MAN-BITES-DOG BUSINESS CYCLES

KRISTOFFER P. NIMARK

Abstract. The newsworthiness of an event is partly determined by how unusual it is and this paper investigates the business cycle implications of this fact. We present a tractable model that features an information structure in which some types of signals are more likely to be observed after unusual events. The proposed information structure can help us understand why we observe (i) large changes in macro economic aggregate variables without a correspondingly large change in underlying fundamentals (ii) persistent periods of high macroeconomic volatility and (iii) a positive correlation between absolute changes in macro variables and the cross-sectional dispersion of expectations as measured by survey data. These results are consequences of optimal updating by agents when the availability of some signals is positively correlated with tail-events. The model is estimated by likelihood based methods using raw survey data and a quarterly time series of total factor productivity along with standard aggregate time series. The estimated model suggests that there have been episodes in recent US history when the impact on output of innovations to productivity of a given magnitude were up to twice as large compared to normal times.

1. Introduction

A well-known journalistic dictum states that "*dog-bites-man* is not news, but *man-bites-dog* is news". That is, unusual events are more likely to be considered newsworthy than events that are commonplace. This paper investigates the business cycle implications of this aspect of news reporting. Particularly, we will demonstrate that a single and relatively simple mechanism can help us understand three features of business cycles. First, there can be large changes in aggregate variables like CPI inflation and GDP growth, but without an easily identifiable change in fundamentals of comparable magnitude. Second, there appear to be persistent episodes of increased macroeconomic volatility in the data. Third, measures of uncertainty as well as measures of cross-sectional dispersion of expectations are positively

Date: December 12, 2011 The author is grateful for comments and suggestions from Tobias Broer, Cosmin Ilut, Jarkko Jaaskela, Leonardo Melosi, Emre Ozdenoren, Pontus Rendahl, Martin Schneider, Jaume Ventura, Mirko Wiederholt, Jacob Wong, participants at the LBS workshop on *The Macroeconomics of Incomplete Information: Empirical and Policy Perspectives*, the LAEF conference at UC Santa Barbara on *Putting Information into (or taking it out of) Macro Economics*, ESSIM 2011, the Riksbank conference on *Beliefs and Business Cycles*, ASSET 2011, seminars at the University of Adelaide, Reserve Bank of Australia, University of Sydney and University of Cambridge. Financial support from Ministerio de Ciencia e Innovacion (ECO2008-01665), Generalitat de Catalunya (2009SGR1157), Barcelona GSE Research Network and the Government of Catalonia is gratefully acknowledged.
Address: CREI, Universitat Pompeu Fabra, Ramon Trias Fargas 25-27, Barcelona 08005.
e-mail: knimark@crei.cat *web*: www.kris-nimark.net.

correlated with absolute magnitudes of changes in macro economic aggregates. These features can be explained by Bayesian agents optimally updating to signals that are more likely to be available about unusual events.

Conceptually, information about the current state of the world can be divided into at least three categories. What we may call *local information* is information that agents observe directly through their interactions in markets, e.g. through buying and selling goods or through participating in the labor market. But there is also information that is collected and summarized by organizations and made available to a broader public. One such type of information is what we may call *statistics*. Statistics are often collected by government agencies and are normally reported regardless of the realized values of the variable that they refer to. Another source of information is the news media, such as newspapers and television, which may be the main source of information for a large section of the general population. Due to constraints on the number of pages or the time available for broadcasting, news organizations need to select which news stories to report. The man-bites-dog dictum referred to above suggests that more unusual events are more likely to be selected for dissemination and are thus more likely to become *news*. In order to have a terminology that is distinct from the one used by the literature studying how information about future productivity affect the economy today (e.g. Beaudry and Portier 2006 and Jaimovich and Rebelo 2009) "news" in the sense meant here will be referred to as man-bites-dog signals.

A prime example of man-bites-dog news reporting is the *Movers* segment on Bloomberg Television. In a typical segment, the price movements of a few stocks are reported along with short statements on the probable causes of these movements. The stocks in question are a small sub-sample of all stocks traded and are selected on the basis of having had the largest price movements during the day. Unusual price movements are thus more likely to be reported than more common price movements.

That some types of signals are more likely to be available about unusual events may suggest that we should be better informed about unusual events. However, the flip-side of this argument is that if there is a man-bites-dog signal available, the probability of an unusual event has increased. In fact, and as will be demonstrated below, observing a man-bites-dog signal can actually increase agents' uncertainty about an event. This may at first seem counter intuitive, but consider again the example of the *Movers* segment. If you were told that a certain stock had been mentioned in the *Movers* segment, but not told if its price had risen or fallen, how would you revise your beliefs about the price of the stock in question? Arguably, your conditional variance of the stock price should increase. Stated differently, the fact that the stock was mentioned in the *Movers* section tells you something about the second moment of the distribution of the stock's price. After observing the actual price change, your conditional variance decreases (to zero in this case). This example illustrates the two effects at work: Conditioning only on the *availability* of a man-bites-dog signal increases uncertainty, while observing the actual contents of the signal decreases the uncertainty. Below we show in a tractable setting that when signals are noisy, the former effect can dominate so that the posterior uncertainty after observing a man-bites-dog signal is larger than when no such signal is observed.

In a dynamic setting, a larger posterior uncertainty in period t translates into a larger prior uncertainty in period $t+1$. By embedding a man-bites-dog information structure in a simple

business cycle model similar to that of Lorenzoni (2009), we show that the propagation of uncertainty through time endogenously generates periods of persistently higher volatility in output and inflation. The mechanism is the following. Agents in the model need to solve a dynamic filtering problem in order to make optimal consumption and price setting decisions. In a given period, the weight agents put on new information is inversely related to the precision of their priors. Since a man-bites-dog signals in period t can increase the prior uncertainty in period $t+1$, agents may put more weight on all signals in period $t+1$ relative to the case when there was no man-bites-dog episode in period t. The increased sensitivity to new information can persist for several periods and implies that the impact of an exogenous disturbance of a given magnitude can also be larger than usual for several periods. This mechanism can endogenously generate periods of higher volatility of macro economic aggregates as observed in the data and documented by Engle (1982), Stock and Watson (2003), Primiceri (2005) and Fernandez-Villaverde and Rubio-Ramirez (2010) among others.

Bloom (2009) presents a model in which firms respond to increased uncertainty by adopting a wait-and-see approach to capital investment and recruiting so that an increase in the second moment of the exogenous shocks generates a fall in output. The paper by Bloom thus provides a story for how firms respond to increased uncertainty and where uncertainty has a direct effect on the level of output. The man-bites-dog mechanism provides a story for how economic agents come to understand that conditional uncertainty has increased. In the model presented here, a large shock in levels is more likely to generate a man-bites-dog signal and a large shock in levels is thus more likely to lead to an increase in the conditional uncertainty. In related work, Bachmann and Moscarini (2011) propose an alternative mechanism that also implies that causality runs from realizations of first moment shocks to conditional uncertainty. In their paper, imperfectly informed firms tend to be more likely to experiment with prices after a large negative shock. This behavior allows firms to get a better estimate of the demand elasticity of the good that they are producing, but increases the conditional variance of profits in the short run.

In the business cycle model presented below, man-bites-dog signals are public signals. This means that information about unusual events will be more correlated across agents than information about common events. For a signal to be public in the strong common knowledge sense of the word, it is not enough that a signal is publicly available. Instead, the signal must not only be observed by everybody, but the fact the everybody observes the signal must also be common knowledge. The concept of a public signal in this sense is thus much stronger than the everyday meaning of the term "public information". Arguably, while the theoretical dichotomy between purely private or purely public signals is somewhat artificial, information about unusual events are more likely to be closer to the theoretical ideal of a public signal than information about more commonplace events. For instance, in a segmented news market where different news outlets cater to partially different audiences, there are some events that virtually all news organizations will report. Though there are many aspects of an event that determines how newsworthy it is judged to be, everything else equal, the more unusual an event is, the larger is the number of news organizations that will consider it newsworthy.

As mentioned above, observing a man-bites-dog signal can increase the posterior uncertainty relative to the case when no man-bites-dog signal is observed. The same parameter restrictions that ensure that posterior variances increases after a man-bites-dog signal also imply that the cross-sectional dispersion of expectations increases. This holds even if the man-bites-dog signal is public (in the strong common knowledge sense of the word). In the data, we observe a positive correlation between the cross-sectional dispersion of forecasts (as measured by the Survey of Professional Forecaster) and the absolute magnitudes of changes in macro aggregates. Interpreted through the lens of the model, this suggest that the empirically relevant specification of the model may be one where the increase in uncertainty from conditioning on the availability of a man-bites-dog signal is dominating the increased precision due to the content of the signal.

In order to quantify the importance of the the man-bites-dog aspect of news reporting I estimate the model on US data. In addition to standard macro variables like GDP, CPI inflation and the Federal Funds rate, I also use the quarterly time series of Total Factor Productivity constructed by John Fernald (2010) as well as raw survey data from the *Survey of Professional Forecasters*. Using raw survey data, i.e. the cross-section of individual survey responses rather than a mean or median response, has at least two advantages. First, and as documented by Mankiw, Reis and Wolfers (2004) and Swanson (2006a) there is significant time variation in the dispersion of forecasts reported by survey respondents. Since the model can potentially fit this fact, raw survey data can be exploited when estimating the model allowing for a sharper inference about the precision of signals observed by agents. The second advantage of using raw survey data rather than a median or mean expectation stems from the fact that the number of survey respondents varies over time. For instance, the number of respondents forecasting nominal GDP growth and CPI inflation varies between 9 (1990:Q2) and 50 (2005:Q4) in a sample that covers the period 1981:Q3 to 2010Q4. Using raw survey data and likelihood based estimation methods naturally incorporates that we have a (presumably) more representative sample of the population with 50 observations than with 9.

Most existing macro models imply that the dispersion of cross-sectional expectations is either zero, as in the full information rational expectations models, or non-zero but constant as in models with private but time-invariant information structures, e.g. Lorenzoni (2009), Mackowiak and Wiederholt (2009), Graham and Wright (2009), Nimark (2008, 2010) Angeletos and La'O (2009, 2010) or Melosi (2011). One exception is the sticky information models of Mankiw and Reis (2002) and Reis (2006a, 2006b). In sticky information models, only a fraction of agents update their information in each period and those who update, all observe the state perfectly. This results in a cross-sectional distribution of expectations that is a mixture of a degenerate and a dispersed distribution. It is because the cross-sectional dispersion of expectations in the model presented here is time varying but continuous that it it possible to estimate the structural parameters of the model using likelihood based methods and "raw" survey data. One methodological contribution of the paper is to demonstrate how dynamic models with a time varying information structure can be solved and estimated. As far as I know, this is the first paper to do so and this may be of independent interest to some readers.

The paper is structured as follows. In the next section, the concept of man-bites-dog signals is introduced formally. In the static setting of that section, many results can be derived analytically and it is demonstrated that the implications of man-bites-dog signals crucially depend on two parameters (i) the parameter governing how unusual an event has to be to significantly increase the probability of observing a signal (ii) the variance of the noise in the generated signal. The larger either of these parameters are, the more likely it is that the economy will display the features described here in the introduction. Section 3 presents a simple business cycle model similar to that of Lorenzoni (2009), but with a man-bites-dog type of information structure. Section 4 discusses how the model is solved and how the parameters are estimated. Section 5 contains the main empirical results of the paper. Section 6 briefly discusses two alternative specifications of the model and Section 7 concludes.

2. Signals and unusual events

This section presents a simple and tractable framework in which some signals are more likely to be available when something unusual has happened. There are no economic decisions made by agents beyond forming an expectation about a latent variable. The purpose of the section is to build intuition for how the filtering problems of Bayesian agents change when the availability of some signals is positively correlated with tail events. These insights also carry over to the business cycle model presented in a later section.

2.1. **Reverse engineering the conditional probability of observing a signal.** Agents indexed by $j \in (0,1)$ want to form an expectation about the latent variable x. Its unconditional probability density function is denoted $p(x)$ and x has zero mean and unconditional variance σ_x^2. The indicator variable S takes the value 1 when the signal y about x is available and zero otherwise. We want to find a specification such that the conditional probability of observing a signal is increasing in the distance of the realized value of x from its mean. That is, we want to find a joint distribution of x and S such that the inequality

$$\frac{\partial p\left(S=1 \mid x\right)}{\partial \left|x\right|} \geq 0 \tag{2.1}$$

holds. The main idea is illustrated in Figure 1. There, the unconditional distribution $p(x)$ of x is plotted (solid line) together with the probability $p(S=1 \mid x)$ of observing the signal y conditional on the realization of x. At the mean, there is approximately a 40% chance of observing the signal y. As realizations of x further from the mean are considered, the conditional probability of observing a signal increases towards 1 so that a signal is generated almost surely when the realization of x is far enough away from the mean.

It turns out to be easier to maintain tractability by instead of directly specifying the conditional probability of observing a signal $p\left(S=1 \mid x\right)$ in the numerator of (2.1) to instead specify the distribution of x conditional on $S=1$. By Bayes' rule, we have that the conditional probability of observing a signal is given by

$$p\left(S=1 \mid x\right) = \frac{p\left(x \mid S=1\right)}{p\left(x\right)} p\left(S=1\right). \tag{2.2}$$

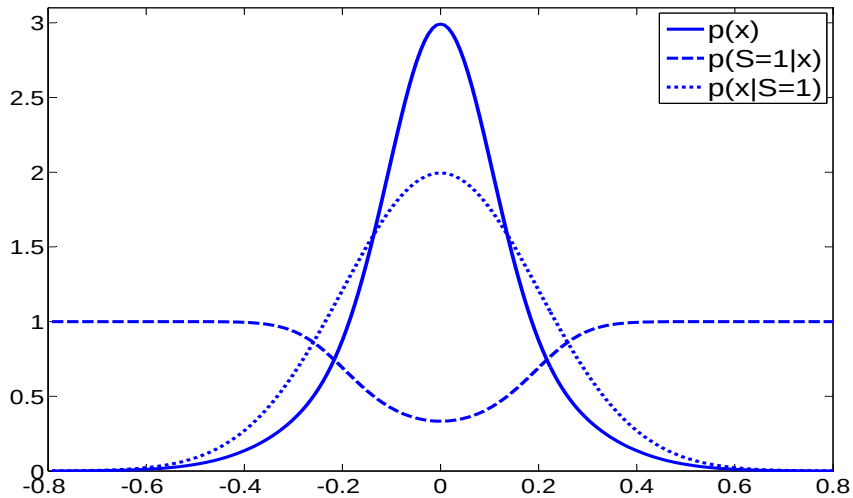

FIGURE 1. Unconditional distribution of x and conditional probability of observing the signal y.

Since $p(x)$ is uniformly decreasing in $|x|$ for any symmetric unimodal distribution $p(x)$, we can ensure that the inequality (2.1) holds by choosing a conditional distribution $p(x \mid S=1)$ that is decreasing at a slower rate than $p(x)$ in $|x|$. For unimodal symmetric distributions $p(x)$ and $p(x \mid S=1)$ we thus need to choose a $p(x \mid S=1)$ that is more dispersed than $p(x)$.

The link between the conditional and unconditional distributions of x and the conditional probability of observing a signal is perhaps easiest to illustrate in a graph. In Figure 1 the dotted line is the probability of x conditional on $S=1$. The relationship (2.2) implies that the conditional probability of observing the signal y is the ratio of the dotted line and the solid line times the normalizing constant $p(S=1)$. To ensure that the probability of observing a signal increases as values of x further from its mean are considered, we thus need to choose a conditional probability $p(x \mid S=1)$ that decreases at a slower rate than $p(x)$ in $|x|$. Visually, this is obvious if the more peaked distribution (solid curve) is compared to the flatter distribution (dotted curve). At the points where the solid and the dotted curves cross, the ratio on the right hand side of (2.2) is one, implying that the unconditional probability of observing a signal in this example can be read off the dashed line on a straight line downward from the point of the crossing curves. In the particular example used to construct Figure 1, $p(S=1)$ is 50%.

One specification that delivers a conditional probability of observing the signal y that is increasing in $|x|$ and maintains tractability of agents' filtering problem is to let x be distributed as the mixture normal

$$x \sim (1-\omega) N\left(0, \sigma^2\right) + \omega N\left(0, \gamma\sigma^2\right) \tag{2.3}$$

where ω is the unconditional probability of observing y, i.e. $\omega = p(S = 1)$ so that

$$p(x \mid S = 0) = N\left(\mu_x, \sigma^2\right) \tag{2.4}$$

$$p(x \mid S = 1) = N\left(\mu_x, \gamma\sigma^2\right) \tag{2.5}$$

The parameter ω determines how unusual the outcomes associated with the signal y are in the unconditional sense. The parameter γ determines how the conditional probability of observing a public signal depends the realizations of x. Setting $\gamma > 1$ ensures that the conditional probability $p(S = 1 \mid x)$ is increasing in the deviation of x from its mean, i.e. that the inequality (2.1) holds. Intuitively, the inequality (2.1) will hold since a more dispersed distribution is "flatter" at all points of the support except at the mean, than a more concentrated distribution. Note that this argument works in both directions: Bayes' rule implies that *any* conditional distribution of the probability of observing a signal that is uniformly increasing in $|x|$ must be associated with a conditional distribution $p(x \mid S = 1)$ that has relatively more probability mass further out in the tails compared to the unconditional distribution of x.[1]

The advantage of modeling the information structure this way is that since the conditional distributions (2.4) and (2.5) are Gaussian, the expectations of x conditional on S are known in closed form. In the set up of this paper, agents will never need to evaluate the unconditional distribution $p(x)$ since whether S equals 1 or 0 will always be known to agents. So while it is more tractable to specify the conditional distributions $p(x \mid S = 0)$ and $p(x \mid S = 1)$ directly, we may still want to treat the unconditional distribution of x as the primitive, perhaps because it corresponds to a particular data moment. For a given unconditional variance σ_x^2 and parameters ω and γ we can back out the implied σ^2 by using that the variance of the mixture normal distribution (2.3) is given by

$$\sigma_x^2 = \omega\gamma\sigma^2 + (1 - \omega)\sigma^2 \tag{2.6}$$

so that

$$\sigma^2 = \frac{\sigma_x^2}{\omega\gamma + (1 - \omega)} \tag{2.7}$$

For a given unconditional variance σ_x^2 we can thus adjust the slope and level of the dashed line $p(S = 1 \mid x)$ by choosing different values of ω and γ while holding the variance of the unconditional distribution $p(x)$ fixed by letting σ^2 be determined by (2.7).[2]

2.2. **The filtering problem.** Agents indexed by j form an estimate of x conditional on all available information. When $S = 0$, agent j observes only a private signal x_j which is the sum of the true x plus an idiosyncratic noise term

$$x_j = x + \varepsilon_j : \varepsilon_j \sim N\left(0, \sigma_\varepsilon^2\right) \ \forall \ j. \tag{2.8}$$

[1]In the Appendix it is shown that a conditional distribution with more probability mass in the tails also can be generated endogenously from an unconditionally normal distribution of x and a directly specified man-bites-dog signal technology. I.e. instead of specifying $p(x \mid S)$ and $p(x)$ as is done here, we specify $p(x)$ and simple man-bites-dog selection mechanism. The conditional distribution $p(x \mid S)$ can then be derived endogenously and it is shown that it has more mass in the tails compared to $p(x)$. However, the expressions for the conditional expectations of x are less tractable.

[2]In settings where higher unconditional moments matter, one need to be careful since some values of ω and γ may introduce kurtosis in the unconditional distribution of x.

where the variance of the idiosyncratic noise is common across agents. When $S = 1$ agents also observe the public signal y

$$y = x + \eta : \eta \sim N\left(0, \sigma_\eta^2\right). \tag{2.9}$$

The information set of agent j is thus given by

$$\Omega_j^0 = x_j, S \tag{2.10}$$

if $S = 0$ and

$$\Omega_j^1 = y, x_j, S \tag{2.11}$$

if $S = 1$. The fact that all agents observe y when $S = 1$ is common knowledge (though this does not really matter until later). The conditional expectations of x are then given by the standard formulas for multiple signals with independent Gaussian noise processes

$$E\left(x \mid \Omega_j^0\right) = \frac{\sigma_\varepsilon^{-2}}{\sigma_\varepsilon^{-2} + \sigma^{-2}} x_j \tag{2.12}$$

and

$$E\left(x \mid \Omega_j^1\right) = \frac{\sigma_\varepsilon^{-2}}{\sigma_\varepsilon^{-2} + \sigma_\eta^{-2} + \gamma^{-1}\sigma^{-2}} x_j + \frac{\sigma_\eta^{-2}}{\sigma_\varepsilon^{-2} + \sigma_\eta^{-2} + \gamma^{-1}\sigma^{-2}} y \tag{2.13}$$

That is, the conditional expectations of x are linear functions of the signals where the weights are determined by the relative precision of the individual signals. The posterior variances are also standard for normally distributed signals with independent noise processes and given by

$$E\left[x - E\left(x \mid \Omega_j^0\right)\right]^2 = \left(\sigma_\varepsilon^{-2} + \sigma^{-2}\right)^{-1} \tag{2.14}$$

and

$$E\left[x - E\left(x \mid \Omega_j^1\right)\right]^2 = \left(\sigma_\varepsilon^{-2} + \sigma_\eta^{-2} + \gamma^{-1}\sigma^{-2}\right)^{-1} \tag{2.15}$$

2.3. **The implications of *man-bites-dog* signals.** The set up described above allows us to prove two results that at first may appear counter intuitive. First, the posterior uncertainty can be larger after the signal y has been observed, compared to when it has not been observed. Secondly, the dispersion of expectations about x may increase after y is observed, even though y is a public signal.

Proposition 1. *The posterior uncertainty about x can be larger when the signal y is observed relative to when it is not. I.e. there are parameter values such that the inequality*

$$E\left[x - E\left(x \mid \Omega_j^0\right)\right]^2 < E\left[x - E\left(x \mid \Omega_j^1\right)\right]^2 \tag{2.16}$$

holds.

Proof. Directly comparing the posterior variances (2.14) and (2.15) implies that the proposition holds if

$$\sigma_\varepsilon^{-2} + \sigma^{-2} > \sigma_\varepsilon^{-2} + \sigma_\eta^{-2} + \gamma^{-1}\sigma^{-2} \tag{2.17}$$

Rearranging this expression gives

$$\sigma_\eta^2 > \frac{\sigma^2}{(1 - \gamma^{-1})} \tag{2.18}$$

so that if $\gamma > 1$ there exists some $\sigma_\eta^2 > \sigma^2$ such that the inequality in the proposition holds. □

Proposition 1 states that if the signal y is noisy enough and $\gamma > 1$, the posterior variance may be larger when y is observed compared to when it is not. This is possible due to the conditional nature of the probability of observing y. The fact that y is observed indicates that something unusual is more likely to have happened, in the sense that larger deviations of x from its mean are more likely if $\gamma > 1$. When y is very noisy, this effect dominates. Conversely, if γ is large so that only very unusual events are more likely to generate a signal y, the signal y need not be very noisy for the increased uncertainty to dominate. However, since the right hand side of (2.18) has a minimum of σ^2 when $\gamma \to \infty$, that $\sigma_\eta^2 > \sigma^2$ is a necessary condition for the posterior uncertainty to be larger. Recognizing that the fact that a signal is available tells agents something about the (second moments of the) state can thus cause posterior uncertainty to be larger after a public signal is observed as compared to when it is not.

Corollary 1. *When the inequality*

$$\sigma_\eta^2 > \frac{\sigma^2}{(1-\gamma^{-1})} \tag{2.19}$$

holds, the cross sectional dispersion of expectations about x is larger when y is observed than when it is not.

The proof follows directly from that the denominator in the weight on the private signal is the same as the denominator in the posterior variances. The cross sectional dispersion is increasing in the weight on the private signal, holding the variance of the idiosyncratic noise constant. The same conditions that deliver a higher posterior variance thus also deliver more weight on the private signal and the intuition is also similar. If the public noise variance is high and the conditional likelihood of a tail event is high, agents will put more weight on other (e.g. private) sources of information.

It is straight forward to show that the total weight agents put on all signals increases when $S = 1$.

Proposition 2. *The average expectation of x responds stronger to x when S=1 than when S=0.*

Proof. In the Appendix. □

The proof of Proposition 2 simply entails verifying that the sum of the coefficients on the two signals is larger when $S = 1$ than the coefficient on the single private signal when $S = 0$. The result is a direct consequence of the decrease in precision of agents beliefs "posterior" to observing $S = 1$ but "prior" to observing y. The weight agents put on y and x_j is inversely related to this precision and the sum of these weights thus go up unambiguously.

This section demonstrated that Bayesian updating to signals that are more likely to be available after an unusual event can have quite different properties compared to Bayesian updating to signals whose availability is uncorrelated with the underlying variable. Particularly, it was demonstrated that both posterior uncertainty and the cross-sectional dispersion of posterior beliefs can increase when a public man-bites-dog signal is observed. It was also shown that on average, agents expectations of the latent variable respond stronger to a given

deviation of the latent variable from its mean when there is a man-bites-dog signal available. In the next section, the man-bites-dog information structure is embedded in a simple business cycle model in order to illustrate the dynamic consequences of having signals whose availability is positively correlated with tail events.

3. A business cycle model

This section presents a simple business cycle model, following closely that of Lorenzoni (2009) but with a man-bites-dog information structure. As in the model by Lorenzoni, there are informationally separated islands that are subject to both common and island specific productivity shocks that cannot be distinguished by direct observation. Instead, island inhabitants need to form an estimate of the common component of productivity in order to make optimal consumption and price setting decisions. In the model, there is both island specific and public information and the man-bites-dog signal is specified as a public signal about aggregate productivity. This means that information about large changes to productivity will be more correlated across agents than information about small changes. The model also features strategic complementarities. It is well-known that strategic complementarities tend to increase the importance of public signals, e.g. Morris and Shin (2002). In the model presented here, this effect will be further reinforced by the man-bites-dog mechanism, generating an amplification mechanism implying that aggregate variables may respond more than proportionally to larger productivity shocks.

3.1. **Preferences and technology.** The set up is a stationary version of the island economy described in Lorenzoni (2009). There is a continuum of islands indexed by $j \in (0,1)$ and on each island there is a continuum of firms indexed by $i \in (0,1)$ producing differentiated goods. On each island there is a representative household that consume goods and supply labor $N_{j,t}$ to the firms on the island. Households are thus also indexed by $j \in (0,1)$ and a household on island j maximizes

$$E\sum_{s=0}^{\infty}\beta^s\left[\exp(d_{j,t})\ln C_{j,t} - \frac{N_{j,t}^{1+\varphi}}{1+\varphi} \mid \Omega_t(j)\right] \tag{3.1}$$

where $\Omega_t(j)$ is the information set of inhabitants of island j and $C_{j,t}$ is the consumption bundle consumed by island j households defined as

$$C_{j,t} = \left(\int_{\mathcal{B}j,m}\int_0^1 C_{i,m,j,t}^{(\delta-1)/\delta} di\ dm\right)^{\delta/(1-\delta)} \tag{3.2}$$

As in Lorenzoni (2009), households only consume a subset $\mathcal{B}_{j,m} \subseteq (0,1)$ of the available goods in the economy and the set $\mathcal{B}_{j,m}$ is drawn by nature in each period. The shock $d_{j,t}$

$$d_{j,t} = d_t + \zeta_{j,t} : \zeta_{j,t} \sim N\left(0, \sigma_\zeta^2\right) \tag{3.3}$$

is a demand disturbance that is correlated across islands and the common component d_t follows an AR(1) process

$$d_t = \rho_d d_{t-1} + u_t^d : u_t^d \sim N\left(0, \sigma_d^2\right) \tag{3.4}$$

The demand disturbance d_t is not present in the original model by Lorenzoni (2009) but is needed here in order to avoid stochastic singularity when the model is estimated. Firm i on island j produce good i, j using the technology

$$Y_{i,j,t} = \exp\left(a_{j,t}\right) N_{i,j,t} \tag{3.5}$$

(The log of) productivity $a_{j,t}$ is the sum of a common component a_t and the island specific component $\varepsilon_{j,t}$

$$a_{j,t} = a_t + \varepsilon_{j,t} : \varepsilon_{j,t} \sim N\left(0, \sigma_\varepsilon^2\right) \ \forall \ \ j, t. \tag{3.6}$$

The common productivity component a_t follows an AR(1) process

$$a_t = \rho_a a_{t-1} + u_t^a : u_t^a \sim N\left(0, \sigma_a^2\right) \tag{3.7}$$

Firms on island j are owned by island j households and set prices $P_{j,t}$ to maximize discounted expected profits $\Pi_{j,t}$

$$E\left[\sum_{s=t}^{\infty} \theta^s \beta^s \frac{C_{j,t}}{C_{j,t+s}} \Pi_{j,t} \mid \Omega_t^j\right] = E\left[\sum_{s=t}^{\infty} \theta^s \beta^s \frac{C_{j,t}}{C_{j,t+s}} \left(P_{j,t+s} Y_{j,t+s} - W_{j,t+s} N_{j,t+s}\right) \mid \Omega_t^j\right] \tag{3.8}$$

where θ is the probability of not changing the price of a given good in a given period. The intertemporal budget constraint of households on island j is given by

$$\frac{B_{j,t+1}}{R_t} + \int_{\mathcal{B}j,m} \int P_{i,j,t} C_{j,m} di \ dm = B_{j,t} + W_{j,t} N_{j,t} + \int \Pi_{i,j,t} di \tag{3.9}$$

where the (log of the) nominal one period interest rate R_t follows a Taylor rule

$$r_t = \phi_\pi \pi_t + \phi_y y_t + \phi_r r_{t-1} + u_t^r : u_t^r \sim N\left(0, \sigma_r^2\right) \tag{3.10}$$

and $B_{j,t}$ are holdings of nominal bonds that pay one dollar in period t by households on island j.

3.2. **Linearized equilibrium conditions.** The model presented above can be log linearized around a non-stochastic steady state, yielding the following equilibrium conditions. (i) An Euler equation determining the optimal intertemporal allocation of consumption

$$c_{j,t} = E\left[c_{j,t+1} \mid \Omega_t^j\right] - r_t + E\left[\overline{\pi}_{\mathcal{B}j,t+1} \mid \Omega_t^j\right] + d_{j,t} \tag{3.11}$$

where $\overline{\pi}_{\mathcal{B}j,t+1}$ is the inflation of the goods basket consumed on island j in period $t+1$. (ii) A labor supply condition equating marginal disutility of labor supply with the marginal utility of consumption multiplied by the real wage

$$w_{j,t} - \overline{p}_{\mathcal{B}j,t} = c_{j,t} + \varphi n_{j,t}. \tag{3.12}$$

(iii) A demand schedule for good j depending on aggregate output and the relative price of good j

$$y_{j,t} = \int_{\mathcal{C},j,t} c_{m,t} dm - \delta\left(p_{j,t} - \int_{\mathcal{C},j,t} \overline{p}_{m,t} dm\right) \tag{3.13}$$

where $\int_{\mathcal{C},j,t} \overline{p}_{m,t} dl$ is the log of the relevant price sub index for consumers from other islands buying goods from island j. (iv) An island j Phillips curve relating inflation on island j to the nominal marginal cost on island j and expected future island j inflation is given by

$$\begin{aligned} p_{j,t} - p_{j,t-1} &= \lambda \left(\overline{p}_{\mathcal{B}j,t} + c_{j,t} - p_{j,t} - a_{j,t} \right) + \lambda \varphi \left(y_{j,t} - a_{j,t} \right) \\ &\quad + \beta E \left(p_{j,t+1} - p_{j,t} \mid \Omega_t^j \right) \end{aligned} \tag{3.14}$$

where $\overline{p}_{\mathcal{B}j,t}$ is the relevant price subindex for consumers on island j and $\lambda = (1-\theta)(1-\theta\beta)/\beta$. The steps required to arrive at the linearized equilibrium conditions (3.11) - (3.14) are identical to those described in Lorenzoni (2009).

3.3. **Local information.** The inhabitants of each island observes their own productivity and demand disturbances $a_{j,t}$ and $d_{j,t}$. Since these contain a component that is common across islands, these local variables are informative about the aggregate state. In addition to these exogenous local signals, firms and households can also extract information about the aggregate state from observing the demand for their own good $y_{j,t}$ and the price of the goods basket that they purchase. Lorenzoni (2009) assumes that the islands visited by the inhabitant of island j while shopping are drawn so that the price index of the goods basket consumed by island j inhabitants is equal to the aggregate price level plus a normally distributed island j specific shock. We will make a similar assumption, with an adjustment to the mean of of the normally distributed shock such that the signal is conditionally stationary. That is, the set $\mathcal{B}_{j,t}$ is drawn such that

$$\overline{p}_{\mathcal{B}j,t} = p_t + \xi_{j,t}^1 : \xi_{j,t}^1 \sim N(p_{j,t-1} - p_{t-1}, \sigma_{\xi 1}^2) \tag{3.15}$$

Since $p_{j,t-1}$ is observable by the inhabitants of island j the price index of the good purchased by island j inhabitant is thus a noisy measure of aggregate inflation $p_t - p_{t-1}$, rather than of the aggregate price level as in Lorenzoni (2009). Reformulating the signal structure this way does not change anything substantial in the model but simplifies the representation of agents' filtering problems since all other variables are stationary (while the price level is not). Similarly, we assume that the set of islands $\mathcal{C}_{j,t}$ are drawn such that (3.13) takes the form

$$y_{j,t} = y_t - \delta \left(p_{j,t} - p_t \right) + \xi_{j,t}^2 : \xi_{j,t}^2 \sim N(\delta \left[p_{j,t-1} - p_{t-1} \right], \sigma_{\xi 2}^2) \tag{3.16}$$

Again, the adjustment of the mean of the disturbance relative to Lorenzoni (2009) is made in order to keep signals stationary and does not change anything of substance. As in Lorenzoni's original model, the shocks $\xi_{j,t}^1$ and $\xi_{j,t}^2$ are introduced in order to prevent local interactions from perfectly revealing the aggregate state.

3.4. **The joint distribution of signals and shocks.** The man-bites-dog signal structure is embedded in the business cycle model in a similar way as in the static setting discussed above. The unobservable variable of interest is the common component of productivity a_t and I will specify a joint distribution of the indicator variable s_t and the innovations u_t^a in (3.7) such that a man-bites-dog signal is more likely to be generated when there has been a large (in absolute terms) innovation to the common productivity process a_t. The indicator variable s_t takes the value 1 when a man-bites-dog signal is generated in period t which

occurs with unconditional probability ω. Similar to the static setting of Section 2, a mixture normal density for u_t^a will be used to keep the filtering problem tractable

$$u_t^a \sim (1-\omega)\, N\left(0, \sigma^2\right) + \omega N\left(0, \gamma\sigma^2\right). \tag{3.17}$$

The unconditional variance of the productivity innovations u_t^a is then given by

$$E\left(u_t^a\right)^2 = (1-\omega)\,\sigma^2 + \omega\gamma\sigma^2 \tag{3.18}$$

To complete the description of the joint distribution of innovations and signals, it is further assumed that when $s_t = 1$ all households observe an additional public signal $z_{a,t}$ given by

$$z_{a,t} = a_t + \eta_t : \eta_t \sim N\left(0, \sigma_\eta^2\right) \tag{3.19}$$

The signal z_t^a is thus the man-bites-dog signal and the vector of observables $\mathbf{z}_{j,t}$ available to households and firms on island j in period t then is

$$\mathbf{z}_{j,t} = \left[a_{j,t}, d_{j,t}, y_{j,t}, \overline{p}_{j,t}, r_t, s_t\right] \tag{3.20}$$

if $s_t = 0$ and

$$\mathbf{z}_{j,t} = \left[a_{j,t}, d_{j,t}, y_{j,t}, \overline{p}_{j,t}, r_t, s_t, z_t^a\right] \tag{3.21}$$

if $s_t = 1$. The information set $\Omega_{j,t}$ of firms and households on island j evolves as

$$\Omega_{j,t} = \left\{\mathbf{z}_{j,t}, \Omega_{j,t-1}\right\} \tag{3.22}$$

This completes the description of the model. It is perhaps worth noting that the original model of Lorenzoni (2009) is nested in the model presented here, i.e. by setting $\rho_a = \omega = \gamma = 1$ and $\sigma_d^2 = 0$.

4. Solving and estimating the model

There are two features of the model presented above that make standard solution methods for linear rational expectations models inapplicable. First, there is island specific information about variables of common interest to all islands. Natural state representation then tend to become infinite, due to the well-known problem of the infinite regress of expectations that arises when agents need to "forecast the forecasts of others" (see Townsend 1983 and Sargent 1991). Second, the precision of agents' information is a function of the realized history of s_t and thus varying across time. In this section, I first describe how the method to solve dynamic models with private information proposed in Nimark (2011) can be modified to solve a model with a time varying information structure. Details are kept at a minimum and focus is on aspects of the solution method that help intuition for how time varying information sets translate into time varying equilibrium dynamics, though a complete description of the solution algorithm is given in the Appendix. When solved, the model is in standard state space form, albeit with time varying parameters and the section ends with a description of how a posterior estimate of the joint distribution of the parameters of the model and the history of s_t can be constructed using a version of the Metropolis-Hastings algorithm.

4.1. **Solving dynamic models with private information.** In Nimark (2011) it is demonstrated that if the variance of the unobservable exogenous state is finite, the equilibrium dynamics of a linear rational expectations model with privately informed agents can be approximated to an arbitrary accuracy by a finite dimensional representation. There are two parts to deriving this result. The first step puts structure on higher order expectations, i.e. expectation about other agents expectations, by exploiting that it is common knowledge that agents form model consistent expectations. The second step uses this structure to show that the impact of orders of expectations is decreasing in the order of expectation. The second step is somewhat involved, and interested readers are referred to the original reference for more details. Here we briefly describe how common knowledge of model consistent expectations help put structure on the dynamics of higher order expectations.

Let $\mathbf{x}_t$ denote a vector containing the exogenous state variables a_t and d_t so that

$$\mathbf{x}_t \equiv \begin{bmatrix} a_t \\ d_t \end{bmatrix} \tag{4.1}$$

The solution method then delivers a law of motion of the state X_t of the form

$$X_t = MX_{t-1} + N\mathbf{u}_t : \mathbf{u}_t \sim N(0, I) \tag{4.2}$$

where

$$X_t \equiv \begin{bmatrix} \mathbf{x}_t \\ \mathbf{x}_t^{(1)} \\ \vdots \\ \mathbf{x}_t^{(\overline{k})} \end{bmatrix} \tag{4.3}$$

and

$$\mathbf{x}_t^{(k+1)} \equiv \int E\left[\mathbf{x}_t^{(k)} \mid \Omega_t(j)\right] dj \tag{4.4}$$

The state thus consists of average higher order expectations of current productivity a_t and the common demand shock d_t. The constant $\overline{k}$ is the maximum order of expectation considered.

The intuition for how common knowledge of model consistent expectations puts structure on the dynamics of higher order expectations so that the law of motion (4.2) can be found is the following: As usual in rational expectations models, first order expectations $\mathbf{x}_t^{(1)}$ are optimal, i.e. model consistent estimates of the actual state $\mathbf{x}_t$. The knowledge that other islands have model consistent estimates allow inhabitants of an individual island to treat average first order expectations as a stochastic process with known properties when they form second order expectations. Common knowledge of the model thus implies that second order expectations $\mathbf{x}_t^{(2)}$ are optimal estimates of $\mathbf{x}_t^{(1)}$ given the law of motion for $\mathbf{x}_t^{(1)}$. Imposing this structure on all orders of expectations allows us to find the law of motion (4.2) for the complete hierarchy of expectations as functions of the structural parameters of the model. In practise, this is implemented by letting the agents use the Kalman filter to form an estimate of the state X_t so that agent j's state estimate evolves according to the update equation

$$E\left[X_t \mid \Omega_t(j)\right] = (I - K_t D)ME\left[X_{t-1} \mid \Omega_{t-1}(j)\right] + K_t \mathbf{z}_{j,t} \tag{4.5}$$

where K_t is the Kalman gain. The law of motion (4.2) of the state will thus depend on the update equation (4.5) since by the definitions (4.3) and (4.4) we have that

$$X_t \equiv \left[\begin{array}{cc} I_{2k} & \mathbf{0}_{2k\times 2} \end{array} \right] \left[\begin{array}{c} \mathbf{x}_t \\ \int E\left[X_t^{(k)} \mid \Omega_t(j)\right] dj \end{array} \right] \tag{4.6}$$

i.e.. the state is partly made up of higher order expectations about the exogenous state vector $\mathbf{x}_t$.

For a given law of motion (4.2), the endogenous variables can be computed as a linear function of higher order expectations about the current state $\mathbf{x}_t$ and the aggregate shocks $\mathbf{u}_t$ by using that the structural equations of the linearized model can be written in the form

$$\left[\begin{array}{c} \pi_t \\ y_t \end{array} \right] = A \int E\left(\left[\begin{array}{c} \pi_{t+1} \\ y_{t+1} \end{array} \right] \mid \Omega_{j,t} \right) + (B+C)\, X_t + G_u \mathbf{u}_t + G_r r_{t-1}$$

Conjecturing a solution of the form

$$\left[\begin{array}{c} \pi_t \\ y_t \end{array} \right] = GX_t + G_u \mathbf{u}_t + G_r r_{t-1} \tag{4.7}$$

together with method of undetermined coefficients then implies that matrix G must satisfy the equation

$$G = AGMH + (B+C) \tag{4.8}$$

where H is a matrix defined so that

$$\int E\left[X_t \mid \Omega_t(j)\right] dj = HX_t \tag{4.9}$$

i.e. H moves a vector of average higher order expectations one step "up" in orders of expectations. For a given A and M, the equation (4.8) can be solved for G by function iteration.

The fact that the state of the model is made up of expectations and higher order expectations means that the law of motion of the state in a given period is a function of the precision of agents priors as well as the precision of period t signals. Since these quantities will depend on the history of s_t, the law of motion of the state will be time varying. That the law of motion will be time varying then in turn implies that the function mapping the current state into endogenous variables will also be time varying.

4.2. **Regimes and endogenous time varying dynamics.** Changes in the regimes s_t will in general have persistent effects on the dynamics of the model. The reason is that if, for example, there is a man-bites-dog episode in period t that increases the posterior uncertainty in period t, then this will translate into an increase in prior uncertainty in period $t+1$. How much weight agents put on signals observed in period $t+1$ depends on the precision of the signals relative to the precision of the prior. A large posterior uncertainty in period t thus translates into more weight being put on signals observed in period $t+1$. Since the state is made up of the (higher order) estimates of the exogenous states, the dynamics of the state in a given period depend on the weight agents put on signals in that period so a single man-bites-dog episode may have a persistent effect on the dynamics of the state. This makes the model more interesting and allow us to potentially fit more facts in the data. It also allows us to exploit the variation in the cross-sectional dispersion of forecasts as measured

by the SPF to make sharper inference about the precision of agents information at different points in time. But since a period t regime s_t has persistent effects we need to keep track of the history of s_t (which we denote s^t) to determine period t equilibrium dynamics. Here we describe how this is done in practice.

As long as the filtering errors of agents follows a stable process, we do not need to keep track of the entire history but only the most recent realizations of s_t. How far back in time the realizations of s_t matter of course depends on the parameters of the model and more specifically, it will depend on the eigenvalues of the process that propagates the variance of the filtering errors through time. If the underlying processes are not very persistent, or if information is very accurate, filtering errors do not tend to be long lived and only a few of the past realizations of s_t influence current dynamics. In general, how many lags of s_t that are relevant for period t dynamics depend on the parameters of the model and needs to be checked on a case-by-case basis. However, that the matrix M in the law of motion is stable, is sufficient to ensure that the current state uncertainty can be approximated accurately by a finite history s^t.

We will proceed by specifying a maximum lag of s_t, say $s_{t-\mathcal{T}}$, that matters for current dynamics and then check whether increasing the lag $\mathcal{T}$ changes the dynamics sufficiently to motivate the increased computational burden. If we for instance set $\mathcal{T} = 4$ we thus implicitly assume that period t dynamics are similar for all histories s^{t-4}. In practise, we need to ensure that $\mathcal{T}$ is large enough so that the Kalman gain K_t in agents' update equation (4.5) is invariant to whether $s^t = \{..., 1, 0, 0, 0, 1\}$ or $s^t = \{..., 0, 0, 0, 0, 1\}$. If that is the case, we only need to keep track of the $\mathcal{T}$ most recent realizations of s_t. This strategy implies that if there are 2 different exogenous regimes (i.e. $s_t \in \{0, 1\}$) we have $2^{\mathcal{T}}$ relevant different histories s^t. The number of endogenous regimes thus increases exponentially in $\mathcal{T}$. For instance, if only the last four periods matter for current dynamics we have $2^4 = 16$ different endogenous regimes, but if the last 8 periods matter the number of endogenous regimes is $2^8 = 256$. [3] It is perhaps worth pointing out here that while we truncate the history of s_t used to compute the Kalman gain K_t, agents still condition on the entire history of vector of observables $\mathbf{z}_{j,t}$.

4.3. **Solving the model with time varying information sets.** As described in Section 4.1 above, the fact that the state of the model is made up of expectations and higher order expectations means that how agents update their state estimates determines the actual law of motion of the very same state that agents form estimates about. Here, a brief overview of how the solution method needs to be amended when the information structure is time-varying is given, though most of the details are relegated to the Appendix.

First, with time varying information sets, the law of motion now depends on the history of regimes s^t so that

$$X_t = M(s^t)X_{t-1} + N(s^t)\mathbf{u}_t : \mathbf{u}_t \sim N(0, I) \tag{4.10}$$

[3]The endogenous regimes are quite similar which reduces the computational burden. Setting $\mathcal{T} = 12$ which implies that there are $2^{12} = 4096$ different endogenous regimes result in a computational time of about 10 seconds to solve the model once on a standard desktop PC. This can be compared to a solution time of about 0.3 seconds for $\mathcal{T} = 5$ at the posterior mode of the estimated model.

The endogenous variables can still be written as linear functions of the state, but now with time varying parameters

$$\begin{bmatrix} \pi_t \\ y_t \end{bmatrix} = G(s^t)X_t + G_u \mathbf{u}_t + G_r r_{t-1} \tag{4.11}$$

The matrices $G_u$and G_r captures the effect of the monetary policy shock and lagged interest rate respectively. Since interest rate is observable these matrices do not depend on s^t. The matrix $G(s^t)$ can be computed by noting that for a given conjectured law of motion (4.10) and a given linear function (4.11) we have that

$$\begin{bmatrix} \pi_t \\ y_t \end{bmatrix} = A \int E\left(\begin{bmatrix} \pi_{t+1} \\ y_{t+1} \end{bmatrix} \mid \Omega_{j,t} \right) + (B+C)\, X_t + G_r r_{t-1} \tag{4.12}$$

When forming expectations about future inflation and output agents will need to take into account that there may a man-bites-dog signal available in the next period. Expectations of output and inflation thus depend on the probability ω that s_{t+1} will take the value 1. The Euler equation for inflation and output (4.12) can thus be written as

$$\begin{aligned} \begin{bmatrix} \pi_t \\ y_t \end{bmatrix} &= \omega A G\left(s_1^{t+1}\right) M\left(s_1^{t+1}\right) H X_t \\ &\quad + (1-\omega)\, A G\left(s_0^{t+1}\right) M\left(s_0^{t+1}\right) H X_t \\ &\quad + (B+C)\, X_t + G_u \mathbf{u}_t + G_r r_{t-1} \end{aligned}$$

where s_n^{t+1} denotes the history s^{t+1} with $s_{t+1} = n$. Using the conjectured form (4.11) and equating coefficients implies

$$\begin{aligned} G\left(s^t\right) &= \omega A G\left(s_1^{t+1}\right) M\left(s_1^{t+1}\right) H \\ &\quad + (1-\omega)\, A G\left(s_0^{t+1}\right) M\left(s_0^{t+1}\right) H \\ &\quad + (B+C) \end{aligned} \tag{4.13}$$

With a constant information structure the matrices $G\left(s^t\right)$ and $G\left(s_1^{t+1}\right)$ are the same and (4.13) simplifies to (4.8). When information sets are time-varying, the solution algorithm becomes a fixed point problem over the $\mathcal{S}$ relevant histories of s_t. A fixed point can again be found by function iteration. The Appendix describes an algorithm that for given $G(s^t), M(s^t)$ and $N(s^t)$ computes new values of the same matrices. Nimark (2011) uses a version of Brouwer's fixed point theorem to prove for the time invariant case that a solution to this iteration exists and that there is a finite maximum order of expectation $\overline{k}$ that results in equilibrium dynamics that are arbitrarily close to the dynamics of the limit $\overline{k} \to \infty$. For these results to apply it is sufficient that the variance of the exogenous state $\mathbf{x}_t$ is finite and these results thus also apply to the current model as long as $\mathbf{x}_t$ follows a stable process.

4.4. **Estimating the model.** The solved model is a state space system with time varying parameters and standard likelihood based methods are applicable to estimate the parameters of the model. In addition to the 18 structural parameters of the model, which we denote

$$\Theta = \{\rho_a, \rho_\xi, \sigma_a, \sigma_d, \sigma_r, \sigma_\varepsilon, \sigma_\zeta, \sigma_{\xi 1}, \sigma_{\xi 2}, \sigma_\eta, \delta, \eta, \phi_\pi, \phi_y, \phi_r, \theta, \beta, \omega, \gamma\} \tag{4.14}$$

we also want to construct a posterior estimate of the indicator variable s_t that keeps track of whether there was a man-bites-dog signal available in period t or not. Below we describe how this can be done by sampling from the two conditional distributions $p\left(\Theta \mid s^T, Z^T\right)$ and $p\left(s^T \mid \Theta, Z^T\right)$. Dividing the sampling from the joint posterior distribution of Θ and s^T into two blocks also lets us get around the problem that unlike the agents inside the model, we as econometricians do not observe the regimes s_t. Conditional on a draw from s^T, the model is linear-Gaussian and it is easy to evaluate the likelihood.

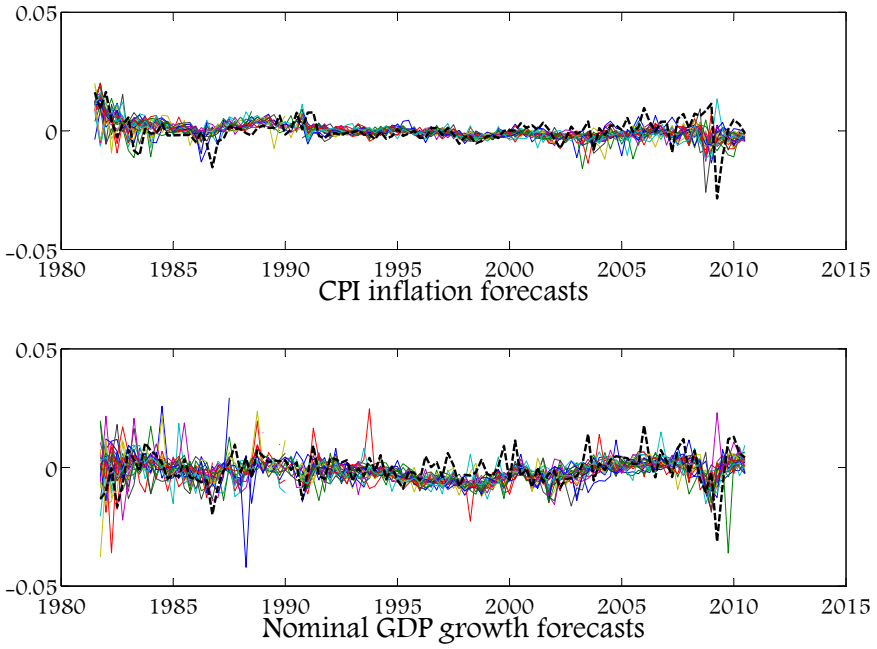

FIGURE 2. Actual and SPF Forecasts of CPI inflation and Nominal GDP growth

4.5. **The data.** The time series used to estimate the model are US CPI inflation, the Federal Funds rate, US real GDP, the quarterly time series of total factor productivity constructed by John Fernald (2010) and raw survey data from the Survey of Professional Forecasters (SPF). The data is quarterly and the sample ranges from 1981:Q3 to 2010:Q4. The start date is chosen based on the availability of survey data for inflation forecasts and the end date is the date of the most recent data on real GDP and total factor productivity. CPI inflation is de-trended using a linear trend and the same trend is taken out of the Federal Funds rate. Real GDP is de-trended using the HP-filter with a smoothing coefficient of 1600. The raw survey data is one quarter ahead forecasts of CPI inflation and nominal GDP growth taken from the Survey of Professional Forecasters available from the web site of the Federal Reserve Bank of Philadelphia. The (de-trended) survey data is plotted in Figure 2 along with the actual outcomes of CPI inflation and Nominal GDP growth. Inflation forecasts are

de-trended using the CPI inflation trend and nominal GDP growth forecasts are de-trended by subtracting the inflation trend and the growth in the real GDP (HP-filter) trend. We denote the vector of observables in period t Z_t. All elements in Z_t have natural counterparts in the model. Specifically, TFP will be taken as a noisy measure of the common productivity component and the vectors of survey forecasts $\mathbf{f}'_{t,\pi}$ and $\mathbf{f}'_{t,\pi+\Delta y}$ are taken to be representative of the expectations of the privately informed agents in the model.

The observables in period t are thus given by

$$Z_t = \left[\begin{array}{cccccc} a_t & \pi_t & y_t & r_t & \mathbf{f}'_{t,\pi} & \mathbf{f}'_{t,\pi+\Delta y} \end{array} \right]' \tag{4.15}$$

and linked to the state of the model by the measurement equation

$$Z_t = \overline{D}(s^t)X_t + D_r r_{t-1} + \overline{R}(s^t)u_t \tag{4.16}$$

Due to the fact that the number of survey respondents is not constant in the sample, the dimensions of both $\overline{D}$ and R are varying across time. More interestingly, the entries of $\overline{D}$ and $\overline{R}$ also varies over time. The matrix $\overline{D}$ is time varying since the function mapping the state into endogenous variables is time varying. The matrix $\overline{R}$ is time varying since the cross-sectional dispersion of forecasts are time varying in the model and each survey entry can be viewed as a noisy measure of the average expectation where the variance of the "noise" is the cross-sectional variance of forecasts implied by the model. Individual respondents can be tracked in the SPF so while it is in theory possible to exploit the (limited) panel dimension of the SPF responses, it is not feasible in practise. Doing so would require that in order to evaluate the likelihood, we as econometricians would need to carry along an individual state for each respondent in the SPF, thereby increasing the state dimension by a multiplier of 50. Instead, we treat individual survey responses $f_{t,\pi}$ and $f_{t,\pi+\Delta y}$ in each period as independent draws from the distributions

$$f_{t,\pi} \sim N\left(\int E\left[\pi_t \mid \Omega_{j,t}\right] dj, \sigma^2_{f\pi}(s^t)\right) \tag{4.17}$$

and

$$f_{t,\pi+\Delta y} \sim N\left(\int E\left[\Delta y_{t+1} + \pi_t \mid \Omega_{j,t}\right] dj, \sigma^2_{f\pi+\Delta y}(s^t)\right) \tag{4.18}$$

where the variances $\sigma^2_{f\pi}(s^t)$ and $\sigma^2_{f\pi+\Delta y}(s^t)$ are the model implied cross-sectional variance of inflation and nominal GDP growth expectations. The cross-sectional variances vary across time and are a function of the structural parameters Θ as well as the history s^t.

There are four observable variables in Z_t that map into aggregate variables in the theoretical model, but in most periods there are only three aggregate disturbances, i.e. the innovations to common productivity, the demand shocks and to the interest rate rule. To avoid stochastic singularity when evaluating the likelihood function, we treat the TFP series as well as the in CPI inflation and real GDP growth as being measured with (very small) noise.[4]

[4]The variances of the measurement errors are about $1/100^{th}$ of the variance of the actual time series.

4.6. **Estimation Algorithm.** The posterior distribution of Θ and s^T are estimated using a Multiple-Block Metropolis algorithm (see Chib 2001). It exploits that conditional on a history of man-bites-dog regimes s^T, the model is in linear-Gaussian state space form. The estimation algorithm can be described as follows.

(1) Specify initial values Θ_0 and s_0^T.
(2) Repeat for $j = 1, 2, ..., J$
  (a) Block 1: Draw from $p\left(\Theta \mid s^T, Z^T\right)$
    (i) Propose $\Theta^* \sim N\left(\Theta_{j,1}, c\Sigma\right)$ where c is a constant a nd Σ is a matrix with the absolute with diagonal $|\Theta_0|$ and zeros on the off-diagonal. The constant c is chosen to achieve an acceptance rate of about 22%.
    (ii) Calculate $\alpha_j^{\Theta} = \min\left\{\frac{L\left(Z^T|s_{j-1}^T,\Theta^*\right)}{L\left(Z^T|s_{j-1}^T,\Theta_{j-1}\right)}, 1\right\}$
    (iii) Set $\Theta_{j+1} = \Theta^*$ if $U(0,1) \leq \alpha_j^{\Theta}$ and $\Theta_{j+1} = \Theta_j$ otherwise.
  (b) Block 2: Draw from $p\left(s_\tau \mid s_{-\tau}^T, \Theta, Z^T\right)$ where $s_{-\tau}^T \equiv \{s_1, s_2, ...s_{\tau-1}, s_{\tau+1}, ..., s_{T-1}, s_T\}$
    (i) Draw a time period τ uniformly from the set $\{1, 2, ...T\}$
    (ii) Set $s_{-\tau}^{*T} = s_{-\tau,j-1}^T$
    (iii) Propose $s_\tau^* = 1$ if $U(0,1) \leq \omega$ and 0 otherwise.
    (iv) Calculate $\alpha_j^s = \min\left\{\frac{L\left(Z^T|s^{*T},\Theta_j\right)p(s^{*T})}{L\left(Z^T|s_{j-1}^T,\Theta_j\right)p(s_{j-1}^T)}, 1\right\}$
    (v) Set $s_j^T = s^*$ if $U(0,1) \leq \alpha_j^s$ and $s_j^T = s_{j-1}^T$ otherwise.
(3) Return values $\{\Theta_0, \Theta_1, ..., \Theta_J\}$ and $\{s_0^T, s_1^T, ..., s_J^T\}$

The log of the likelihood function $L\left(Z^T \mid s^T, \Theta\right)$ can be evaluated as

$$\log L\left(Z^T \mid s^T, \Theta\right) = -\frac{1}{2}\left\{\sum_{t=1}^{T} 2\pi \dim(Z_t) + \log\left|\Sigma(s^t)\right| + \widetilde{Z}_t'\Sigma(s^t)_t^{-1}\widetilde{Z}\right\} \tag{4.19}$$

where $\widetilde{Z}_t$ is the innovation to the observation vector Z_t defined as

$$\widetilde{Z}_t \equiv Z_t - E\left(Z_t \mid Z^{t-1}\right) \tag{4.20}$$

and

$$\Sigma(s^t) \equiv E\left(\widetilde{Z}_t\widetilde{Z}_t'\right) \tag{4.21}$$

The innovation $\widetilde{Z}_t$ can for a given history s^t and parameter vector Θ be computed as the Kalman filter innovations to the state space system made up of the measurement equation (4.16) and the state equation (4.10).

4.7. **Posterior parameter estimates.** The posterior estimates of Θ from 500 000 draws from the Metropolis algorithm are reported in Table 1.

Table 1 (Preliminary)
Posterior Parameter Estimates 1981:Q3-2010:Q3

Θ	Mode $\widehat{\Theta}$	Prior dist.	Posterior 2.5%-97.5%
		Preferences etc	
φ	1.10	$U(0,10)$	$(1.00-1.11)$
δ	1.07	$U(0,10)$	$(1.00-1.09)$
β	0.98	$U(0.96,1)$	$(0.97-0.99)$
θ	0.74	$U(0,1)$	$(0.72-0.75)$
		Exogenous aggregate processes	
ρ_a	0.87	$U(0,0.99)$	$(0.85-0.88)$
ρ_d	0.72	$U(0,0.99)$	$(0.69-0.72)$
σ_a	0.018	$U(0,0.99)$	$(0.014-0.020)$
σ_d	0.010	$U(0,1)$	$(0.008-0.011)$
σ_r	0.014	$U(0,1)$	$(0.012-0.016)$
		Island specific processes	
σ_ε	0.25	$U(0,1)$	$(0.24-0.30)$
$\sigma_{\xi 1}$	0.29	$U(0,1)$	$(0.29-0.31)$
$\sigma_{\xi 2}$,	0.31	$U(0,1)$	$(0.29-0.31)$
σ_ζ,	0.21	$U(0,1)$	$(0.19-0.22)$
		Man-bites-dog parameters	
ω	0.051	$U(0,1)$	$(0.046-0.054)$
γ	4.27	$U(1,10)$	$(4.01-4.29)$
σ_η	0.49	$U(0,10)$	$(0.48-0.53)$
		Taylor rule parameters	
ϕ_r	0.08	$U(0,1)$	$(0.07-0.10)$
ϕ_π	1.55	$U(0,10)$	$(1.50-1.56)$
ϕ_y	0.24	$U(0,10)$	$(0.20-0.25)$
Log likelihood at $\widehat{\Theta}$: 2933.7			

5. Estimated Man-bites-dog Dynamics

The estimated model can be used to quantify the contribution of the man-bites-dog mechanism to business cycle dynamics. This section presents impulse response functions of various objects as well as estimates of when man-bites-dog episodes have been more likely to have occurred in the past. We also discuss how man-bites-dog episodes can be interpreted at the aggregate level.

5.1. **The impact of innovations to productivity.** Figure 3 illustrates the impulse response function of output and inflation to a one (unconditional) standard deviation innovation to the common productivity shock a_t. The black lines are the responses when there is no man-bites-dog signal, the blue lines are the responses when there is a man-bites-dog signal in the impact period but none thereafter. It is clear from the figure that the response of output to a productivity shock of a given magnitude is substantially (and significantly) larger if it coincides with the generation of a man-bites-dog signal. A man-bites-dog signal

approximately doubles the impact of an innovation to productivity on output and inflation falls by approximately an additional third of the response when there is no man-bites-dog signal. Since the exogenous shock is the same for both cases, the difference must be driven by different responses of expectations.

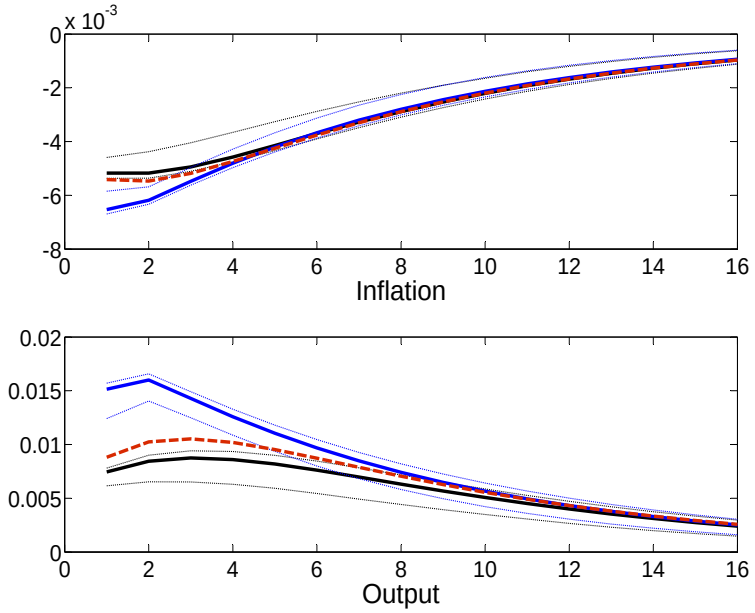

FIGURE 3. Impulse responses of inflation (top panel) and output (bottom panel) to a 1 s.d. innovation to productivity. Blue lines are for man-bites-dog signal in period 1, i.e. $s^t = 1, 0, 0, 0, 0$ and black lines for no man-bites-dog signal. Dotted black and blue lines show the 95 per cent posterior credible intervals. The dashed red lines are the responses of inflation and output when the availability of the signal z_t^a is uncorrelated with the innovation to productivity.

The stronger responses of endogenous variables to an innovation of given magnitude are caused by three conceptually different effects. First, since there is an additional signal available, average expectations respond stronger. Second, since the man-bites-dog signal is public (in the strong common knowledge sense of the word) higher order expectations will respond stronger than they would if the man-bites-dog signal was private. In equilibrium, there are strategic complementarities between individual price and consumption decisions and aggregate price and consumption outcomes. As pointed out by Morris and Shin (2002), public signals can then appear to be disproportionately influential on affecting equilibrium outcomes as they work as "focal points" for expectations. Third, the man bites-dog effect makes agents willing to update their expectations further for a given magnitude of the surprise component in their signal vector since the availability of a man-bites-dog signal suggests that a large innovation is more likely to have occurred. We can get a handle on the

relative importance of the man-bites-dog effect by solving the model under the alternative assumption that the availability of the man-bites-dog signal is entirely random, that is, under the assumption that a public signal about productivity is as likely to be observed when there has been a large innovation to productivity as when there has been a small innovation. The results are illustrated by the red dashed line in Figure 3. While the first two effects contribute a little to the stronger responses of output and inflation, it is clear that the man-bites-dog effect explains most of the difference between the black and blue lines. The "more information" and the "coordination effect" only increases the impact of productivity on output by from 0.0076 to 0.0088, while the response with a man-bites-dog signal virtually doubles the impact multiplier to 0.015.

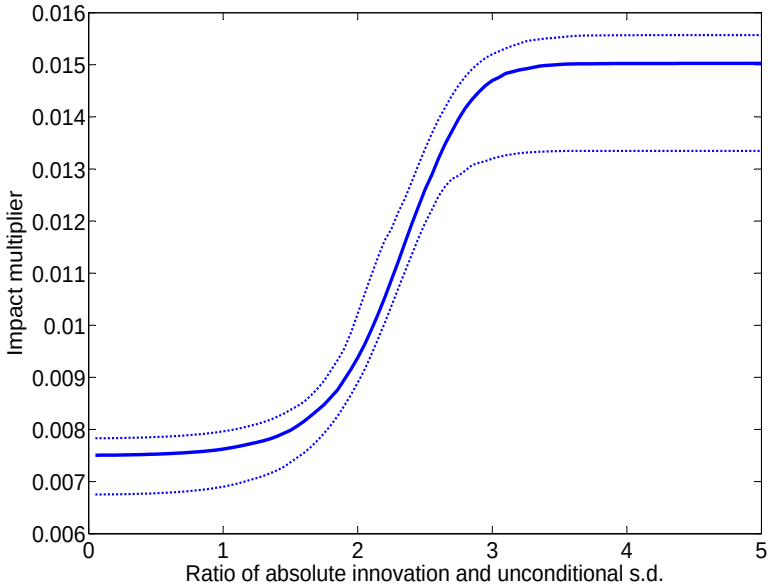

FIGURE 4. Expected output multiplier on innovation to productivity conditional on magnitude of innovation.

Since the probability of generating a man-bites-dog signal is increasing in the absolute size of the shock, the larger a shock is, the more likely it is that the response to the shock will be described by the blue impulse responses. Small shocks on the other hand are more likely to generate the weaker responses. This conditionality of the magnitude of the impulse responses is illustrated in Figure 4, where the expected impact multiplier (y-axis) of a productivity shock on output is plotted for different ratios (x-axis) of the size of the shock relative to its unconditional standard deviation. We can see in the figure that the expected impact is larger for larger shocks. For productivity innovations smaller than one standard deviation in magnitude, inflation and output will almost surely respond according to the black impulse responses. Innovations larger than three standard deviations will almost surely generate

the stronger responses illustrated by the blue lines. The man-bites-dog information structure thus introduces a non-linearity in the endogenous variables' responses to exogenous shocks with the consequence of making both recessions and booms sharper than they would otherwise be.

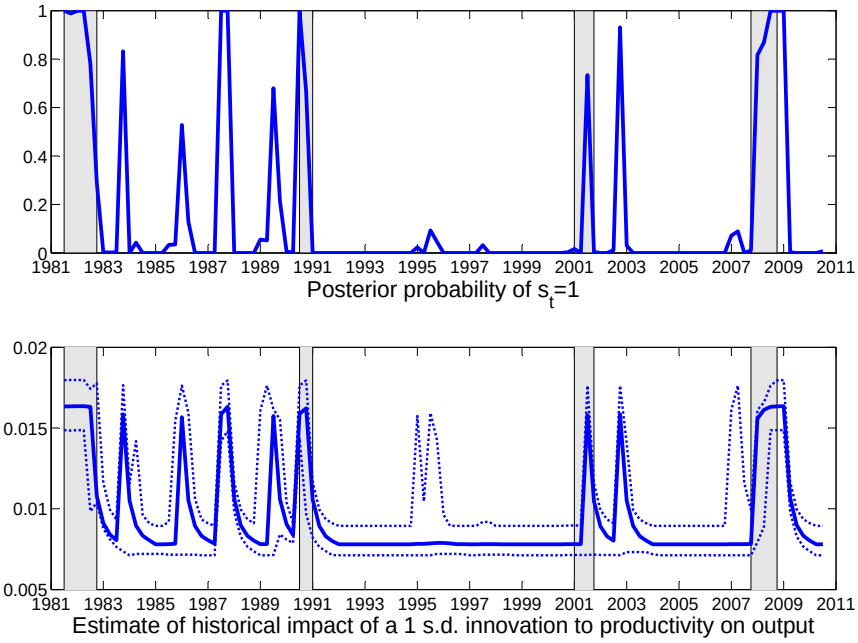

FIGURE 5. Posterior historical probability of $s_t = 1$ (top panel) and posterior estimate of historical impact of a 1 (unconditional) s.d. innovation to productivity on output with 95 per cent probability interval.

5.2. **Historical man-bites-dog episodes.** In the top panel of Figure 5 the posterior probabilities of man-bites-dog events are plotted against the sample dates with shaded areas representing NBER dated recessions. There are four episodes (some of them lasting several quarters) that have an estimated probability of a man-bites-dog event close to unity: The recession in the early 1980s, the 1987 stock market crash, the 1991 recession and the current "great recession". In addition to these major events, there are a handful of episodes for which the probability of a man-bites-dog event is estimated to be larger than 50 per cent. Of those, three are in the 1980s, one coincides withe the 2001 recession and one occur in the last quarter of 2002.

As suggested by the theory, the quarters that are assigned a high probability of being a man-bites-dog episode are also quarters with larger than average (in absolute terms) productivity innovations. The average absolute innovation from the quarters assigned a probability larger than 10 per cent is 0.0068 compared to 0.0052 for the full sample. The corresponding

numbers for quarters with a probability of a man-bites-dog event above 50 and 90 per cent are 0.0076 and 0.0092, respectively.

There appear to be a slight bias towards quarters associated with negative productivity innovations (relative to trend) to be assigned a large probability of being man-bites-dog episodes. This bias becomes more pronounced the more stringent one defines a man-bites-dog episode. For instance, of the 18 quarters in which the probability of a man-bites-dog event is estimated to be larger than 50 per cent, 14 are negative and 4 are positive innovations (as computed at the mode of posterior estimate of ρ_a). For the quarters with a probability larger than 90 per cent, the numbers are 9 negative and 1 positive innovation. This may suggest that bad news are perhaps more newsworthy, though an alternative explanation could be that most of the largest (in absolute terms) innovations in the sample are negative (the 3 largest innovations are all negative). Another feature of the estimates that perhaps makes this bias appear stronger than what is actually the case is that recessions typically feature several quarters in a row that have both negative innovations and are assigned a high probability of being a man-bites-dog episode. If we were to count episodes rather than quarters, then there are 5 episodes that are associated with bad news, the four recessions and the 1987 stock market crash (though the latter did not coincide with a large negative productivity innovation). The three non-recession non-stock market crash episodes in the 1980s are all associated with positive innovations to productivity.

5.3. **Interpreting historical man-bites-dog episodes.** In the introduction, the *Movers* segment on Bloomberg Television was given as an example of man-bites-dog news reporting. One can think of other types of news events that are more likely to affect the macro economy than the movements of individual stock prices. For instance, the news that a governor of an important central bank held an unscheduled press conference could certainly be interpreted as increasing the likelihood of unusual macroeconomic developments. If little concrete information was given beyond reassuring words of "policy vigilance" etc, this may very well increase uncertainty. Still, it seems unlikely that individual events, perhaps with the exception of the 1987 stock market crash, directly causes what is identified as man-bites-dog episodes by the model. A more plausible interpretation of these episodes is the following. At certain times the economy for various reasons become one of the main news stories, dominating network news and newspaper front pages. According to the man-bites-dog dictum, this should be more likely to happen when macro economic developments are in some sense unusual. Arguably, one reasonable reaction to intense media focus on the economy is to ask "Why are all the newspapers talking about the economy? Maybe something unusual is going on." and thus making people more sensitive to news about the economy.

The interpretation of man-bites-dog events as periods when economic issues dominate the news also motivates modeling man-bites-dog signals as public signals. If unusual economic events are more likely to dominate the news media as compared to more common place events, it seems reasonable to think of signals about these type of events as being a closer approximation to the theoretical ideal of a public signal in the strong common knowledge sense of the word. It seems clear that, for instance, there are more people who knows approximately how much the stock market fell in the crash of 1987, than who knows what

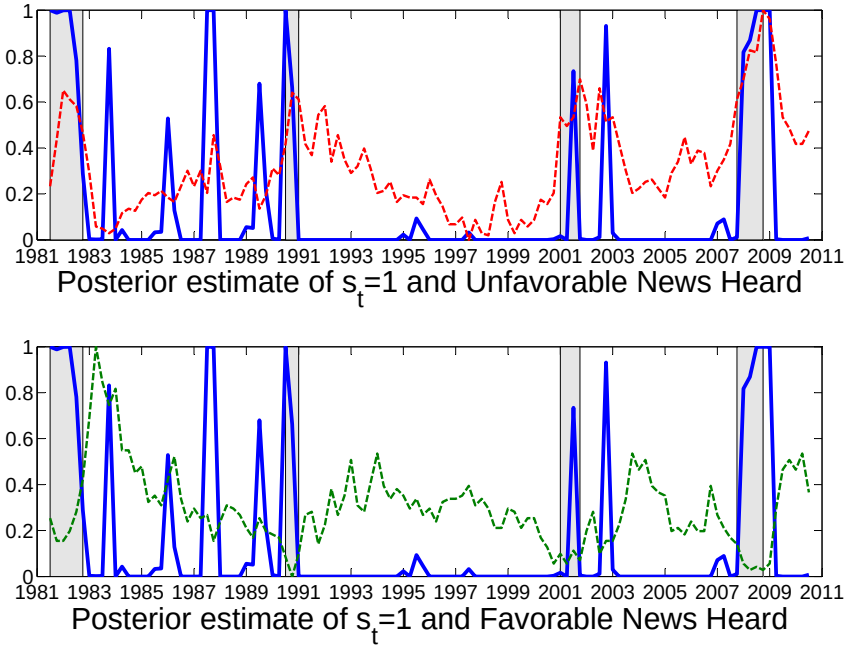

FIGURE 6. Posterior historical probability of $s_t = 1$ (blue lines) and Unfavorable News Heard (red dashed line, top panel) and Favorable News Heard (green dashed line, bottom panel). Data on News Heard are from Michigan Survey and are re-scaled to range between zero and one.

the Dow Jones index did in 2005, though both pieces of information are clearly publicly available.

One way to check more directly whether what the model interprets as man-bites-dog events are indeed related to the intensity of news coverage we can compare the posterior probabilities that $s_t = 1$ with the fraction of respondents in the Michigan Survey that have heard unfavorable or favorable news "during the last months". This data was not used in estimation and thus provides an independent check on the reasonableness of the estimates produced by the model. Figure 6 contains the posterior probabilities of $s_t = 1$ and the number of respondents that have heard unfavorable news (top panel) and unfavorable news (bottom panel). The numbers are re-scaled to lie in the range between 0 and 1. It is clear that the unfavorable news heard index increases around recessions and around the stock market crash in Q4 of 1987. Of the remaining episodes that are assigned a high probability of being a man-bites-dog event, but that are not associated with these events, two out of three are associated with a significantly higher than normal index of "favorable news heard". There is thus only one episode in late 1989 that is assigned a high probability of being a man-bites-dog event but where there is little movement in the either the unfavorable or

favorable news index. The fact that the model identifies episodes that are indeed associated with a higher news penetration as man-bites-dog events should increase our confidence in it.

5.4. **Endogenous persistence in volatility.** The model estimates suggest that man-bites-dog events are associated with stronger responses to productivity innovations. The posterior estimate of the history of the output multiplier on productivity innovations is plotted in the bottom panel of Figure 5. There it can be seen that when there is man-bites-dog event, the impact multiplier increases sharply, but only returns back to a more normal level gradually. That is, a single man-bites-dog signal increases the sensitivity of output to productivity shocks for several periods. The reason is that if a man-bites-dog signal is observed in period t then this will affect the posterior covariance of the state estimation error in period t. As explained above, the posterior covariance in period t translates into the prior covariance in period $t+1$, implying that the man-bites-dog signal in period t has persistent effects on the law of motion of the state, through its effect on the sensitivity of agents to new information. These effects can be quite long-lived, depending on the parameters and generally last longer when uncertainty is relatively large and when the persistence of the exogenous state is high.

Figure 6 illustrates the output multiplier on a productivity shock the periods after a man-bites-dog signal is observed. That is, Figure 7 plots the posterior estimate of the relevant elements of

$$G(s^t)N(s^t) : t = 1, 2, ... \tag{5.1}$$

from the model solution (4.10) - (4.11) for

$$s^t = 1, 0, 0, 0, 0, ... \tag{5.2}$$

That is, Figure 6 plots an impulse response of the time varying parameter that governs how much output increases after a one standard deviation innovation to productivity. As can be seen from the figure, the impact a productivity shock has when a man-bites-dog signal is generated is persistently higher after the impact period. The man-bites-dog information structure thus generates something similar to ARCH dynamics or stochastic volatility in the endogenous variables even though no such persistence is present in the volatility of the exogenous productivity process. (A picture similar to Figure 6 can also be plotted for the inflation multiplier.) It is worth noting that the positive autocorrelation observed in Figure 6 is an empirical finding, not a necessary implication of the man-bites-dog mechanism. For instance, if the man-bites-dog signal was estimated to be very precise, posterior covariances would decrease in period t, making agents less sensitive to new information in period $t+1$ and we would see a negative correlation between the impact multiplier in period t and period $t+1$.

Coibion and Gorodnichenko (2011) document that expectations are updated faster, in the sense that average expectations in surveys respond proportionally stronger, during more volatile periods. Though Coibion and Gorodnichenko are looking at a different sample period and focus on low frequency movements, this is exactly the qualitative prediction made by the model presented here: A man-bites-dog event increases the volatility of macro aggregates *because* expectations are updated faster, since a man-bites-dog signal makes agents put more weight on new information.

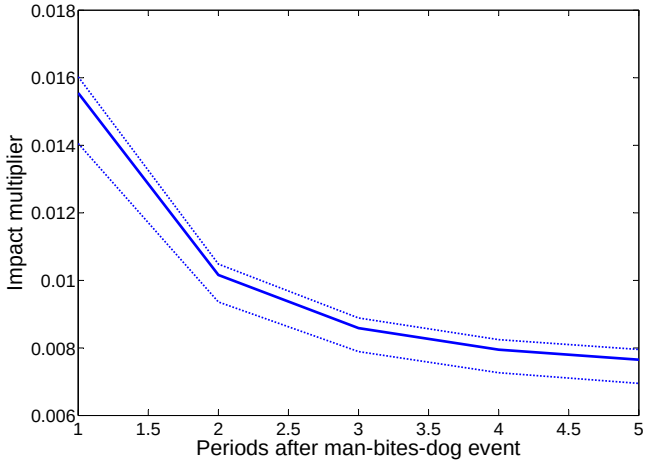

FIGURE 7. Illustration of persistent increase in endogenous volatility after man-bites-dog event.

5.5. **Man-bites-dog signals and the cross-sectional dispersion of expectations.** As was shown in the static setting of Section 2, uncertainty can either increase or decrease when a man-bites-dog signal is observed relative to when it is not. Whether this effect is present depends on the value of the parameter γ governing how unusual an event has to be to significantly increase the probability of observing a man-bites-dog signal and how noisy the man-bites-dog signal is when it is observed. When γ is large or the man-bites-dog signal is very noisy, uncertainty is more likely to increase after a man-bites-dog signal and the cross-sectional dispersion of first order expectations increases. In Figure 8 we have plotted how the cross-sectional dispersion of first order expectations changes after a man-bites-dog signal at the posterior mode of the estimated model. It can be seen that at the mode, the uncertainty effect of a man-bites-dog signal appear to dominate and the cross-sectional dispersion increases at the time of the man-bites-dog signal. Afterwards, the dispersion gradually returns to the more concentrated distribution associated with periods when there are no man-bites-dog events in the relevant history of s_t.

These results can be contrasted to those of Kondor (2010). He shows that in a setting where two classes of agents are constrained in what type of private information they can acquire, a public signal may increase the dispersion between first and second order expectations. In the model presented here, a man-bites-dog signal decreases dispersion between different orders of expectation (not shown) but increases the cross-sectional dispersion of first order expectations.

6. ALTERNATIVE SPECIFICATIONS

In this section I investigate the implications of changing the empirical specifications in two different ways. First, I analyze how the empirical estimates change when the cross-sectional component is excluded from the sample data and the model is instead estimated using only

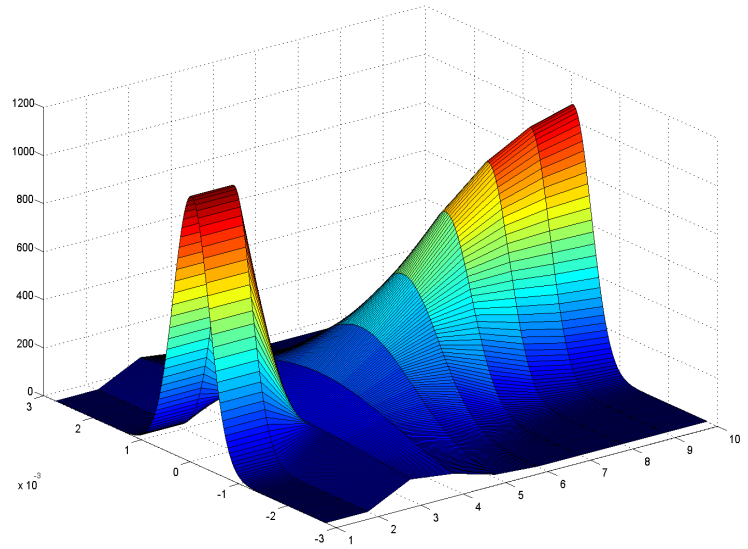

FIGURE 8. Impulse response of cross sectional distribution of expectations after a man-bites-dog signal, i.e. $s^t = 0, 1, 0, 0, 0, ...$ (Cross sectional dispersion increases (widens) in period 2 when $s_2 = 1$.)

mean forecasts from the Survey of Professional Forecasters. Second, I re-estimate the model allowing man-bites-dog signal to be more likely to be observed when there has been a large innovation to demand, rather than to productivity.

6.1. **The information in the cross-section of the survey data.** In the benchmark specification above, we used up to 50 survey observations in a given period to estimate the parameters of the model as well as the historical probabilities that $s_t = 1$. Since there is quite a bit of time-variation in the cross-sectional dispersion in the SPF, and since changes in the cross-sectional dispersion of expectations in the model are caused by man-bites-dog episodes, it may be interesting to to try to gauge the influence of the cross-sectional dimension of the data on the estimates of s_t . In order to investigate the influence of using the cross-sectional information, the model was re-estimated using the within period cross-sectional average forecast of inflation and nominal GDP as noisy indicators of the average expectations in the model. In addition to the standard parameters of the model, the standard deviation of the measurement errors associated with the average forecasts then also need to be estimated.

There are three main implications of this change. First, and unsurprisingly, credible intervals for the posterior parameter estimates become wider as the the number of indicators is decreased. Second, the only episodes that are identified by the model as having a probability of being a man-bites-dog event of more then 50 per cent are the recession in the early 1980s and the current recession. The cross-sectional dimension thus appears to play a crucial role in identifying historical man-bites-dog episodes. Third, the difference between the impulse

responses to a productivity innovation with and without a man-bites-dog signal becomes much larger. At the posterior mode, the impact multiplier on output is 4 times larger when there is a man-bites-dog signal available compared to when there is not. This last point is perhaps natural, given that only the two most extreme events in the sample are identified as man-bites-dog episodes. The difference between man-bites-dog episodes and normal times are then naturally also found to be larger.

6.2. **Man-bites-dog demand shocks.** In the benchmark specification, large innovations to productivity are more likely to generate a man-bites-dog signal. An alternative specification is to let the man-bites-dog signals be about demand shocks so they are more likely to be observed when there has been a large innovation to aggregate demand. When the model is estimated under this alternative specification, there are two major changes. First, the fit in terms of the log likelihood at the posterior mode is significantly worse, on the order of 400 log points. Secondly, the posterior estimate of s_t now suggest that there were up to 40 man-bites-dog episodes in the sample, more or less evenly distributed over the sample period. Based on the evidence from the Michigan Survey referred to above, it seems like the model with man-bites-dog demand episodes is capturing something that operates at a different frequency compared to the News Heard indices plotted in Figure 6. One possible explanation for this difference is that we do not have an observable time series of the demand disturbance, while we do treat TFP as observable in estimation. The model is thus less disciplined by the data in assigning probabilities of man-bites-dog events. Given the worse fit, implausible estimate of s_t and the that it is an econometrically less disciplined exercise, there seems to be little reason to prefer a specification with demand driven man-bites-dog episodes.

7. Conclusion

That some types of signals are more likely to be available about unusual events may suggest that we should be better informed about unusual events. However, in this paper we have showed that the flip-side of this argument is that the distribution of events conditional on there being a man-bites-dog signal available is different from the unconditional distribution in ways that may undo the effect of having more information. It was demonstrated that conditioning only on the *availability* of a man-bites-dog signal increases uncertainty, while observing the actual contents of the signal decreases uncertainty. If the likelihood of observing the signal only increases substantially after very unusual events or if the signal is sufficiently noisy, the former effect may dominate so that the posterior uncertainty after observing a man-bites-dog signal is larger than it would be if no such signal was available. We also showed that even though the man-bites-dog signal is public, the cross-sectional dispersion of (first order) expectations may increase.

In the second part of the paper, a simple business cycle model was presented in which large innovations to productivity is more likely to generate a public signal. When estimated, the model was used to quantify the potential effects of man-bites-dog signals on business cycle dynamics. According to the model, there were several episodes in recent US history in which the impact of an innovation to productivity on aggregate output was about twice as strong as the corresponding impact during normal times. The increased sensitivity of macro aggregates to productivity innovations were found to be persistent, lasting about 5

quarters after a single man-bites-dog event occurred. We also presented corroborative and independent evidence from the Michigan Survey that the episodes identified by the model as man-bites-dog events were indeed associated with higher than normal news coverage of the economy.

Arguably, some features of the model captures a "crisis mentality" in which there is an intense media focus on the economy. Yet, while there is more information produced and broadcast about the economy during recessions, uncertainty and sensitivity to new information appears to increase. While a majority of the episodes identified as man-bites-dog event by the model were indeed negative events, we also found several episodes in which a similar effect appears to be at work when economic news are favorable.

The model was estimated by likelihood based methods using both the quarterly total factor productivity time series constructed by Fernald (2010) and "raw" survey data from the Survey of Professional Forecasters along with more standard macro indicators. Using a time series of TFP as an observable variable has obvious advantages in terms of disciplining the model, especially since one of the aims of the paper has been to quantify the extent of time variation in the impact of TFP shocks on other variables. Using raw survey data allowed us to incorporate the information in the time variation in the second moment of survey responses into the posterior estimates of the parameters of the model. Particularly, we showed that the cross-sectional dimension in the Survey of Professional Forecasters is informative about the the timing of man-bites-dog events. In order to exploit the time variation in the cross-sectional dispersion of the survey data, it is necessary to have a model that can fit this fact and the paper makes a methodological contribution by demonstrating how a model with time varying information sets can be solved and estimated. This may be of separate interest to some readers.

Finally, it may also be worth pointing out a difference between the mechanism proposed here and models featuring stochastic volatility. In a stochastic volatility model, the variance of the exogenous shocks hitting the economy is a persistent time-varying process. In the model presented here, the sensitivity of endogenous variables to a given magnitude of shocks varies over time. While the man-bites-dog mechanism is more likely to be triggered by a large exogenous shock, the persistent increase in sensitivity of the endogenous variables to exogenous shocks is entirely endogenous. To the extent that we can observe the exogenous shocks directly, this distinction is a testable difference between the two approaches.

References

[1] Angeletos, G-M. and J. La'O, 2009, "Incomplete information, higher-order beliefs and price inertia", *Journal of Monetary Economics*, Volume 56, Supplement 1, ppS19-S37.

[2] Angeletos, G-M. and J. La'O, 2010, "Noisy Business Cycles", *NBER Macroeconomics Annual 2009*, Volume 24, pp319-378.

[3] Bachmann, R. and G. Moscarini, 2011, "Business Cycles and Endogenous Uncertainty", working paper, Yale University.

[4] Beaudry, P. and F. Portier, 2006, "Stock Prices, News, and Economic Fluctuations," American Economic Review, vol. 96(4), pages 1293-1307.

[5] Bloom, Nicholas,, 2009, "The Impact of Uncertainy Shocks", *Econometrica*, Vol. 77, No. 3, pp623–685.

[6] Carroll, Christopher D. 2003, "Macroeconomic Expectations of Households and Professional Forecasters", *Quarterly Journal of Economics* 118(1), pp269–298.

[7] Campbell, John C., 1994, "Inspecting the mechanism", *Journal of Monetary Economics*, 33, pp463-506.
[8] Chib, Siddhartha, 2001, "Markov Chain Monte-Carlo Methods: Computation and Inference", *Handbook of Econometrics*, Ch.57, vol 5, edited by J.J. Heckman and E. Leamer, Elsevier.
[9] Coibion, O. and Y. Gorodnichenko, 2011, "Information Rigidity and the Expectations Formation Process: A Simple Framework and New Facts", mimeo, UC Berkeley.
[10] Doms, M. and N. Morin, 2004, "Consumer sentiment, the economy, and the news media," Working Papers in Applied Economic Theory 2004-09, Federal Reserve Bank of San Francisco.
[11] Engle, Robert F, 1982, "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," *Econometrica*, vol. 50(4), pages 987-1007.
[12] Fernald, John, 2010, "A Quarterly, Utilization-Adjusted Series on Total Factor Productivity", working paper, Federal Reserve Bank of San Francisco.
[13] Fernández-Villaverde, J. and J. F. Rubio-Ramírez, 2010, "Macroeconomics and Volatility: Data, Models, and Estimation", NBER Working Papers 16618.
[14] Graham, L. and S. Wright, 2010, "Information, heterogeneity and market incompleteness", *Journal of Monetary Economics*, vol. 57(2), pp64-174.
[15] Haario, H, E. Saksman and J. Tamminen, (2001), "An Adaptive Metropolis Algorithm", *Bernoulli*, 7(2), 223-242.
[16] Jaimovich, N. and S. Rebelo, 2009, "Can News about the Future Drive the Business Cycle?", *American Economic Review*, vol. 99, issue 4, pp1097-1118.
[17] Kondor, P. 2009, "The more we know, the less we agree: higher-order expectations, public announcement and rational inattention", mimeo, Central European University.
[18] Lorenzoni, Guido, 2009, "A Theory of Demand Shocks," American Economic Review, American Economic Association, vol. 99(5), pages 2050-84, December.
[19] Mankiw, G.N., R. Reis and J. Wolfers, 2004, "Disagreement about Inflation Expectations," NBER Chapters, in: NBER Macroeconomics Annual 2003, Volume 18, pages 209-270.
[20] Mankiw, G. and R. Reis, 2002, "Sticky Information Versus Sticky Prices: A Proposal To Replace The New Keynesian Phillips Curve," The Quarterly Journal of Economics, vol. 117(4), pages 1295-1328.
[21] Melosi, Leonardo, 2010, "Public's Inflation Expectations and Monetary Policy", mimeo, London Business School.
[22] Morris, S. and H.S. Shin, 2002, "The social value of public information", *American Economic Review* 92, pp1521-1534.
[23] Nimark, K., 2008, "Dynamic Pricing and Imperfect Common Knowledge", *Journal of Monetary Economics*, pp365-382.
[24] Nimark, K., 2011, "Dynamic Higher Order Expectations", working paper, Universitat Pompeu Fabra.
[25] Primiceri, G., 2005, "Time Varying Structural Vector Autoregressons and Monetary Policy", *Review of Economic Studies,* vol 72, pp821-852.
[26] Reis, Ricardo, 2006a, "Inattentive Producers," Review of Economic Studies, vol. 73(3), pages 793-821,
[27] Reis, Ricardo, 2006b, "Inattentive consumers," Journal of Monetary Economics, vol. 53(8), pages 1761-1800.
[28] Sargent, Thomas J., 1991, "Equilibrium with Signal Extraction from Endogenous Variables", *Journal of Economic Dynamics and Control* 15, pp245-273.
[29] Stock, J.H. and M.W. Watson, 2003, "Has the Business Cycle Changed and Why?", *NBER Macroeconomics Annual 2002*, Volume 17, pp159-230.
[30] Swanson, Eric T., 2006a, "Have Increases in Federal Reserve Transparency Improved Private Sector Interest Rate Forecasts?," Journal of Money, Credit and Banking, vol. 38(3), pages 791-819.
[31] Swanson, Eric T., 2006b, "Optimal nonlinear policy: signal extraction with a non-normal prior", Journal of Economic Dynamics and Control, vol. 30(2), pp185-203.
[32] Townsend, Robert M., 1983, Forecasting the Forecasts of Others, *Journal of Political Economy*, vol 91, pp546-588.

Appendix A. Generating conditional heteroscedasticity

In Section 2 of the main text a man-bites-dog information structure was reversed engineered by directly specifying the conditional distributions $p(x \mid S)$ for $S = 0, 1$. Here we show that if more unusual events are more likely to be reported, this endogenously generates heteroscedastic conditional distributions an unconditionally normal distribution of x.

A.1. **Set up.** The basic set up is the following. There are two latent variables of interest; $x_1 \sim N(0,1)$ and $x_2 \sim N(0,1)$. We can interpret x_1 as being the state of the macro economy and x_2 as being "everything else". There exists what we may think of as a newspaper that observes noisy measure z_1 and z_2 of x_1 and x_2

$$z_1 = x_1 + \varepsilon_1 : \varepsilon_1 \sim N(0,1) \tag{A.1}$$

$$z_2 = x_2 + \varepsilon_2 : \varepsilon_2 \sim N(0,1) \tag{A.2}$$

The newspaper will report either a noisy signal of x_1 or x_2 but due to page constraints is assumed to be unable to report both. Instead, it will report a signal about the variable with the most unusual realization (in expectations) Since

$$E[x_i \mid z_i] = \frac{1}{2} z_i : i \in \{1, 2\} \tag{A.3}$$

it will report the signal y_1

$$y_1 = x_1 + \eta_1 : \eta_1 \sim N(0,1) \tag{A.4}$$

if $|z_1| > |z_2|$ and the signal y_2

$$y_2 = x_2 + \eta_2 : \eta_2 \sim N(0,1) \tag{A.5}$$

otherwise.

A.2. **Implied conditional distribution.** Let the variable S take the value 1 when y_1 is reported and zero otherwise. Rearranging Bayes' rule from (2.2) gives

$$p(x_1 \mid S = 1) = \frac{p(S = 1 \mid x_1)\, p(x_1)}{p(S = 1)} \tag{A.6}$$

which we can use to find the distribution of x_1 conditional on the availability of the signal y_1. With symmetric variables we have that $p(S = 1) = 0.5$ and $p(x_1)$ is by assumption the standard normal distribution. To find $p(S = 1 \mid x)$ note that

$$p(S = 1 \mid x_1) = p(|z_1| > |z_2| \mid x_1) \tag{A.7}$$

$$= \int_{-\infty}^{\infty} p(\varepsilon_1) \left[1 - 2\left(1 - \Phi\left(|x_1 + \varepsilon_1|\right)\right)\right] d\varepsilon_1 \tag{A.8}$$

where Φ is the cdf of z_2 which is zero mean normally distributed with variance 2. The quantity $1 - \Phi(|x_1 + \varepsilon_1|)$ is thus the probability that $z_2 > |z_1|$, the quantity $2(1 - \Phi(|x_1 + \varepsilon_1|))$ is the probability that $|z_2| > |z_1|$ so that $1 - 2(1 - \Phi(|x_1 + \varepsilon_1|))$ equals the probability that $|z_1| > |z_2|$ for given x_1 and ε_1. Since we only condition on x_1 we need to integrate over the support of ε_1 weighting each probability $[1 - 2(1 - \Phi(|x_1 + \varepsilon_1|))]$ by the pdf of ε_1. The expression (A.8) can be plugged into (A.6) to get the distribution of x_1 conditional on there

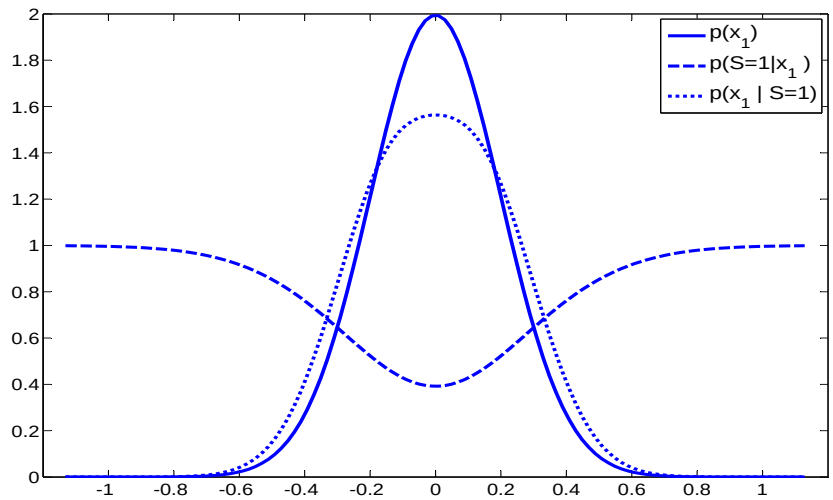

FIGURE 9. Conditional and unconditional distributions of x_1 and conditional probability of observing the signal y_1. (Note: Graph generated by scaling the standard deviation of all shocks by 0.2.)

being a signal y_1 available. All three distributions that make up the expression (A.7) are plotted in Figure 10.

We can see in the figure that the information structure has the man-bites-dog property that the conditional probability of observing a signal is increasing as realizations of x_1 further out in the tails are considered. Also, the distribution of x_1 conditional on $S = 1$ has fatter tails than the unconditional (normal) distribution of x_1. This simple example thus displays the same properties that are essential for the results in Section 2, though it is less tractable. For instance, there is no closed form solution available for the expectation of x_1 conditional on S and y_1. Readers with a particular interest in modeling filtering from fat-tailed priors may want to also look at Swanson (2006b) for a different approach in an optimal policy context.

APPENDIX B. PROOF OF PROPOSITION 2

Proposition 2. *The average expectation of x responds stronger to x when S=1 than when S=0.*

Proof. We need to show that the sum of the coefficients on the private signal x_j and public signal y in the conditional expectation

$$E\left(x \mid \Omega_j^1\right) = \frac{\sigma_\varepsilon^{-2}}{\sigma_\varepsilon^{-2} + \sigma_\eta^{-2} + \gamma^{-1}\sigma^{-2}} x_j + \frac{\sigma_\eta^{-2}}{\sigma_\varepsilon^{-2} + \sigma_\eta^{-2} + \gamma^{-1}\sigma^{-2}} y \tag{B.1}$$

when $S = 1$ is larger than the coefficient on the private signal

$$E\left(x \mid \Omega_j^0\right) = \frac{\sigma_\varepsilon^{-2}}{\sigma_\varepsilon^{-2} + \sigma^{-2}} x_j \tag{B.2}$$

when $S = 0$. Simply comparing the expected average expectation conditional on x for $S = 0$

$$\int E\left[x \mid \Omega_j^0\right] dx = \int \frac{\sigma_\varepsilon^{-2}}{\sigma_\varepsilon^{-2} + \sigma^{-2}} x_j dj \tag{B.3}$$

$$= \left(1 - \frac{\sigma^{-2}}{\sigma_\varepsilon^{-2} + \sigma^{-2}}\right) x \tag{B.4}$$

and $S = 1$

$$E\left[\int E\left(x \mid \Omega_j^1\right) dx \mid x\right] = \int \frac{\sigma_\varepsilon^{-2}}{\sigma_\varepsilon^{-2} + \sigma_\eta^{-2} + \gamma^{-1}\sigma^{-2}} x_j dj \tag{B.5}$$

$$+ \frac{\sigma_\eta^{-2}}{\sigma_\varepsilon^{-2} + \sigma_\eta^{-2} + \gamma^{-1}\sigma^{-2}} x \tag{B.6}$$

$$= \left(1 - \frac{\gamma^{-1}\sigma^{-2}}{\sigma_\varepsilon^{-2} + \sigma_\eta^{-2} + \gamma^{-1}\sigma^{-2}}\right) x \tag{B.7}$$

$$+ \frac{\sigma_\eta^{-2}}{\sigma_\varepsilon^{-2} + \sigma_\eta^{-2} + \gamma^{-1}\sigma^{-2}} \eta \tag{B.8}$$

means that the proposition is true if the inequality

$$\left(1 - \frac{\sigma^{-2}}{\sigma_\varepsilon^{-2} + \sigma^{-2}}\right) < \left(1 - \frac{\gamma^{-1}\sigma^{-2}}{\sigma_\varepsilon^{-2} + \sigma_\eta^{-2} + \gamma^{-1}\sigma^{-2}}\right) \tag{B.9}$$

holds. The last expression can with a little algebra be rearranged to

$$\gamma^{-1} < 1 + \sigma_\eta^{-2} \tag{B.10}$$

which is always true since $\gamma > 1$ and $\sigma_\eta^{-2} > 0$. □

Appendix C. Solving the model

This section describes how to find the equilibrium dynamics of the model. We first describe how the equilibrium relationships that can be used to pin down inflation and output and the law of motion for the state. We then describe an algorithm for finding a solution.

C.1. **Definitions and conjectures.**

First, define the exogenous state vector

$$\mathbf{x}_{j,t} \equiv \begin{bmatrix} a_{j,t} \\ d_{j,t} \end{bmatrix} \tag{C.1}$$

and note that

$$\int \mathbf{x}_{j,t} dj = \mathbf{x}_t \tag{C.2}$$

Island j's hierarchy of expectations is then given by

$$X_{j,t} \equiv \begin{bmatrix} \mathbf{x}_{j,t} \\ E\left[X_t^{(\bar{k}-1)} \mid \Omega_{j,t}\right] \end{bmatrix} \tag{C.3}$$

where

$$X_t^{(\bar{k}-1)} \equiv \begin{bmatrix} \mathbf{x}_t \\ \mathbf{x}_t^{(1)} \\ \vdots \\ \mathbf{x}_t^{(\bar{k}-1)} \end{bmatrix} \tag{C.4}$$

We will conjecture (and later verify) that consumption and inflation on island j can be written as linear functions of the true state X_t and island j's hierarchy of expectations $X_{j,t}$

$$\begin{bmatrix} c_{j,t} \\ \pi_{j,t} \end{bmatrix} = G_x\left(s^t\right) X_t + G_j\left(s^t\right) X_{j,t} + G_r r_{t-1} + G_u \mathbf{u}_t + G_\xi \begin{bmatrix} \xi_{j,t}^1 \\ \xi_{j,t}^2 \end{bmatrix} \tag{C.5}$$

where $\mathbf{u}_t$ is a vector that contains the aggregate shocks in the economy, i.e. those shocks that have a common effect on all islands. Note that this conjecture implies that aggregate consumption and inflation is given by

$$\begin{bmatrix} c_t \\ \pi_t \end{bmatrix} = \left[G_x\left(s^t\right) + G_j\left(s^t\right)\right] X_t + G_r r_{t-1} + G_u \mathbf{u}_t \tag{C.6}$$

since $\int \xi_{j,t}^n dj = 0$ for $n = 1, 2$. Use that $c_t = y_t$ and define $G(s^t) \equiv G\left(s^t\right) + G_j\left(s^t\right)$ so that inflation and output can be written as

$$\begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = G(s^t) X_t + G_r r_{t-1} + G_u \mathbf{u}_t \tag{C.7}$$

It will also be convenient to partition $G\left(s^t\right)$ into row vectors

$$\begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} G_y(s^t) \\ G_\pi(s^t) \end{bmatrix} X_t + G_r r_{t-1} + G_u \mathbf{u}_t \tag{C.8}$$

We also need to conjecture (and later verify) that the state X_t follows

$$X_t = M(s^t) X_{t-1} + N(s^t) \mathbf{u}_t : \mathbf{u}_t \sim N(0, I) \tag{C.9}$$

Solving the model implies finding the matrices $G\left(s^t\right), G_r, G_u, M(s^t)$ and $N(s^t)$.

C.2. **Inflation and output as functions of the state.** We can use the Euler equation for island j consumption

$$c_{j,t} = E\left[c_{j,t+1} \mid \Omega_t^j\right] - r_t + E\left[\overline{\pi}_{\mathcal{B}j,t+1} \mid \Omega_t^j\right] + d_{j,t} \tag{C.10}$$

and the island j Phillips curve

$$\begin{aligned} \pi_{j,t} \;=\; & \lambda(1+\varphi\delta)\left(\pi_t - \pi_{j,t}\right) + \lambda c_{j,t} + \lambda\varphi y_t - \lambda 1(+\varphi) a_{j,t} \\ & + \beta E\left(\pi_{j,t+1} \mid \Omega_t^j\right) + \lambda \xi_{j,t}^1 + \lambda\varphi \xi_{j,t}^2 \end{aligned} \tag{C.11}$$

(where we use the definitions (3.15) and (3.16) of $\overline{p}_{\mathcal{B}j,t}$ and $y_{j,t}$) to find the matrices A, B, C, G_r and G_u in

$$\begin{bmatrix} c_{j,t} \\ \pi_{j,t} \end{bmatrix} = A \int E\left(\begin{bmatrix} c_{t+1} \\ \pi_{t+1} \end{bmatrix} \mid \Omega_{j,t} \right) + BX_t^j + CX_t + G_r r_{t-1} + G_u \mathbf{u}_t \tag{C.12}$$

A, B, C, G_r and G_u are then given by

$$A = \begin{bmatrix} 1 & 1 \\ 0 & \beta \end{bmatrix} \tag{C.13}$$

$$B = \begin{bmatrix} e_2 \\ -\lambda(1+\varphi)e_1 \end{bmatrix} + \begin{bmatrix} 0 \\ -\lambda(1+\varphi\delta)e_2 G_j\left(s^t\right) \end{bmatrix} + \begin{bmatrix} 0 \\ \lambda e_1 G_j\left(s^t\right) \end{bmatrix} \tag{C.14}$$

$$\begin{aligned} C &= \begin{bmatrix} \phi_\pi G_\pi(s^t) + \phi_y G_y(s^t) \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ \lambda(1+\varphi\delta)G_\pi(s^t) \end{bmatrix} \\ &\quad + \begin{bmatrix} 0 \\ \lambda\varphi G_y(s^t) \end{bmatrix} + \begin{bmatrix} 0 \\ -\lambda(1+\varphi\delta)G_\pi(s^t) \end{bmatrix} \end{aligned} \tag{C.15}$$

$$G_r = -\begin{bmatrix} \phi_r(1-\phi_r)^{-1} & (\lambda+\lambda\varphi)\,\phi_r(1-\beta\phi_r)^{-1} \end{bmatrix}' \tag{C.16}$$

$$G_u = -\begin{bmatrix} (1-\phi_r)^{-1} & (\lambda+\lambda\varphi)\,(1-\beta\phi_r)^{-1} \end{bmatrix}' \begin{bmatrix} \sigma_r & 0 \end{bmatrix} e_r' \tag{C.17}$$

where G_r and G_u are simply given by direct forward substitution of the effect of the lagged interest rate r_{t-1} and the monetary policy shock u_t^r on future consumption and inflation rates, where the effect on inflation is indirect and works through how consumption affect marginal cost. The row vector e_r' picks out the element of $\mathbf{u}_t$ that corresponds to the monetary policy shock u_t^r in the Taylor rule (3.10).

C.2.1. *Aggregation.* If we are not interested in deriving the dynamics of prices on an individual island, we can proceed by taking averages across islands and use the conjectured form (C.7) to get

$$\begin{aligned} G\left(s^t\right) X_t &= (B+C)\,X_t \\ &\quad + \omega AG\left(s_1^{t+1}\right) M\left(s_1^{t+1}\right) HX_t \\ &\quad + (1-\omega)\,AG\left(s_0^{t+1}\right) M\left(s_0^{t+1}\right) HX_t \end{aligned} \tag{C.18}$$

The expression (C.18) uses that expected output and inflation is given by

$$\begin{aligned} E\left(\begin{bmatrix} y_{t+1} \\ \pi_{t+1} \end{bmatrix} \mid \Omega_{j,t} \right) &= \omega AG\left(s_1^{t+1}\right) M\left(s_1^{t+1}\right) HX_t \\ &\quad + (1-\omega)\,AG\left(s_0^{t+1}\right) M\left(s_0^{t+1}\right) HX_t \end{aligned} \tag{C.19}$$

where s_n^{t+1} denotes the history s^{t+1} with $s_{t+1} = n$. That is, the expectations of period $t+1$ inflation and output have to be weighted by the probability that will be a man-bites-dog signal available in the next period. Equating coefficients then implies that $G\left(s^t\right)$ must satisfy

$$\begin{aligned} G\left(s^t\right) &= (B+C) \\ &\quad + \omega AG\left(s_1^{t+1}\right) M\left(s_1^{t+1}\right) H \\ &\quad + (1-\omega)\,AG\left(s_0^{t+1}\right) M\left(s_0^{t+1}\right) H \end{aligned} \tag{C.20}$$

C.3. **The law of motion of the state.** We now describe how to find the law of motion for the state. We have conjectured above that it will take the form

$$X_t = M(s^t)X_{t-1} + N(s^t)\mathbf{u}_t : \mathbf{u}_t \sim N(0, I) \tag{C.21}$$

and the partly, this law of motion is exogenous. That is, the first two row are given by the law of motion for the exogenous aggregate states a_t and d_t

$$\mathbf{x}_t = \rho \mathbf{x}_{t-1} + \nu_t \mathbf{u}_t$$

where

$$\rho = \begin{bmatrix} \rho_a & 0 \\ 0 & \rho_d \end{bmatrix} \tag{C.22}$$

and

$$\nu_t = \begin{bmatrix} \sigma_a^2 & 0 \\ 0 & \sigma_d^2 \end{bmatrix} \text{ if } \sigma_t = 0 \tag{C.23}$$

$$\nu_t = \begin{bmatrix} \gamma\sigma_a^2 & 0 \\ 0 & \sigma_d^2 \end{bmatrix} \text{ if } s_t = 1 \tag{C.24}$$

C.3.1. *The filtering problem of agent j.* Since X_t is made up of agents' higher order expectations about $\mathbf{x}_t$, the endogenous part of $M(s^t)$ and $N(s^t)$ depend on the how agents update their higher order expectations. For given $M(s^t)$ and $N(s^t)$ and for given $G(s^t)$ the filtering problem of the agents can be described by the state space system made up of the law of motion (C.21) and the measurement equation

$$\mathbf{z}_{j,t} = D(s^t)X_t + \begin{bmatrix} R_u(s_t) & R_j(s_t) \end{bmatrix} \begin{bmatrix} \mathbf{u}_t \\ \mathbf{u}_{j,t} \end{bmatrix} \tag{C.25}$$

The matrices $D(s^t)$ and $R_u(s_t)$ and $R_j(s_t)$ vary over time and are given by

$$D(s^t) = \begin{bmatrix} e_1 \\ e_2 \\ G_\pi(s^t) \\ G_y(s^t) + G_\pi(s^t) \\ \phi_y G_y(s^t) + \phi_\pi G_\pi(s^t) \end{bmatrix}, R_u(s_t) = \begin{bmatrix} \mathbf{0}_{4\times 4} & \mathbf{0}_{4\times 1} \\ 0 & \sigma_r \end{bmatrix}, \tag{C.26}$$

$$R_j(s_t) = \begin{bmatrix} \sigma_{j,a} & 0 & 0 & 0 & 0 \\ 0 & \sigma_{j,d} & 0 & 0 & 0 \\ 0 & 0 & \sigma_{\xi 1} & 0 & 0 \\ 0 & 0 & 0 & \sigma_{\xi 2} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{C.27}$$

when $s_t = 0$ and

$$D(s^t) = \begin{bmatrix} e_1 \\ e_2 \\ G_\pi(s^t) \\ G_y(s^t) + G_\pi(s^t) \\ \phi_y G_y(s^t) + \phi_\pi G_\pi(s^t) \\ e_1 \end{bmatrix}, R_u(s_t) = \begin{bmatrix} \mathbf{0}_{4\times4} & \mathbf{0}_{4\times1} & \mathbf{0}_{4\times1} \\ \mathbf{0}_{1\times4} & \sigma_r & 0 \\ \mathbf{0}_{1\times4} & 0 & \sigma_\eta \end{bmatrix}, \tag{C.28}$$

$$R_j(s_t) = \begin{bmatrix} \sigma_{j,a} & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{j,d} & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{\xi1} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{\xi2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{C.29}$$

when $s_t = 1$. Agent j's state estimate

$$X_{j,t|t} \equiv E\left[X_t \mid \Omega_{j,t}\right] \tag{C.30}$$

evolves according to the Kalman update equation

$$X_{j,t|t} = M(s^t)X_{j,t-1|t-1} + K(s^t)\left[\mathbf{z}_{j,t} - D(s^t)M(s^t)X_{j,t-1|t-1}\right] \tag{C.31}$$

where the Kalman gain is given by

$$K(s^t) = \left[P(s^t)D'(s^t) + N'(s^t)R(s_t)\right]\left[D(s^t)P(s^t)D'(s^t) + R(s_t)R'(s_t)\right]^{-1} \tag{C.32}$$

where

$$R(s_t) \equiv \begin{bmatrix} R_u(s^t) & R_j(s^t) \end{bmatrix} \tag{C.33}$$

and $P(s^t)$ is defined as

$$P(s^t) = E\left(X_t - X_{j,t|t-1,s^t}\right)\left(X_t - X_{j,t|t-1,s^t}\right)' \tag{C.34}$$

with

$$X_{j,t|t-1,s^t} \equiv E\left(X_t \mid z_j^{t-1}, s^t\right)$$

i.e. the state estimation error covariance $P(s^t)$ is variance of error prior to observing $z_{j,t}$ but "posterior" to observing s_t. The state estimation error are thus conditionally normally distributed with covariance $P(s^t)$ given by

$$\begin{aligned} P(s^t) &= M(s^t)\begin{bmatrix} P(s^{t-1}) - \left[P(s^{t-1})D'(s^{t-1}) + N'(s^{t-1})R(s_t)\right]\times \\ \left[D(s^{t-1})P(s^{t-1})D'(s^{t-1}) + R(s_{t-1})R'(s_{t-1})\right]^{-1}\times \\ \left[P(s^{t-1})D'(s^{t-1}) + N'(s^{t-1})R(s_{t-1})\right]' \end{bmatrix} M'(s^t) \\ &\quad + N(s^t)N(s^t) \end{aligned} \tag{C.35}$$

(The term in brackets is the posterior state estimation error in period $t-1$.) The expressions for $K(s^t)$ and $P(s^t)$ are standard formulas for the Kalman filter recursions for state space systems with time varying (but known) parameters (see for instance Anderson and Moore 1979).

C.3.2. *The average expectation hierarchy.* Substituting in the expression for $\mathbf{z}_t(j)$ into (C.31) gives

$$\begin{aligned} X_{j,t|t} &= \left[I - K(s^t)D(s^t)\right] M(s^t)X_{j,t-1|t-1} \\ &\quad + K(s^t)D\left(s^t\right)\mathbf{X}_t + K(s^t)R\left(s^t\right)\begin{bmatrix} \mathbf{u}_t \\ \mathbf{u}_{j,t} \end{bmatrix} \end{aligned} \tag{C.36}$$

Taking averages across agents and using law of motion for X_t to get gives

$$\begin{aligned} X_{t|t} &= \left[I - K(s^t)D(s^t)\right] M(s^t)X_{t-1|t-1} \\ &\quad + K(s^t)D\left(s^t\right)\left[M(s^t)X_{t-1} + N(s^t)\mathbf{u}_t\right] \\ &\quad + K(s^t)R_u\left(s_t\right)\mathbf{u}_t \end{aligned} \tag{C.37}$$

since $\int \mathbf{u}_t(j)dj = 0$. Amending the result to the exogenous law of motion for $\mathbf{x}_t$ results in the law of motion for the entire system

$$\begin{aligned} \begin{bmatrix} \mathbf{x}_t \\ X_{t|t} \end{bmatrix} &= \begin{bmatrix} \rho & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \mathbf{x}_{t-1} \\ X_{t-1|t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ K(s^t)D(s^t)M(s^t) \end{bmatrix}\begin{bmatrix} \mathbf{x}_{t-1} \\ X_{t-1|t-1} \end{bmatrix} \\ &\quad + \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \left[I - K(s^t)D(s^t)\right]M(s^t) \end{bmatrix}\begin{bmatrix} \mathbf{x}_{t-1} \\ X_{t-1|t-1} \end{bmatrix} \\ &\quad + \begin{bmatrix} \nu(s_t) \\ K(s^t)D(s^t)N(s^t) \end{bmatrix}\mathbf{u}_t + K(s^t)R_u\left(s_t\right)\mathbf{u}_t \end{aligned} \tag{C.38}$$

For each s^t we thus need to find a fixed point of

$$M\left(s^t\right) = \begin{bmatrix} \rho & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ K(s^t)D(s^t)M(s^t) \end{bmatrix} + \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \left[I - K(s^t)D(s^t)\right]M(s^t) \end{bmatrix} \tag{C.39}$$

$$N\left(s^t\right) = \begin{bmatrix} \nu(s_t) \\ K(s^t)D(s^t)N(s^t) \end{bmatrix} + K(s^t)R_u\left(s_t\right) \tag{C.40}$$

C.4. **Algorithm for finding the solution.**

(1) Start by making initial guesses for the $\mathcal{S}$ different versions of the matrices $G\left(s^t\right), M(s^t)$ and $N(s^t)$. A good initial guess is to set all of them equal to the full information solution.
(2) For given matrices $G\left(s^t\right), G_r, G_u, M(s^t)$ and $N(s^t)$ compute $\mathcal{S}$ "new" $G\left(s^t\right)$ using (C.20). (That is, one need to loop through the $\mathcal{S}$ different matrices $G\left(s^t\right)$.
(3) For given matrices $G_r, G_u, M(s^t)$ and $N(s^t)$ and the "new" $G\left(s^t\right)$ compute $\mathcal{S}$ "new" $M(s^t)$ and $N(s^t)$ using (C.39) and (C.40)
(4) Iterate on 2 and 3 until convergence.

To keep track of the $\mathcal{S}$ different versions of the matrices $G\left(s^t\right), M(s^t)$ and $N(s^t)$ it is helpful to use the following indexing strategy. Define $m \times n \times \mathcal{S}$ arrays such that each "slice" of the array correspond to the matrices $G\left(s^t\right), M(s^t)$ and $N(s^t)$ for a given history s^t. Each unique finite history of s_t can be expressed as a decimal number using the mapping between binary and decimal numbers. The history

$$s^t = \{0, 0, 0, 1\} \tag{C.41}$$

will thus be assigned the index number 1 since the binary number 0001 equals 1 in the decimal system. Similarly, the history

$$s^t = \{1, 1, 0, 1\} \tag{C.42}$$

is given the index number 13 since the binary number 1101 equals 13 in the decimal system, and so on. The number of relevant histories of s_t equals 16 in this case where it is assumed that the last last 4 periods of s_t are relevant for current dynamics. If the number of different relevant histories equal 8, we have $\mathcal{S} = 2^8$ so that we would need to use 8 bit binary numbers to keep track of the 256 different relevant histories of s_t and so on.