^a)$. But based on this information set, equation (19) defines a news representation for $\{\varepsilon^x_t\}$ with i.i.d. news processes. Therefore, we can apply Lemma (1) to the composite shock process, which gives the relations stated in the proposition. □

where $| \ | < 1$ is equal to the expression stated in the proposition. Because $H_t(s)$ is unchanged from representation (12) for all $t \in \mathbb{Z}$, it follows that $\hat{x}_{j;t} \equiv E[x_{t+j}|H_t(s)]$ is also unchanged for any $j \in \mathbb{Z}$. Therefore these two representations are observationally equivalent. $\square$

By writing out the corresponding law of motion for $\{\tilde{m}_t(\delta)\}$ and then taking limits as δ approaches one from below, we obtain the law of motion for $\{m_t\}$ stated in the proposition. In a similar manner, we can obtain the law of motion for $\{x_t\}$ in terms of $\{m_t\}$ by using the spectral characteristic $\varphi(\cdot)^{-1}$. Finally, the definition of the noise process $\{\tilde{v}_t(\delta)\}$ in equation (23) implies that

$$f_v(\lambda;\delta) = \frac{1}{2\pi}\frac{\sigma^2\sigma_v^2}{|\theta|^2}\left|\frac{(1-\delta e^{-i\lambda})(1-\theta e^{-i\lambda})(1-\rho e^{-i\lambda})}{(1-\delta e^{-i\lambda})(1-\theta e^{-i\lambda})^2}\right|^2;$$

where $|\theta| < 1$ is equal to the expression stated in the proposition. By letting δ tend to one from below, we obtain the law of motion for $\{v_t\}$. Because $H_t(\tilde{s})$ is unchanged from representation (13) for each $\delta \in [0, 1)$ and all $t \in \mathbb{Z}$, it follows that

$$\hat{x}_{j,t} \equiv \lim_{\delta \to 1} E_t[\tilde{x}_{t+j}(\delta)|H_t(\tilde{s})]$$

is also unchanged for any $j \in \mathbb{Z}$. Therefore these two representations are observationally equivalent. □

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-Grohe 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,

$$E\left[\sum_{t=0}^{\infty}\beta^t\zeta_t\frac{(C_t - bC_{t-1} - \psi h_t^{\theta} S_t)^{1-\sigma} - 1}{1-\sigma}\right];$$

subject to a standard sequence of constraints,

$$S_t = (C_t - bC_{t-1})^{\gamma} S_{t-1}^{1-\gamma}$$

$$C_t + A_t I_t + G_t = \frac{W_t}{\mu_t}h_t + r_t u_t K_t + P_t$$

$$K_{t+1} = (1-\delta(u_t))K_t + z_t^I I_t\left[1 - \mathcal{S}\left(\frac{I_t}{I_{t-1}}\right)\right]$$

Relative to the standard real business cycle model, this model features investment adjustment costs (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 (u_t); one period internal habit formation in consumption, controlled by $0 < b < 1$; a potentially low wealth e ect on labor supply, when $0 < \quad < 1$ approaches its lower limit; and monopolistic labor unions, which e ectively reduce the wage rate by an amount $_t$ each period but rebate pro ts lump sum to the household through P_t.

Output is produced by a representative rm, which combines capital K_t, labor h_t, and a xed factor of production L

Intermediate goods rms are monopolistically competitive and take the demands of nal goods rms as given. They each have a production function of the form $Y_t(j) = A_t K_t(j) \; N_t(j)^{1}$. Each intermediate rm 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 .

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

$$Y_t^k(\;) = \quad \frac{I_t(\;)}{K_t(\;)} \quad K_t(\;);$$

where is an increasing and concave function. The aggregate capital stock evolves according to $K_t = \; (I_t{=}K_t)K_{t\;1} + (1 \quad)K_{t\;1}$, where $0 < \; < 1$ is the depreciation rate. The aggregate resource constraint is $Y_t = C_t + I_t + G_t$ (ignoring resources lost due

P_t is the price level, T_t is a lump sum tax, R

while the nominal interest rate is given by the e ective federal funds rate. Data were downloaded from the St. Louis Federal Reserve Database, FRED, on October 25,

	Economic parameters	BS info	BLL info
	habit	0.3145	0.0252
	Frisch elasticity	4.9976*	4.9999*
	capital adj. cost	5.2093	3.4670
	Calvo price	0.9420	0.9309
	Taylor in ation	4.8073	4.8897
y	Taylor output growth	0.0042	0.0484
i	interest smoothing	0.5072	0.4074
i	s.d. policy shock	0.1343	0.1629
"i	autocorr. policy	0.9989*	0.9892
	BS info parameters		
	autocorr. growth	0.9231	
	s.d. growth shock	0.2190	
	s.d. surprise* shock	0.8716	
	s.d. noise* shock	0.0001*	
	BLL info parameters		
	autocorr. growth		0.8581
	s.d. growth shock		1.3638
	s.d. noise* shock		0.0010

Table 5: Estimated parameters for alternative versions of the Barsky and Sims (2012) model.

Economic parameters		BS info	BLL info	BS info + ex wage	BLL info + ex wage
h	habit	0.8145	0.7066	0.6209	0.4726
	inverse Frisch elasticity	0.2000*	0.2000*	0.2000*	0.2000*
	cap. util. cost	0.5023	0.0079	0.4628	0.0010*
	inv. adj. cost	15.0000*	15.0000*	15.0000*	15.0000*
$_p$	Calvo price	0.8771	0.8645	0.8654	0.8929*
$_w$	Calvo wage	0.9013	0.8708	-	-
	Taylor in ation	4.2259	3.8640	1.0100*	1.0100*
$_y$	Taylor output gap	0.0010*	0.0010*	0.4742	0.4135
$_r$	interest smoothing	0.4686	0.4540	0.2813	0.1127
$_q$	s.d. policy shock	0.3394	0.2792	0.3089	0.4112
$_q$	autocorr. policy	0.9990*	0.9990*	0.9425	0.9481
BS info parameters					
	autocorr. growth	0.9166		0.8980	
	s.d. growth shock	0.2553		0.4430	
	s.d. surprise* shock	0.9762		1.2358	
	s.d. noise* shock	0.0001*		0.0001*	
BLL info parameters					
	autocorr. growth		0.8911		0.8068
	s.d. growth shock		1.3025		1.8876
	s.d. noise* shock		0.0001*		0.0001*

Table 6: Estimated parameters for alternative versions of the Blanchard et al. (2013) model.