Understanding Uncertainty Shocks and the Role of the Black Swan

Anna Orlik and Laura Veldkamp [*]

June 28, 2013

Abstract

A key question in macroeconomics is what shocks drive business cycles. A recent literature explores the effect of uncertainty shocks. But where do uncertainty shocks come from? Researchers often estimate stochastic volatility, using all available data, and equate it with uncertainty. If we define uncertainty to be the conditional variance of a one-period-ahead forecast error, then this stochastic volatility corresponds to the uncertainty of an agent who knows his model's parameters, well before they could ever be estimated, and is certain they are true. We show that a Bayesian forecaster who revises model parameters in real time and accounts for model uncertainty, experiences large, counter-cyclical uncertainty shocks, even if his model is homoskedastic. The exercise teaches us that large uncertainty shocks need not come from exogenous changes in variance. They may also come from "black swans": Events that are unlikely under the previous period's estimated model and cause agents to significantly revise their beliefs about the probability distribution of future outcomes.

Some times feel like uncertain times for the aggregate economy. At other times, events appear to be predictable, volatility is low, confidence is high. An active emerging literature argues that changes in uncertainty can explain asset pricing, banking crises, business cycle fluctuations, and the 2007 recession in particular (Stock and Watson, 2012). Uncertainty shocks are typically measured as innovations in a GARCH or stochastic volatility model, forecast dispersion or a price of a volatility option. But none of these measures captures

1

an essential source of uncertainty, that people do not know the true model of the world. People in 1950 could not not estimate a model on a 65-year post-war data sample, as an econometrician does when inferring 1950 volatility today. In 2006, financial analysts' models told them that a nationwide decline in house prices was nearly impossible. Those models proved incorrect. Models are selected and estimated with error. We use real GDP data and Bayesian methods to measure uncertainty of a forecaster who considers errors in his model and its parameters to be a source of uncertainty. Even when this forecaster uses a model which rules out the possibility of changes in the volatility of its innovations, uncertainty shocks still arise from the learning process itself.

Our forecasting model allows us to explore what kinds of event generate the largest uncertainty shocks and why. It reveals that the largest shocks to uncertainty often come from events we dub "black swans." Under the probability distribution estimated based on the previous period's information set, these events would be highly improbable. Upon observing such an event, Bayes' law leads our forecaster to place a lower probability weight on his previous forecasting model and more weight on models that would make the observed event less unlikely. This shift of probability weight among various models raises model uncertainty. Thus, the black swan causes our forecaster to be less certain about what stochastic process generates the data he sees, which is an uncertainty shock.

We define macroeconomic uncertainty as the variance of next-period GDP growth y_{t+1}, conditional on all information observed through time-t: $Var[y_{t+1}|\mathcal{I}_t]$. We use this definition because in most models, this is the theoretically-relevant moment. When there is an option value of waiting, beliefs with a higher conditional variance (imprecise beliefs) raise the value of waiting to observe additional information. Thus, it is uncertainty in the form of a higher conditional variance that typically delays consumption or investment and thus depresses economic activity.

A key premise of our paper is that we should use an explicit model to infer what uncertainty is. Why not just estimate a stochastic volatility model or compute forecast dispersion and call this uncertainty? Indeed, these measures would be uncertainty for an agent with a particular class of forecasting models. By deriving each measure from a forecasting model, this paper makes explicit the set of assumptions on which each measure is based. For example, the volatility of innovations would be equal to uncertainty for an agent

who knows that his model is the true GDP process and knows its true parameters. But the macroeconomy is not governed by a simple, known model and we surely don't know its parameters. Instead, our forecast data (from the survey of professional forecasters) suggests that forecasters estimate simple models to approximate complex processes and constantly use new data to update our beliefs. In such a setting, uncertainty and volatility can behave quite differently.

Two examples illustrate why uncertainty and volatility might differ. In the first example, uncertainty can change with constant volatility. When a shock hits that is low probability under the agent's current model and parameters, it will make the agent shift probability weight to other models and/or parameters. Being less certain about the data-generating process raises uncertainty about future GDP growth. In the second example, volatility rises but the agent does not know that innovations are being drawn from a different distribution. It may take many periods for the agent to learn that volatility has risen. Because the agent is unaware that volatility has increased, his uncertainty remains low. This paper builds and estimates models that capture these kinds of effects. By comparing model-generated measures of uncertainty to its proxies, we can learn about just how different volatility, forecast dispersion and uncertainty are.

Since our uncertainty estimates are model-dependent, we examine a sequence of simple forecasting models and use their ability to match features of professional forecast as a model selection criterion. We use only homoskedastic shocks so that all our uncertainty shocks come from the belief updating process. To build up intuition for how the model works, we start with one of the simplest settings: a 2-state regime switching model of GDP growth. Then, we start adding features to the model to make its predictions resemble those of professional forecasters. Each period t, our forecaster observes time-t GDP growth and uses the complete history of GDP data to estimate the state and forecast the GDP growth in $t+1$. Since professional forecasters are not endowed with knowledge of model parameters, we allow our forecaster to re-estimate all the parameters of the model each period. The resulting uncertainty series has large, counter-cyclical shocks. Uncertainty rises 10-20% above its mean in every recession.

However, a linear model with normally distributed shocks also badly misses a key feature of the data: The average forecast is significantly lower than the average GDP

3

growth realization. For an unbiased forecaster with a linear model, this should not be the case. But if GDP growth is a concave transformation of a linear-normal underlying variable, Jensen's inequality tells us that expected values will be systematically lower than the realizations. Adding extra uncertainty, such as parameter uncertainty, amplifies this effect. Therefore, we explore a non-linear forecasting model and find that non-linearity alone generates modest uncertainty shocks. But combining non-linear forecasting with parameter uncertainty reveals that agents may experience large increases in uncertainty, particularly in recessions.

Finally, we notice that the GDP estimates have higher average errors than the forecasters do. One way to bring the model and data closer on this dimension is to give forecasters additional signals about future GDP. This captures the fact that forecasters examine a large amount of macro data before forming a forecast. Introducing additional signals also allows the model to speak to forecast dispersion and to achieve a closer match with the moments of the forecaster data. But because signals are normal and the model is still linear, it does not do any better at explaining uncertainty shocks than the models without signals.

We compare our model-based uncertainty series to commonly-used uncertainty proxies and find that it is less variable, but more persistent than the proxy variables. The best proxies are the price of a volatility option (VIX) or forecast dispersion. But neither achieves more than a 50% correlation with our uncertainty measure.

Related literature A new and growing literature uses uncertainty shocks as a driving process to explain business cycles (e.g., Bloom, Floetotto, Jaimovich, Sapora-Eksten, and Terry (2012), Basu and Bundick (2012) Christiano, Motto, and Rostagno (2012), Bianchi, Ilut, and Schneider (2012)), to explain investment dynamics Bachmann and Bayer (2012a), to explain asset prices (e.g., Bansal and Shaliastovich (2010), Pastor and Veronesi (2012)), and to explain banking panics (Bruno and Shin, 2012). These papers are complementary to ours. We explain where uncertainty shocks come from, while these papers trace out the economic and financial consequences of the shocks.[1]

More similar to our exercise is a set of papers that measure uncertainty shocks in various

---

[1]In contrast, Bachmann and Bayer (2012b) argue that there is little impact of uncertainty on economic activity.

4

ways. Bloom (2009), Baker, Bloom, and Davis (2012), Stock and Watson (2012), Jurado, Ludvigson, and Ng (2013) Fernandez-Villaverde, Kuester, Guerron, and Rubio-Ramirez (2012), Fernandez-Villaverde, Guerron-Quintana, Rubio-Ramirez, and Uribe (2011), Naka-mura, Sergeyev, and Steinsson (2012) and Born, Peter, and Pfeifer (2011) document the properties of uncertainty shocks in the U.S. and in emerging economies, while Bachmann, Elstner, and Sims (2012) use forecaster data to measure ex-ante and ex-post uncertainty in Germany. While our paper also engages in a measurement exercise, what we add to this literature is a quantitative model of how and why such shocks arise.

The theoretical part of our paper grows out of an existing literature that estimates Bayesian forecasting models with model uncertainty. Cogley and Sargent (2005) use such a model to understand the behavior of monetary policy, while Johannes, Lochstoer, and Mou (2011) estimate a similar type of model on consumption data to capture properties of asset prices. While the mechanics of model estimation are similar, the focus on non-linear filtering, uncertainty shocks and uncertainty proxies distinguish our paper. Nimark (2012) also generates increases in uncertainty by assuming that only outlier events are reported. Thus, the publication of a signal conveys both the signal content and information that the true event is far away from the mean. Such signals can increase agents' uncertainty. But that paper does not attempt to quantitatively explain the fluctuations in uncertainty mea-sures. Our paper is in the spirit of Hansen's Ely Lecture (Hansen, 2007) which advocates putting agents in the model on equal footing with an econometrician who is learning about his environment over time and struggling with model selection.

1 A Forecasting Model

The purpose of the model is to explain how much and why uncertainty varies. In the model, an agent will observe a time-series of data and forecast the next period's realization of that time-series. In order to fully understand what elements of the model are responsible for which features of the uncertainty process, we build up the model in stages. In the first stage, the agent knows the model and its parameters. He uses Bayes' law to forecast the probability of being in each state tomorrow, which implies a GDP forecast. In the second stage, the agent knows the model, but not its parameters. Each period, he estimates the parameters and the state. For a problem with this many interdependent objects to

learn about, inference gets tricky. We need to use a Metropolis-Hastings algorithm to estimate the model objects. Those estimates allow the agent to forecast next period's productivity. His expected forecast error determines his degree of uncertainty. Finally, we add additional signals about GDP and uncertainty about whether innovations are normally or t-distributed. In each period, the agent sees a new piece of data, re-estimates states, parameters and the model probabilities, and as a result, is faced with a new degree of uncertainty.

1.1   Data Description

There are two pieces of data that we use to evaluate and estimate our forecasting model. The first is GDP data from the Bureau of Economic Analysis. The variable we denote y_t is the growth rate of GDP. Specifically, it is the log-difference of the seasonally-adjusted real GDP series, times 400, so that it can be interpreted as an annualized percentage change.

We use the second set of data, professional GDP forecasts, to evaluate our forecasting models. We describe below the four key moments that we use to make that assessment. The data come from the Survey of Professional Forecasters, released by the Philadelphia Federal Reserve. The data are a panel of individual forecaster predictions of real US output for both the current quarter and for one quarter ahead from quarterly surveys from 1968 Q4 to 2011 Q4. In each quarter, the number of forecasters varies from quarter-to-quarter, with an average of 40.5 forecasts per quarter.

Formally, $t \in \{1, 2, \ldots, T\}$ is the quarter in which the survey of professional forecasters is given. Let $i \in \{1, 2, \ldots, I\}$ index a forecaster and $I_t \subset \{1, 2, \ldots, I\}$ be the subset of forecasters who participate in a given quarter. Thus, the number of forecasts made at time t is $N_t = \sum_{i=1}^{I} \mathbb{I}(i \in I_t)$. Finally, let y_{t+1} denote the GDP growth rate over the course of period t. Thus, if GDP_t is the GDP at the end of period t, observed at the start of quarter $t + 1$, then $y_{t+1} \equiv ln(GDP_t) - ln(GDP_{t-1})$. This timing convention may appear odd. But we date the growth $t + 1$ because it is not known until the start of date $t + 1$. The growth forecast is constructed as $E_{it}[y_{t+1}] = ln(E_{it}[GDP_t]) - ln(GDP_{t-1})$.

A forecast of period-t GDP growth made at the start of period t, by forecaster i is denoted $E_{it}[y_t]$. Forecasters' forecasts will differ from the realized growth rate. This difference is what we call a forecast error.

6

Definition 1. *An agent i's forecast error is the distance, in absolute value, between the forecast and the realized growth rate: $FE_{i,t+1} = |y_{t+1} - E[y_{t+1}|\mathcal{I}_{it}]|$.*

We date the forecast error $t + 1$ because it depends on a variable y_{t+1} that is not observed at time t. Similarly, an average forecast error is

$$\bar{FE}_{t+1} = \frac{1}{N_t} \sum_{i=1}^{I_t} FE_{i,t+1}. \tag{1}$$

We compare forecast errors from the survey of professional forecasters to those from the model to select a forecasting model that performs well.

1.2 A General Forecasting Model

A model, denoted $\mathcal{M}$, is a probability distribution over a sequence of outcomes. We examine models with the following structure. Let $\{y_i\}_{i=0}^t$ denote a series of data available to the forecaster at time t. We postulate the following two-state hidden Markov regime switching process drives the dynamics of y_t[2]

$$y_t = \mu_{S_t} + \sigma e_t \tag{2}$$

where $e_{t+1} \sim \phi$. Models will differ in their probability distribution ϕ and in the information set available to forecasters. In every model, the information set will include y^t, the history of y observations up to and including time t. The state $S_t \in \{1, 2\}$ is a two-state Markov variable. State transitions are governed by an 2×2 transition probability matrix whose elements are $q_{ij} \equiv \Pr(S_t = j|S_{t-1} = i)$ with $\sum_j^N q_{ij} = 1 \; \forall i$. The transition matrix controls the persistence and stochastic evolution of the mean of the y_t process.

Each model has a vector θ of parameters. The elements of θ are the variance parameter σ, two state-dependent means μ_1, μ_2, and the transition probabilities $q_{11}, q_{12}, q_{21}, q_{22}$.

The state S_t is never observed. Thus (2) is the observation equation of a filtering problem. It maps an unobservable state S_t into an observable outcome y_t. The standard deviation of the unexpected innovations to this process is what we call volatility. Note that

---

[2]We have also explored a version of the model with a continuous hidden state. The results are very similar and are reported in Appendix B.

when we compute what is an unexpected innovation, we take the model and its parameters as given.

Definition 2. *Volatility is the standard deviation of the unexpected innovations to the observation equation, taking the model and its parameters as given:*

$$VOL_t = \sqrt{E\left[(y_{t+1} - E[y_{t+1}|y^t, \theta, \mathcal{M}])^2 \,\middle|\, y^t, \theta, \mathcal{M}\right]}.$$

The agent, who we call a forecaster and index by i, is not faced with any economic choices. He simply uses Bayes' law to forecast future y outcomes. Specifically, at each date t, the agent conditions on his information set $\mathcal{I}_{it}$ and forms beliefs about y_{t+1}. We call the expected value $E(y_{t+1}|\mathcal{I}_{it})$ and agent i's *forecast* and the square root of the conditional variance $Var(y_{t+1}|\mathcal{I}_{it})$ is what we call *uncertainty*.

Definition 3. *Uncertainty is the standard deviation of the time-$(t+1)$ GDP growth, conditional on an agent's time-t information:* $U_{it} = \sqrt{E\left[(y_{t+1} - E[y_{t+1}|\mathcal{I}_{it}])^2 \,\middle|\, \mathcal{I}_{it}\right]}.$

Volatility and uncertainty are both ex-ante measures because they are time-t expectations of $t + 1$ outcomes, which are time-t measurable. However, forecast errors are an ex-post measure because it is not measurable at the time when the forecast is made.

Substituting definition 1 into definition 3 reveals that $U_{it} = \sqrt{E_t[FE_{i,t+1}^2|\mathcal{I}_t]}$. So, uncertainty squared is the same as the expected squared forecast error. Of course, what people measure with forecast errors is not the *expected* squared forecast error. It is an average of *realized* squared forecast errors: $\sqrt{1/N \sum_i FE_{i,t+1}^2}$.

Definition 4. *The uncertainty generated by imperfect information about the model and its parameters (hereafter "process uncertainty") is $P = U - V$.*

The difference between volatility and uncertainty is due to the difference between conditioning on an agent's information set, and conditioning on $\{y^t, \mathcal{M}, \theta\}$. If the agent's information set includes the the model $\mathcal{M}$, parameters θ as well as the t-history of GDP observations, then process uncertainty is zero. But if the agent is uncertain about the model or parameters, U and V will differ and that difference in the uncertainty due to lack of knowledge of the true data-generating process. Note that unlike U or V, P can be, and sometimes is, negative. For example, suppose the highest-probability model has high-variance innovations, but the probability of that model is just over 50%. Then accounting

for model uncertainty will lead the agent to take some draws from the high-variance model and some from the low-variance model. The resulting set of outcomes may be less uncertain than they would be if the high-variance model were known to be the truth.

1.3 Forecasting with Only State Uncertainty (Model 1)

Many papers equate volatility, uncertainty and squared forecast errors. To understand when such assumptions are warranted, we begin with a setting where volatility and uncertainty are equivalent and both are equal to the square root of expected squared forecast errors. In this section only, we assume that the forecaster does know the model $\mathcal{M}$ and the parameters of that model θ.

Model 1 assumptions All agents have an identical information set: $\mathcal{I}_{it} = \{y^t, \theta, \mathcal{M}\}$, $\forall i$. Furthermore, the transitory innovations to the Markov process (2) are standard, normal variables: $e_{t+1} \sim N(0, 1)$.

By examining definitions 2 and 3, we can see that when $\mathcal{I}_{it} = \{y^t, \theta, \mathcal{M}\}$, uncertainty and volatility are equivalent. Thus, when a forecaster knows his model and its parameters with certainty, volatility measures uncertainty. But even in this model, neither is always equal to the forecast error since FE_{it} depends on a random $t+1$ outcome.

To compute the process for uncertainty, we use Bayesian updating to form the conditional variance in definition 3, as follows: Since the history of states $\{S_1, \ldots, S_t\}$ is not observed, the agent will use the Bayes rule to filter the hidden state. Each period, he will use the new y_t data to update his belief about which state might have been realized, in the following way. The agent is interested in computing the predictive data density $p\left(y_{t+1}|y^t\right)$ and using it to make following forecast

$$E\left(y_{t+1}|y^t\right) = \int y_{t+1} p\left(y_{t+1}|y^t\right) dy_{t+1}. \tag{3}$$

Let $\psi_{n,t} = \Pr\left(S_t = n|y^t\right)$ with $\sum_{n=1}^{N} \psi_{n,t} = 1$ be the forecaster's belief about state n in period

9

t. Then, the forecast is constructed as

$$
\begin{aligned}
E\left(y_{t+1}|y^t\right) &= \sum_{n=1}^{2} \Pr\left(S_{t+1} = n|y^t\right) \int y_{t+1} p\left(y_{t+1}|y^t, S_{t+1} = i, S_t = j\right) dy_{t+1} \\
&= \left[q_{11}\psi_{1,t} + q_{21}\left(1 - \psi_{1,t}\right)\right]\mu_1 + \left[q_{12}\psi_{1,t} + q_{22}\left(1 - \psi_{1,t}\right)\right]\mu_2 \qquad (4)
\end{aligned}
$$

where the last equality uses the fact that the expectation, conditional on the current and future Markov state, is $E\left(y_{t+1}|y^t, S_{t+1} = i, S_t = j\right) = \mu_i$.

The conditional variance is

$$
Var(y_{t+1}|y^t) = \left[q_{11}\psi_{1,t} + q_{21}\left(1 - \psi_{1,t}\right)\right]\left\{1 - \left[q_{11}\psi_{1,t} + q_{21}\left(1 - \psi_{1,t}\right)\right]\right\}\left(\mu_1 - \mu_2\right)^2 + \sigma^2
$$

The key source of uncertainty shocks is the variation in the posterior state belief $\psi_{i,t+1}$ The agent starts with a prior belief $\psi_{n,0}$ and forms posterior beliefs using Bayes law:

$$
\psi_{i,t+1} = \frac{\sum\limits_{j=1}^{N} p\left(y_{t+1}|y^t, S_{t+1} = i, S_t = j\right) q_{ji}\psi_{j,t}}{\sum\limits_{i=1}^{N}\sum\limits_{j=1}^{N} p\left(y_{t+1}|y^t, S_{t+1} = i, S_t = j\right) q_{ji}\psi_{j,t}} \qquad (5)
$$

Comparing model 1 to forecast data

If this model with state uncertainty only produces forecasts that closely resemble those in the forecast data, then that would support the idea that volatility and uncertainty are approximately interchangeable.

The parameters that are in the forecaster's information set come from estimating the model with maximum-likelihood, on the full sample of data.[3] The data is GDP growth rates 1968:Q4-2011:Q4. The procedure for computing forecasts and uncertainty is as follows: We endow the agent with knowledge of the parameters in table 1 and with the initial belief that there is an equal probability of being in each state ($\psi_{n,0} = 0.5$). Each period, the agent updates the current state belief using (5) and forecasts y_{t+1} using (4).

In the 2-state model, the agent believes that he is in the good (high-growth) state, most

---

[3]The appendix reports results where parameters come from estimating maximum likelihood parameters, using the data from 1947-1968.

Table 1: **Maximum Likelihood Parameters**

These are maximum likelihood estimates of (2), using real GDP growth data from 1968-2011.

Parameter	Value
μ_1	-2.47%
μ_2	3.55%
σ^2	$(2.9\%)^2$
q_{11}	0.69
q_{22}	0.96

of the time. When low GDP growth is realized, the agent starts to place more probability weight on the bad, low-growth state. Being less certain of which state currently prevails makes the forecaster's uncertainty rise. Because of the negatively skewed pattern of GDP growth, the agent puts high probability weight on the good state, which is frequently realized. This makes low-growth realizations more likely to generate increases in uncertainty. Thus a 2-state model can generate the pattern of counter-cyclical uncertainty observed in the data.

The problem with this model is that with only 2 states, the forecaster only attached probability weights to 2 possible growth estimates. The model doesn't allow him to adjust the states and attain more nuanced forecasts of future output. In the results in table 2, this shows up as a forecast error that is too large, too variable and too autocorrelated. Moreover, the forecasts don't track GDP growth very well. Their correlation is only 0.12, compared to 0.72 in the forecaster data. Furthermore, the agent's average GDP forecast is too low, causing the average forecast error to be much higher than in the data. So while the 2-state model provides simple laboratory for exploring how forecaster's update their beliefs in simple forecasting models, it is not the best framework for explaining how economic agents (professional forecasters) form expectations.

1.4 Do Agents Know Models and Parameters?

Obviously no one knows the true model of the economy in all its detail. But every model involves some fiction. The relevant question is how good or bad an approximation this assumption is for the purposes of measuring macroeconomic uncertainty. One way to quantify the importance of process uncertainty is to compare estimates of volatility

	data	state unc	param unc	signals	nonlinear
Avg forecast	2.24%	3.78%	2.72%	2.61%	2.24%
Avg forecast error	2.20%	2.53%	2.43%	1.96%	2.35%
Std forecast error	1.42%	2.48%	2.41%	1.40%	2.31%
Autocorr forecast error	0.28	0.28	0.17	0.13	0.15
Corr(forecast$_{t+1|t}$, GDP_{t+1})	0.71	0.12	0.25	0.74	0.27

Table 2: **Properties of forecasts in model and data.** Column 1 labeled 'state unc' uses equation (4) to forecast y_{t+1}. Columns 2, 3, 4 and 5 use equations (4), (7), (10) and (11) to forecast y_{t+1}. For the data and model 3 with heterogeneous signals, the standard deviation is the time-series standard deviation of the cross-section average, i.e. $std(1/N \sum_i |E_{it}[y_{t+1}] - y_{t+1}|)$. Similarly, the autocorrelation and correlation with GDP use the average forecast error or average forecast at each date t.

with mean-squared forecast errors. To see why this comparison is informative, imagine that forecasters know the true data-generating process for the macroeconomy. In that case, their forecast error is $\left(y_{t+1} - E[y_{t+1}|y^t, \theta, \mathcal{M}]\right)$ and the average squared forecast error is $(1/N) \sum_i \left(y_{t+1} - E[y_{t+1}|y^t, \theta, \mathcal{M}]\right)^2$. Recall that volatility is defined as VOL_t^2 $= E\left[\left(y_{t+1} - E[y_{t+1}|y^t, \theta, \mathcal{M}]\right)^2 \Big| y^t, \theta, \mathcal{M}\right]$, which is just the expected value of this sample average. If there is a large sample of forecasters and we apply the law of large numbers, then knowledge of the model and parameters would imply that mean-squared forecast errors and volatility are equal.

Computing the mean squared forecast error for the professional forecasts for GDP in each quarter is straightforward. In contrast, computing volatility depends on the model and parameters we condition on. Therefore, we consider three possible volatility estimates: The first is the estimate from our forecasting model with state uncertainty. The second is from our forecasting model with state, parameter and model uncertainty. The third volatility estimate is from an estimated GARCH model. Section 2.1 describes this model and estimation procedure in more detail.

What we find is that mean-squared forecast errors have a lower mean, more variation, and less persistence than any of our volatility measures. In particular, the fact that the coefficient of variation of the forecast error looks so much larger than the variation in volatility suggests that uncertainty shocks arise partly from changes in volatility, but partly from some other source. Something besides volatility makes forecasts more accurate in some times than in others. These stark differences between mean-squared forecast errors and

Table 3: **Comparing volatility and mean-squared forecast errors**

	Mean -squared forecast error	Volatility state unc	Volatility model unc	Volatility GARCH
mean	2.64%	3.34%	3.78%	3.54%
coeff var	0.58	0.10	0.32	0.37
autocorr	0.48	0.62	0.84	0.90

volatility leads us to believe that either agents know more than just the past data series and their true forecasting model, and/or that they do not know the true forecasting model. In reality, both must be true: professional forecasters use additional information to make their forecasts and they are uncertain about which model to use. The fact that forecasters have additional data shows up as a mean-squared forecast error that is lower, on average, than volatility. The forecasters' additional information is what allows them to make such precise forecasts. However, additional information alone does not make the forecast error have a large coefficient of variation, relative to volatility. Of course, one could assume that signals are more precise in some times than in others. But this gets back to assuming variations in uncertainty. Our learning model explains why forecast precision and uncertainty might vary.

How much fluctuation in forecast errors comes from small samples?   Of course, the difference between forecast errors and volatility could arise because of a small sample of forecasts. To quantify this effect, we create an artificial panel of forecasts with the same number of forecasters on average and homoskedastic error.[4] We then compute mean-squared forecast errors on this artificial panel. Of course, there is some variation period-to-period in the errors, even though volatility is fixed. But that variation is only 0.40, about two-thirds of what is observed in the data. In fact, among 10,000 runs of this artificial

---

[4]Our simulated GDP growth process is a series of i.i.d. normal variables with mean 1 and the same variance as true GDP growth. Signals are GDP growth plus two error terms $e + \eta$. Both errors are mean-zero, normal random variables. e is common to all forecasters and η is independent across forecasts. Forecasts are the posterior expectation of GDP growth. The variance of e and η are chosen to match the average mean-squared error and the average dispersion in forecasts, 0.66% and 0.39% respectively. Simulating 10,000 runs of 172 quarters each, with 35 forecasters, we found that the coefficient of variation of uncertainty averaged 0.40, with a minimum of 0.31, a maximum of 0.50, and a standard deviation of 0.026.

process, not one produced variation in uncertainty that is as large as what we see in the data. Even when we re-introduce heteroskedasticity in GDP by creating a panel of forecasts that are true GDP data plus homoskedastic noise, we only get a coefficient of variation of MSE's of 0.41, instead of 0.40. This is far short of the 0.58 coefficient observed in the data. What we learn from this exercise is that there is a component of uncertainty shocks that is unexplained by time-varying volatility, even after accounting for the noise introduced by the small sample of forecasts.

This fluctuation in forecast errors does seems to capture some time-variation in the ability to forecast accurately, not just random forecast noise.

1.5 Forecasting with State and Parameter Uncertainty (Model 2)

So far, we have assumed that the forecaster does not observe economic regimes but that he knows all the parameter values. That is obviously unrealistic because we used data from the entire sample period to estimate these parameters. A forecaster at any date before the last date in our sample would not have access to such information. Relaxing this assumption is important because if we want to understand uncertainty, part of uncertainty is not just whether there is a boom or recession (volatility), but also what the model and parameters are (process uncertainty). Solving a model where uncertainty about parameter values is a component of uncertainty helps us understand what role parameter uncertainty plays in producing uncertainty shocks. Therefore, the second forecasting model we consider is one where the agent forecasts using the same 2-state hidden Markov process (2), but the agent does not know any of the parameters θ and updates beliefs about them using Bayes' law.

Model 2 assumptions Each forecaster has an identical information set: $\mathcal{I}_{it} = \{y^t, \mathcal{M}\}, \forall i$. The model $\mathcal{M}$ is (2) with transitory innovations that are normally distributed: $e_{t+1} \sim N(0, 1)$.

Our forecaster starts with a prior distribution of parameters, $p(\theta)$, described in table 9. These priors are the maximum likelihood estimates from the data between 1968-2012.[5] Starting in quarter 4 of 1968, each period, the agent observes y_t, and updates his beliefs about parameters and states using Bayes' law. The posterior distribution $p(\theta, S^t|y^t)$ summarizes beliefs after observing the data $y^t = (y_1, y_2, ..., y_t)$. Using Bayes' rule, we can

---

[5]The appendix explores priors estimated on data from 1947-68.

rewrite $p\left(\theta, S^t | y^t\right) = p\left(\theta | y^t\right) p\left(S^t | y^t, \theta\right)$. Conditional on the history y^t and parameters θ, the agent can update, as in model 1 (equation 5). The challenge is to determine the posterior distribution of parameters $p\left(\theta | y^t\right)$. It a high-dimensional object whose dependence on the data is not necessarily linear.

To compute posterior beliefs about parameters, we employ a Markov Chain Monte Carlo (MCMC) technique.[6] At each date t, the MCMC algorithm produces a sample of possible parameter vectors, $\left\{\theta^d\right\}_{d=1}^{D}$, such that the probability of any parameter vector θ^d being in the sample is equal to the posterior probability of those parameters, $p\left(\theta | y^t\right)$. Therefore, we can compute an approximation to any integral by adding over sample draws: $\int f(\theta) p(\theta | y^t) d\theta \approx 1/D \sum_d f(\theta^d)$. Every parameter draw θ^d implies a probability of being in state i, denoted $\psi_{i,t}^d = \Pr\left(S_t = i | y^t, \theta^d\right)$. Thus, the forecaster can construct his forecast as

$$E\left(y_{t+1} | y^t\right) = \sum_{S^{t+1}} \int y_{t+1} p\left(y_{t+1} | \theta, S^{t+1}\right) p\left(\theta, S^{t+1} | y^t\right) d\theta \tag{6}$$

$$\approx \frac{1}{D} \sum_{d=1}^{D} \left[q_{11}^d \psi_{1,t}^d + q_{21}^d \left(1 - \psi_{1,t}^d\right)\right] \mu_1^d + \left[q_{12}^d \psi_{1,t}^d + q_{22}^d \left(1 - \psi_{1,t}^d\right)\right] \mu_2^d. \tag{7}$$

Comparing model 2 to forecast data Including parameter uncertainty is not only more realistic. It also helps the forecasting model to come closer to the data. Column 3 of table 2 shows that adding parameter uncertainty reduces the average forecast error and its volatility. Although the fit is improved, both moments are still far higher than what the data suggest. Parameter uncertainty did bring the autocorrelation of forecast errors in line with the data and it resulted in a doubling of the correlation of the forecast with GDP. Although, that correlation is still only one-third of what it is in the data.

1.6   A Model with Heterogeneous Signals (Model 3)

One feature of the forecaster data that the model so far does not speak to is the fact that there is heterogeneity in forecasts. A second aspect of forecasting is that obviously, forecasters have access to additional data, beyond just the past realizations of GDP that

---

[6]More details are presented in the Appendix. Also, see Johannes, Lochstoer, and Mou (2011) for a recursive implementation of a similar problem of sampling from the sequence of distributions.

we allow the forecaster in our models to observe. Additional data about GDP might take the form of leading indicators or firm announcements of future hiring and investment plans.

To model this additional information and the heterogeneity of forecasts, we consider a setting where multiple forecasters update beliefs as in the previous model with state and parameter uncertainty. But each period, each forecaster i observes an additional signal z_{it} that is the next period's GDP growth, with common signal noise and idiosyncratic signal noise:

$$z_{it} = y_{t+1} + \eta_t + \epsilon_{it} \tag{8}$$

where $\eta_t \sim N(0, \sigma_\eta^2)$ is common to all forecasters and $\epsilon_{it} \sim N(0, \sigma_\epsilon^2)$ is i.i.d. across forecasters.

We calibrate the two signal noise variances σ_η^2 and σ_ϵ^2 to match two moments. The idiosyncratic signal noise is chosen to match the average dispersion of forecasts. Forecast dispersion in time t is

$$D_t = \sqrt{\frac{1}{N_t} \sum_{i \in I_t} (E_{it}[y_{t+1}] - \bar{E}_t)^2} \tag{9}$$

where $\bar{E}_t = 1/N_t \sum_i E_{it}[y_{t+1}]$ is the average time-t growth forecast and N_t is the number of forecasters making forecasts at time t. So, σ_ϵ is chosen so that $1/T \sum_t D_t$ are equal in the model and in the forecaster data. The value of that average dispersion is 1.6%.

The common signal noise is chosen to match the average forecast error. Specifically, given a σ_ϵ, we choose σ_η so that the average forecast error, $1/T \sum_t FE_t$ is identical in the model and in the data. The value of that average forecast error is 1.91%. Note that this average error is larger than the dispersion. It tells us that signal noise is not exclusively private signal noise.

Our agent uses the following updating equation to form forecasts. (See appendix A.2 for details)

$$E\left(y_{t+1} | y^t, z_i^t\right) = \sum_{S_{t+1}} \int y_{t+1} p\left(y_{t+1} | \theta, S_{t+1}\right) \Pr\left(S_{t+1} | \theta, y^t, z_i^t\right) p\left(\theta | y^t, z_i^t\right) d\theta. \tag{10}$$

Comparing model 3 to forecast data While the model with heterogeneous signal allows us to talk about the part of forecast errors that come from forecast dispersion, the amount of dispersion does not drastically alter our predictions. In a model with must

16

less dispersion (0.4%), we find almost identical levels of uncertainty and correlations of uncertainty with GDP, but larger shocks to uncertainty.

But the fact that forecasters have signals about GDP growth makes a big difference. This model does a better job than any of the others in matching the average forecast error. But recall that we calibrated signal noise in order to match this moment of the data. Reducing the average error also brings down the standard deviation of forecast errors, to a level just slightly above that in the data.

The biggest success of the signal model is that it allows forecasts to be more highly correlated with future GDP growth. The correlation of 0.74 in this model is almost equal to the correlation of 0.72 in the data. While the other forecasting models do not have enormous signal errors because they are close to the true GDP growth number on average, they miss lots of the little ups and downs in GDP and therefore achieve low correlation. With a signal about future GDP, forecasters in this model are more likely to revise their forecasts slightly up when growth will be high and down when it will be low, achieving a higher correlation between forecast and GDP growth. But the bottom line is that building forecast heterogeneity in the model, similar to what we see in the data, is not a mechanism for generating large uncertainty shocks.

Should signal precision vary over time? One potential source of uncertainty shocks could be changes in the precision of signal. Here, we argue that this is an unlikely source of uncertainty shocks because it suggests other features of the data that are counter-factual. Suppose that in periods where the variance of the noise σ_η or σ_ϵ is high, z_t is a relatively poor predictor of y_{t+1}. Since agents' signal about y_{t+1} are low-precision, their uncertainty about y_{t+1}, conditional on this signal, will be high.

Of course, if all else is equal, a more volatile e_t would mean a more volatile y_t, and this story of forecasting variable precision changes would be the same as a story where uncertainty shocks come from volatility shocks. But it is possible that a fall in signal precision (increase in var(e_t)) could generate an uncertainty shock without a volatility shock to the GDP process y_t. For example, if z and e are independent, then $var(y_t) = var(z_t) + var(e_t)$, and a negative relationship between $var(z_t)$ and $var(e_t)$ could leave $var(y_t)$ unchanged. But this structure would imply that more uncertainty from less informative signals (high $var(e_t)$) is associated with lower macro volatility (low $var(z_t)$). There is no such inverse

relationship between uncertainty and the volatility of forecasting variables in the data.

There is another way that $var(y)$ could be constant, despite a shock to signal precision. If e_t and z_t themselves are negatively correlated, then $var(z)$ and $var(e)$ can both rise. Since $var(y) = var(z) + var(e) - 2cov(y, e)$, then a sufficiently high covariance will allow signal precision $1/var(e)$ to change, resulting in an uncertainty shock, without a volatility shock. But the problem with this explanation is that then z is no longer an unbiased forecast of y. If we transform z to make it an unbiased signal, then this negative correlation with the estimation error would disappear. Exploring these possibilities makes the point that it is hard to see how changes in the precision of forecast variables can explain uncertainty shocks, in a way that is consistent with the data.

1.7   Considering a Non-Linear Model of GDP Growth (Model 4)

So far, we have only explored variants of a linear model with normally-distributed innovations. These models all generated average forecasts that were roughly equal to average GDP growth in our sample. But professional forecasters' average forecast is 0.45% below annual average GDP growth. This fact suggests that either the forecasters are not Bayesian or that they are not using a linear model. Therefore, we explore whether a simple non-linear model of GDP growth might explain this fact.

We make this problem tractable by doing a change of measure. Since linear or non-linear doesn't have much meaning when there are only two states, we switch here to a continuous state model. We transform GDP growth into variable x, which is a continuous variable with a hidden persistent state.

$$\begin{aligned}
\widetilde{X}_t &= S_t + \sigma \epsilon_t \\
S_t &= \rho S_{t-1} + \sigma_S \varepsilon_t
\end{aligned}$$

where $\epsilon_t \sim N(0, \sigma_\epsilon^2)$ and $\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$.

The fact that forecasts are downward biased tells us that the mapping from $\widetilde{X}$ to y should be concave. Since the $\widetilde{X}$ variable is normal, it can take on positive or negative values. So we need a concave mapping that is defined over positive and negative real

18

numbers. One such function is a negative exponential:

$$y_t = c - exp(-\widetilde{X}_t) \qquad (11)$$

One economic interpretation of this change of measure is to think of $\widetilde{X}_t$ as an economic fundamental condition. When the economy is functioning very well (high $\widetilde{X}_t$), then improving its efficiency results in a small increase in GDP. But if there is a high degree of dysfunction or inefficiency (low $\widetilde{X}$), then the economy can easily fall into a deep depression. Most modern macroeconomic models are not linear. Many generate exactly this type of effect through borrowing or collateral constraints, other financial accelerator mechanisms, matching frictions, or information frictions. Even a simple diminishing returns story could explain such a concave mapping. A statistical justification for this mapping is that it recognizes the negative skewness present in GDP growth and chooses a skewed non-normal distribution for GDP growth.

What keeps the computation tractable, is that we assume that the forecaster knows the non-linear mapping. He is uncertain about the $\widetilde{X}$ process and what parameters govern it. But he knows the relationship between the variable $\widetilde{X}$ and GDP y. That assumption allows us to take the GDP data, apply the inverse transform in (11) and convert our data into $\widetilde{X}_t$ data. The parameter c in equation (11) affects the skewness of the resulting distribution of GDP growth. Therefore, we choose $c = 24$ so that the skewness of y matches the skewness of GDP growth in the data from 1952-1968:Q3.[7] Then, we can use linear updating MCMC techniques to form beliefs about the $\widetilde{X}_t$ process. For each parameter draw θ_i from the MCMC algorithm, we compute $E[y_t|\mathcal{I}_t, \theta_i]$ and $E[y_t^2|\mathcal{I}_t, \theta_i]$. We average these expectations over all parameter draws and compute uncertainty as $U_t = E[y_t^2|\mathcal{I}_t] - E[y_t|\mathcal{I}_t]^2$.

Comparing model 4 to forecast data   The previous models miss a key feature of the data: The average forecast is 2.2%, while the average GDP growth realization is 2.7%. For an unbiased forecaster with a linear model, this should not be the case. But if GDP growth is a concave transformation of a linear-normal underlying variable, Jensen's inequality tells us that expected values will be systematically lower than the median realization. But by

---

[7]We choose this time period because the agents would know this data before the first forecast is made in 1968. If we instead use the 1968-2012 sample, the results are nearly identical.

itself, Jensen's inequality does not explain the forecast bias because the expected GDP growth and the mean GDP growth should both have a Jensen term. It must be that there is some additional uncertainty in expectations, making the Jensen effect larger for forecasts than it is for the unconditional variance of the true distribution. This would explain why $E[y_{t+1}|\theta] > E[y_{t+1}|y^t]$. If the agent knew the true parameters, he would have less uncertainty about y_{t+1}. Less uncertainty would make the Jensen effect lower and raise his estimate of y_{t+1}, on average. Thus, it is the combination of parameter uncertainty and the non-linear updating model that can explain the forecast bias.

Column 5 of table 2 reveals that this model with its non-linear mapping between the state and GDP does lower the average growth forecast. The average growth forecast almost perfectly matches the average forecast in the survey of professional forecasters. This is not a guaranteed result since we calibrated the nonlinear transformation to match the distribution of GDP growth, not any moments of the forecaster data. Of course, this model also misses on some dimensions: It has forecast errors that are too high and forecasts that have too low a correlation with GDP. But the model with signals teaches us that adding informative signals can remedy these problems.
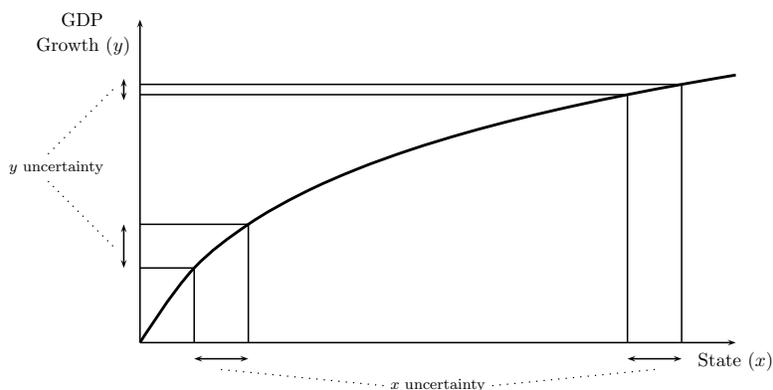

Figure 1: Nonlinear change of measure and state-dependent uncertainty. A given amount of uncertainty about x creates more uncertainty about y when x is low than it does when x is high.

The same concave change of measure that can explain the low GDP forecasts can also generate large uncertainty shocks. Figure 1 shows why. The concave line is mapping from x into GDP growth, y. The slope of this curve is a Radon-Nikodym derivative. A given amount of uncertainty is like a band of possible x's. (If x was uniform, the band would

represent the positive-probability set and the width of the band would measure uncertainty about x.) If that band is projected on to the y-space, the implied amount of uncertainty about y depends on the state x. When x is high, the mapping is flat, and the resulting with of the band projected on the y-axis (y uncertainty) is small. When x is low, the band projected on the y axis is larger and uncertainty is high.

Figure 2 illustrates the large fluctuations in uncertainty and volatility from the nonlinear model, compared to the other models. Volatility is not higher on average, but fluctuates, rising strongly in recessions. Uncertainty is much more persistent. It is highest at the start of the sample, when data is most scarce, and then slowly decays over the rest of the sample. The big low-frequency movements swamp the cyclical fluctuations to leave the correlation with GDP close to zero. But detrended uncertainty does rise in recessions.

Adding parameter uncertainty amplifies the largest volatility shocks. It is in these periods with large outlier GDP growth that agents question their model and its parameters. It takes many years after such a shock for agents to relearn the stochastic process that governs the economy.

1.8 Results: Model-Implied Uncertainty

Figure 2 compares the time series of uncertainty to the time series of volatility in each of our models.[8] In the state-uncertainty model (1), volatility and uncertainty are identical because the agent's information set includes the model and its parameters. In the other three models, U_t and V_t differ. In model 2, with parameter uncertainty only, the two series are still quite similar. At the start of the sample, the presence of unknown parameters raises uncertainty above what the model with state uncertainty alone predicts. It does so directly, but also indirectly, as the unknown parameters make the state harder to infer. But by the end of the sample, beliefs about parameters have largely converged and the uncertainty levels are more similar to the other models.

Adding signals (model 3) clearly lowers and smooths both uncertainty and volatility. The signals are an additional source of information that reduces the conditional variance of agents' estimates. Model 4 generates the largest level of uncertainty. The magnitude

---

[8]In each period t, volatility is computed as standard deviation of one-period ahead expectation of GDP growth conditional on a parameter vector (and a model) associated with a maximum likelihood of the data through to date t from the posterior parameter distribution under each model.
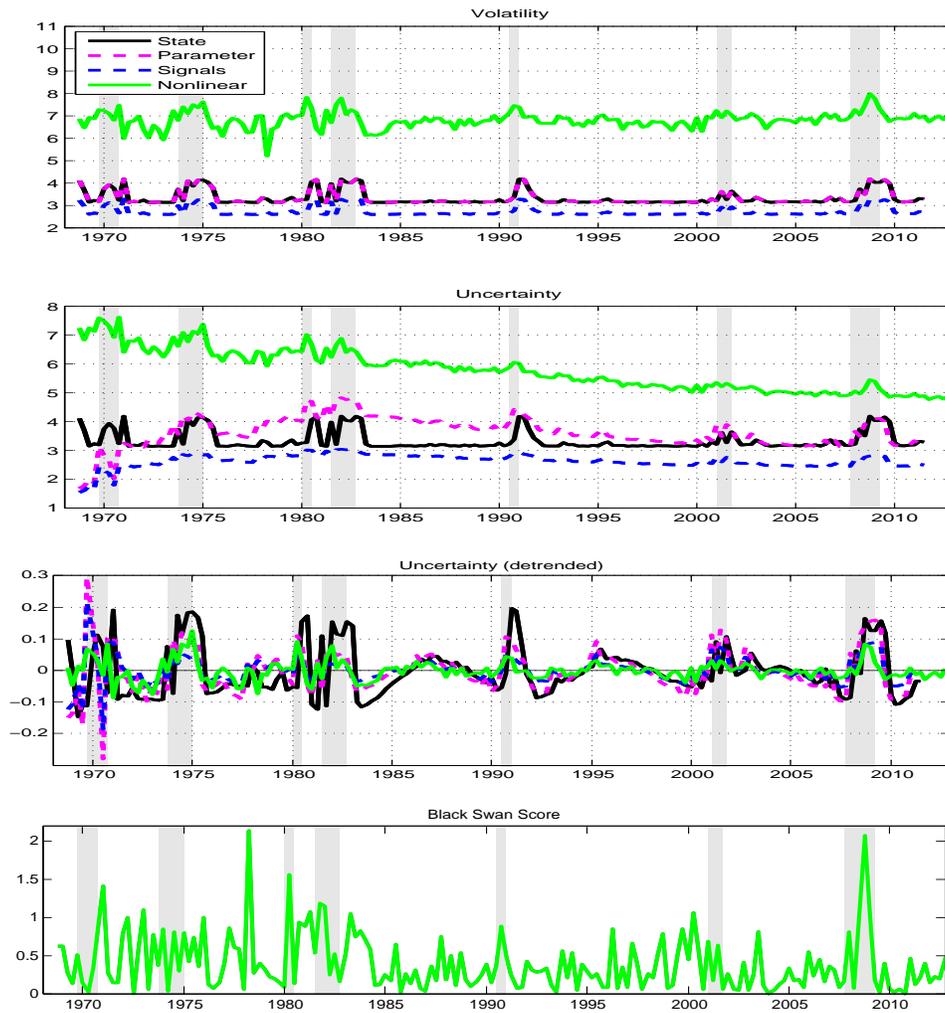
Figure 2: Volatility (V_t) and uncertainty (U_t) implied by each of the models and the black swan score from model 4. The top two panels plot the raw series, while the third shows the same uncertainty series in percentage deviation from trend. The fourth panel plots *Black Swan Score* as defined in (13).

of the shocks is large, but smaller relative to the trend level than the other models. An agent who believes that GDP growth is not normally distributed has a lot more scope for uncertainty.

To isolate uncertainty shocks, we remove the the low-frenquency changes in uncertainty, using a bandpass filter to filter out frequncies lower than once every 32 quarters. Then, we

compute a log deviation from this long-run trend.

$$\tilde{U}_t \equiv ln(U_t) - ln(U_t^{trend}) \tag{12}$$

The resulting series, plotted in the bottom two panels of figure 2, reveals large, highly counter-cyclical uncertainty shocks. In each of the recessions since 1968, uncertainty has risen sharply, 10-20% above trend.

Finally, we return to the idea that events which the forecasting model deems unlikely (black swan events) trigger uncertainty shocks. To see the relationship between uncertainty and unlikely events, we compute a black swan core for every GDP growth observation. The score is simply the number of standard deviations the realized GDP growth is from its forecasted level last period:

$$Black\ Swan\ Score_t = \frac{|y_t - E[y_t|y^{t-1}]|}{U_{t-1}}. \tag{13}$$

The last panel of figure 2 uses the non-linear model 4 to compute the expectation and U_t and report this score. Comparing the black swan score to the uncertainty measure in the plot above it reveals two regularities. First, most but not all spikes in uncertainty coincide with an unusual event. Second, some black swan events do not generate uncertainty shocks. In particular, quarter 2 of 1978 registered 15.4% annualized growth. Such a high growth rate was unanticipated by the model forecasts (and by most forecasters at the time). Thus, it registers as a high black swan score. However, while most unusual events are low-growth events, this is a high-growth outlier. Such a positive black swan events seems to generate much less uncertainty than the negative black swan. The asymmetric change of measure causes positive and negative black swan events to have very different consequenes.

Table 10 compares the moments of uncertainty and volatility from each of our four models. Typically, an econometrician computes volatility by estimating a model on the full sample of data, treating the parameters as true, and then calculating a conditional variance as $V_{t|T} = sqrtE[(y_{t+1} - E[y_{t+1}|y^t, \theta_T])^2|\mathcal{I}_T]$. For each model, that procedure yields an average uncertainty similar to V_t, but with a zero or low standard deviation, because this is a homoskedastic model. The V_t results shown are for V_t estimated using only real-time data. In other words, the θ the econometrician conditions on in definition 2

23

model		state unc	param unc	signals	nonlinear
		(1)	(2)	(3)	(4)
Mean	U_t	3.34%	3.62%	2.63%	5.79%
	V_t	3.34%	3.34%	2.73%	6.82%
Std deviation	U_t	0.32%	0.53%	0.23%	0.71%
	V_t	0.32%	0.32%	0.21%	0.37%
Coef of variation	U_t	0.10	0.15	0.09	0.12
	V_t	0.10	0.10	0.08	0.05
Autocorrelation	U_t	0.62	0.90	0.89	0.95
	V_t	0.62	0.62	0.62	0.47
detrended data moments					
Std deviation	$\tilde{U}_t$	7.77%	6.93%	4.21%	3.18%
	$\tilde{V}_t$	7.77%	7.77%	6.26%	5.11%
$\text{Corr}(\tilde{U}_t, y_t)$		-0.18	-0.54	-0.46	-0.90
$\text{Corr}(\tilde{V}_t, y_t)$		-0.18	-0.18	-0.19	-0.92
$\text{Corr}(\tilde{U}_t, y_{t+1})$		-0.10	-0.13	-0.09	-0.23
$\text{Corr}(\tilde{V}_t, y_{t+1})$		-0.10	-0.10	-0.10	-0.23

Table 4: **Properties of model uncertainty series.** Column 1 labeled 'state unc' uses equation (4) to forecast y_{t+1}. Columns 2, 3 and 4 use equations (7), (10) and (11) to forecast y_{t+1}. Volatilities are computing assuming that the true parameters θ are the maximum likelihood estimates, using all gdp data available at time t.

is the time-t maximum likelihood estimate of θ. For the model with only state uncertainty, uncertainty and volatility are identical, by definition. But for the remaining models, the raw volatility is typically less volatile and less persistent than raw uncertainty.

Uncertainty is a very persistent process, with both low-frequency changes, and fluctuations at the business cycle frequency. As noted by Collin-Dufresne, Johannes, and Lochstoer (2013), the persistent uncertainty process comes from the nature of learning: A single large shock to GDP growth results in a quick reevaluation of the parameter and model probabilities. These revisions in beliefs act as permanent, non-stationary shocks even when the underlying shock is transitory. Since using growth rates of GDP is a form of trend-removal, it only makes sense to correlate a stationary series with another stationary series. Therefore, we detrend volatility and uncertainty in order to discern that nature of their cyclical components. We find that both series are strongly counter-cyclical, particularly from the non-linear model.

2 Data Used to Proxy for Uncertainty

Our model generates an endogenous uncertainty series. Next, we'd like to compare our measure to some of the empirical proxies for uncertainty that are commonly used, including the VIX, forecast dispersion, and GARCH volatility estimates. Although these measures are all supposed to serve as proxies for uncertainty, they have different properties from each other and from the model-generated uncertainty series.

2.1 Time-varying volatility models

One common procedure for estimating the size of volatility shocks is to estimate an ARMA process that allows for stochastic volatility. In order to compare such a volatility measure to our uncertainty series, we estimate a GARCH model of GDP growth, that allows for time-variation in the variance of the innovations. All data is quarterly and all of these series are non-stationary. To obtain stationary series, we use annualized growth rates.

The GARCH process that generates the best fit is one with an AR(1) process for GDP growth (y) and a GARCH(1) process for volatility, which includes 1 lagged variance:

$$y_{t+1} = 3.38 + 0.41y_t + \epsilon_{t+1}, \quad \epsilon_{t+1} \sim \mathcal{N}(0, \sigma_{t+1}^2) \tag{14}$$

$$\sigma_{t+1}^2 = 0.52 + 0.76\sigma_t^2 + 0.24\epsilon_t^2 \tag{15}$$

We assume that the errors in our models, ϵ_t, are Gaussian and estimate the process using our full sample of data (1947-2012). The complete set of estimates, stationarity tests, as well as more detail about the model selection process are reported in appendix B. We also estimate homoskedastic models for each of the time series to test the hypothesis of homoskedasticity. We use the ARCH-LM test to do this. This estimated GARCH process produces an estimate of the shock variance σ_{t+1}^2 in each period. This is what we call the volatility process.

The evidence for time-varying volatility is weak. The log likelihood of the highest-likelihood heteroskedastic model is only 3% higher than the best-fitting homoskedastic model. With an ARCH-LM test, one cannot reject the null hypothesis of homoskedasticity (pvalue is 0.22). We come to a similar conclusion if we estimate the model, starting in 1947,

use a stochastic volatility model, or relax the distributional assumptions. Our baseline analysis assumes that errors ϵ_t have a Gaussian distribution. Using distributions with fatter tails (student-t) yields no difference in estimations or significance. Furthermore, we explore further lags of all variables; either coefficients were not significantly different from zero or the log-likelihood was reduced. Finally, we included different lags of linear terms for ϵ_t and variances σ_t^2 in the GARCH specification. Again, the estimated parameters were not significant or the log-likelihood was reduced.

2.2 Forecast dispersion

Some authors[9] use forecast dispersion (D_t in equation 9) as a measure of uncertainty. One of the advantages of this measure is that it is typically regarded as "model-free." It turns out that dispersion is only equivalent to uncertainty in a particular class of models. In that sense, it is not really model-free.

Whether dispersion accurately reflects uncertainty depends on private or public nature of information. Imprecise private information generates forecast dispersion, while imprecise public information typically does not. To understand the importance of this distinction, we decompose these forecast errors into their public and private components. Any unbiased forecast can be written as the difference between the true variable being forecast and some forecast noise that is orthogonal to the forecast. In other words,

$$y_{t+1} = E[y_{t+1}|\mathcal{I}_t] + \eta_t + e_{it} \tag{16}$$

where the forecast error ($\eta_t + e_{it}$) is mean-zero and orthogonal to the forecast. We can further decompose any forecast error into a component that is common to all forecasters η_t and a component that is the idiosyncratic error e_{it} of forecaster i. Using definition (3), we can write uncertainty as

$$U_{it} = E[(\eta_t + e_{it})^2|\mathcal{I}_t]$$

Since η and e are independent, mean-zero variables, this is

$$U_{it} = var(\eta_t) + var(e_{it})$$

---

[9]See e.g. Baker, Bloom, and Davis (2012) or Diether, Malloy, and Scherbina (2002), or Johnson (2004).

Dispersion depends on the squared difference of each forecast from the average forecast. We can write each forecast as $y_{t+1} - \eta_t - e_{it}$. Then, with a large number of forecasters, we can apply the law of large numbers, set the average e_{it} to 0 and write the average forecast as $\bar{E}[y_{t+1}|\mathcal{I}_t] = y_{t+1} - \eta_t$. Thus,

$$D_t = \frac{1}{N} \sum_i (E[y_{t+1}|\mathcal{I}_t] - \bar{E}[y_{t+1}|\mathcal{I}_t])^2 = \frac{1}{N} \sum_i e_{it}^2$$

If there is a large number of forecasters, each with the same variance of their private forecast noise $var(e_t)$, and we apply the law of large numbers, then

$$D_t = var(e_t).$$

Comparing the formulae for U_{it} and D_t reveals that they are the same under two conditions: 1) there is no common component in forecast errors $(var(\eta_t) = 0)$, and 2) private forecast errors have the same variance $(var(e_{it} = var(e_t), \ \forall i)$. Thus, equating forecast dispersion and uncertainty is implicitly using a forecasting model that has these two properties. Property 2 may be violated and then dispersion may still be a good measure of average uncertainty. If property 1 is violated, but the variance of public signal noise does not change over time $(var(\eta_t) = \bar{v}, \ \forall t)$, then dispersion will covary perfectly with uncertainty. While one can make such an assumption, it does raise the question why private information about the macroeconomy varies in its precision when public sources of information have a precision that is contant. As we will see in the next section, the forecast data does not support such an interpretation.

2.3 Mean-squared forecast errors

A measure related to forecast dispersion that captures both private and common forecast errors is the forecast mean-squared error.

We define a forecast mean-squared error (MSE_{t+1}) of a forecast of y_{t+1} made in quarter t as the square root of the average squared distance between the forecast and the realized value

$$MSE_{t+1} = \sqrt{\frac{\sum_{i \in I_t}(E_{it}[y_{t+1}] - y_{t+1})^2}{N_t}}. \tag{17}$$

If forecast errors were completely idiosyncratic, with no common component, then dispersion in forecasts and mean-squared forecasting errors would be equal. To see this, note that $FE_{jt}^2 = (E_{jt}[y_{t+1}] - y_{t+1})^2$. We can split up FE_{jt}^2 into the sum $((E_{jt}[y_{t+1}] - \bar{E}_t[y_{t+1}]) + (\bar{E}[y_{t+1}] - y_{t+1}))^2$, where $\bar{E}_t[y_{t+1}] = \int_j E_{jt}[y_{t+1}]$ is the average forecast. If the first term in parentheses is orthogonal to the second, $1/N \sum_j FE_{jt}^2 = MSE_t^2$ is simply the sum of forecast dispersion and the squared error in the average forecast: $(E_{jt}[y_{t+1}] - \bar{E}_t[y_{t+1}])^2 + (\bar{E}_t[y_{t+1}] - y_{t+1})^2$.

We can then use this insight along with our forecast data to evaluate the extent to which variation in mean-squared errors (MSE) comes from changes in the accuracy of average forecasts and how much comes from changes in dispersion. If most of the variation comes from changes in disperion, then perhaps the assumption that the common components of forecast errors are zero or negligible is not a bad assumption. But if most of the fluctuation in MSE comes from changes in average forecast errors, then using forecast dispersion as a proxy for uncertainty will miss an important sourcee of variation.

Table 5 examine forecasts of four different macro variables covered by the survey of professional forecasters. For each one, at least half of the variation comes from average forecasts. This finding tells us that the model which equates forecast dispersion with uncertainty, which would imply that dispersion and mean-squared error would be identical, is a model which is not well supported by the data. In other words, forecast dispersion may have very different properties from uncertainty.

	GDP	FedSpend	SLSpend	IntRt
R^2	80%	54%	97%	65%

Table 5: **The fraction of MSE variation explained by average forecast errors.** This is the R^2 of a regression of the forecast mean-squared error, MSE^2, defined in (17) on $(\bar{E}_t[y_{t+1}] - y_{t+1})^2$. The remaining variation is due to changes in forecast dispersion.

Forecast bias   Recent research argues that professional forecasts are biased. For example, Elliott and Timmermann (2008) argue that stock analysts over-estimate earnings growth and the Federal Reserve under-estimates GDP growth. But none of these findings suggest that the bias is volatile. In other words, forecast bias may explain why forecasts are further away from the true outcome. But uncertainty shocks come from fluctuations in the

expected size of forecast errors. A fixed bias does not create such fluctuations.

2.4 VIX and confidence measures

We discuss the volatility and forecast-based measures in most detail because there are measures that we can relate explicitly to our theory. Other proxy variables for uncertainty are interesting but have a less clear connection to our model. The market volatility index (VIX) is a traded blend of options that measures expected percentage changes of the S&P500 in the next 30 days. It captures expected volatility of equity prices. But it would take a rich and complicated model to link macroeconomic uncertainty to precise movements in the VIX. Nevertheless, we can compare its statistical properties to those of the uncertainty measure in our model. Figure 3 does just this.

Another commonly cited measure of uncertainty is business or consumer confidence. The consumer confidence survey asks respondents whether their outlook on future business or employment conditions is "positive, negative or neutral." Likewise, the index of consumer sentiment asks respondents whether future business conditions and personal finances will be "better, worse or about the same." While these indices are indeed negatively correlated with the GARCH-implied volatility of GDP, they are not explicitly questions about uncertainty. Furthermore, we would like to use a measure that we can compare to the forecasts in our model. Since it is not clear what macro variable "business conditions" or "personal finances" corresponds to, it is not obvious what macro variable respondents are predicting.

2.5 Comparing uncertainty proxies to model-generated uncertainty

Figure 3 plots each of the uncertainty proxies. There is considerable comovement, but also substantial variation in the dynamics of each process. These are clearly not measures of the same stochastic process, each with independent observation noise. Furthermore, they have properties that are quite different from our model-implied uncertainty metric. Table 6 shows that our uncertainty metric is less volatile, moderately counter-cyclical, but the raw uncertainty series is more persistent than the proxy variables.
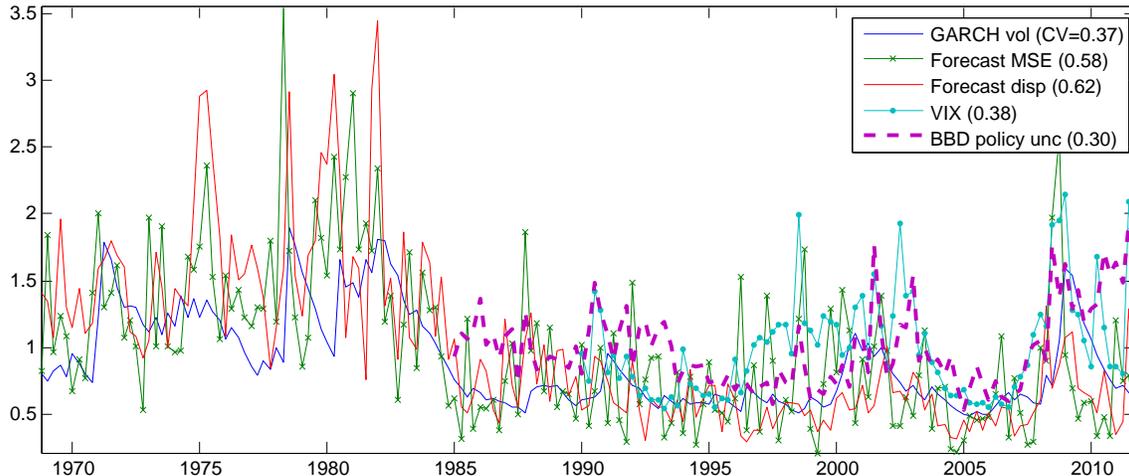
29

Figure 3: Comparing variables used to measure uncertainty in the literature.

2.6 Inferring uncertainty from probability forecasts

One way to infer the uncertainty of an economic forecaster is to ask them about the probabilities of various events. The survey of professional forecasters does just that. They ask about the probability that GDP growth exceeds 6%, is between 5-5.9%, between 4-4.9%, ..., between -1 and -2%, and below -2%. However, the survey only reports a single probability weight that is averaged across all forecasters.

Since this data does not completely describe the distribution of y_{t+1} beliefs, computing a variance requires some approximation. The most obvious approximation is to assume a discrete distribution. For example, when agents assign a probability to $1-2\%$ GDP growth, we treat this as if that is the probability placed on the outcome of 1.5% GDP growth. When the agent says that there is probability $p_{6.5}$ of growth above 6%, we treat this as probability $p_{6.5}$ placed on the outcome $y_{t+1} = 6.5\%$. And if the agent reports probability $p_{-2.5}$ of growth below -2%, we place probability of $p_{-2.5}$ on $y_{t+1} = -2.5\%$. Then the expected rate of GDP growth is $\bar{y} = \sum_{m \epsilon M} p_m m$ for $M = \{-2.5, -1.5, \ldots, 6.5\}$. Finally, the conditional variance of beliefs about GDP growth are $var[y|\mathcal{I}] = \sum_{m \epsilon M} p_m (m - \bar{y})^2$.

The resulting conditional variance series is not very informative. It hardly varies (range is $[0.0072, 0.0099]$). It does not rise substantially during the financial crisis. In fact, it

	Mean	Standard deviation	autocorr	correlation with g_{t+1}	correlation with $\tilde{U}_t$
forecast MSE	2.64%	1.53	0.48	0.04	-3.02%
forecast dispersion	1.54%	0.95	0.74	-0.19	19.98%
GARCH volatility	3.65%	1.35	0.90	0.06	7.26%
GARCH real-time	3.45%	1.04	0.90	-0.07	8.32%
VIX	20.55	7.81	0.58	-0.41	35.80%
BBD policy uncertainty	105.95	31.79	0.65	-0.41	20.95%
model 4 uncertainty ($\tilde{U}_t$)	0	0.03	0.41	-0.23	100%

Table 6: **Properties of forecast errors and volatility series for macro variables.** Forecast MSE and dispersion are defined in (17) and (9) and use data from 1968q4-2011q4. Growth forecast is constructed as $ln(E_t(GDP_t)) - ln(E_t(GDP_{t-1}))$. GARCH volatility is the σ_{t+1} that comes from estimating (15), using data from 1947q2-2012q2. At each date t, GARCH real-time is the σ_{t+1} that comes from estimating (15) using only data from 1947:Q2 through date t. VIX_t is the Chicago Board Options Exchange Volatility Index closing price on the last day of quarter t, from 1990q1-2011q4. BBD policy uncertainty is the Baker, Bloom, and Davis (2012) economic policy uncertainty index for the last month of quarter t, from 1985q1-2011q4. $\tilde{U}_t$ is model 4 uncertainty, which is measured as the log deviation from trend (eq. 12).

suggests that uncertainty in 2008 was roughly the same as it was in 2003. The reason this measure does not detect high uncertainty in exteme events is that the growth rates are top- and bottom-coded. All extremely bad GDP events are grouped in the bin "growth less than 2%." All the uncertainty was about how bad this recession might be. Instead, what the probability bins reveal is a high probability weight on growth below 2%. Since most of the probability is concentrated in one bin, it makes the uncertainty look low. The bottom line is that while using surveys to ask about ex-ante probabilities of GDP events is a promising approach to measuring uncertainty, the available data that uses this approach does not seem useful for our purposes.

2.7   Is uncertainty countercyclical?

Many authors argue that times of high uncertainty trigger recessions. Thus, it is useful to see how our measure of uncertainty and other uncertainty proxies are related to GDP. To be sure, many measures of uncertainty and volatility are higher in recessions than in expansions. But bear in mind that recessions, periods of negative growth, are outliers, where growth is more than 2% below average. This raises the question, is high uncertainty

associated with all unusual events, or primarily with low-growth events?

To answer this question, we do the typical breakdown of uncertainty measures in positive and negative growth periods. But then we do a similar split in above-median and below-median growth periods. Table 7 shows that most of the heightened uncertainty/volatility arises from outlier events. Lower than median growth raises uncertainty by less or even generates less uncertainty. When we extend the GARCH analysis to the entire post-war sample, the relationship between higher-than-median growth and uncertainty flips. Higher growth periods turn out to have higher average GARCH-measured volatility.

	Model 4 uncertainty	Professional Forecaster MSE	GARCH Volatility	GARCH Volatility ('47-'12)
Average conditional on positive GDP growth	-0.92%	2.44%	3.13%	3.65%
Average conditional on negative GDP growth	5.17%	3.72%	4.00%	4.26%
Average conditional on above median GDP growth	-1.96%	2.68%	3.13%	3.78%
Average conditional on below median GDP growth	1.84%	2.56%	3.38%	3.71%

Table 7: Uncertainty in high- and low-growth quarters.

Model, forecast MSEs and first volatility estimates come from data from 1968:Q4–2011:Q4 . Forecast mean-squared errors (MSE) are computed on survey of professional forecaster data, using (17). Second volatility column uses data from 1947q2 – 2012q2. Model 4 is applied to data from 1968q4-2012q4. The model output is measured in log deviation from trend.

Once we measure uncertainty with parameter uncertainty and non-normal innovations, we find much stronger evidence for counter-cyclical uncertainty. Note that the model uncertainty can be negative because it is measured as a log deviation from trend. Thus, on average, uncertainty is 5% below trend in recessions and 2% below trend when GDP growth is lower than its median.

3   Considering Policy Uncertainty

One aspect of uncertainty that our analysis so far has neglected is policy uncertainty. Perhaps the discrepancy between volatility and uncertainty can be explained by changes in the uncertainty about policy shocks. We consider two approaches to this comment. The first is to compare the properties of the Baker, Bloom, and Davis (2012) policy uncertainty

index to our GDP uncertainty measures, just like we do with all the other proxies for uncertainty. The second approach is to use the model to help think about the question: Where do policy uncertainty shocks come from?

Table 6 describes the summary statistics of the BBD policy uncertainty index. This index is a mix of legal changes, textual analysis, forecast dispersion. The nature of some of this data makes it impossible to derive the index from the model. But what we can say is that this index has shocks that are about as large in relative magnitude as our uncertainty measure. But model 4 uncertainty is much more persistent (0.95 vs. 0.65) and more counter-cyclical (-0.90 vs -0.41 correlation with GDP growth).

		Mean	Coeff Var (std/mean)	Sample Period
Fed Govt	forecast MSE	6.36%	0.53	1981:Q3-2012:Q1
Spending	volatility	5.90%	0.01	1981:Q3-2012:Q1
S&L Govt	forecast MSE	6.60%	0.62	1981:Q3-2012:Q1
Spending	volatility	2.36%	0.27	1981:Q3-2012:Q1
Interest	forecast MSE	1.00%	0.82	1981:Q3-2012:Q1
Rate	volatility	0.47%	0.65	1981:Q3-2012:Q1

Table 8: **Properties of forecast errors and volatility series for macro variables.** All variables are in growth rates, except for the interest rate, which is the first difference of the level. So a 1% mean for uncertainty means that the average forecast of the quarterly rate of growth has a 1% standard deviation, relative to the final reported value. Uncertainty denotes the square root of the mean squared error of current period growth forecasts. This growth forecast is constructed as $ln(E_t(x_t)) - ln(E_t(x_{t-1}))$.

Our second approach to thinking about policy uncertainty is to point out that the data susggests that something besides volatility contributes to policy uncertainty shocks. Recall that section 1.4 showed how the difference between mean-squred forecast errors and volatility was a measure of model uncertainty. Table 8 reports the moments of volatility and forecast errors for three policy variables: federal spending, state and local spending and interest rates. For the matching subsample of data, state and local spending has forecast MSE shocks that are 3.4 times larger (coefficient of variation is 0.62) than the shocks to volatility (coefficient of variation is 0.18). For federal spending, the differences in coefficients of variation is even larger (0.53 vs 0.01). This large gap between volatility shocks and forecast error shocks could have a couple of possible explanations. One is

that forecasts contain lots of public information whose quality is changing dramatically over time. Another possibility is that model and parameter uncertainty explains this gap. Estimation errors in model and parameters create time-variation in forecast errors, but not volatility.

For interest rates, the average size of forecast errors is very small. In other words, agents have very precise information about future interest rates. Of course, interest rates are lower in level than the size of government spending. But, keep in mind that all series are log-differenced. So, we are comparing percentage (log) changes in spending and GDP to changes in interest rates, which makes them comparable. Thus, the low level of uncertainty and volatility really reflects the fact that the degree of uncertainty about the other variables is much greater.

The conclusion we draw from this discrepancy is that, just like GDP uncertainty shocks do not seem to come primarily from volatility shocks, policy uncertainty shocks also do not seem to be driven by changes in policy volatility. They may well be a product of a learning process, which this paper explores.

4   Conclusions

The data typically used to measure uncertainty offer a muddy picture of what uncertainty shocks look like: Survey data reveals large swings in the ability of agents to make accurate forecasts. Yet, there is weak evidence of time-varying variance of macro aggregates and the estimated magnitude of these shocks is much smaller than what would explain the survey data. Uncertainty shocks appear to come not from the properties of the data being forecast, but from some state of mind of the forecaster himself.

Our model reconciles these findings. It takes the actual series of GDP as an input and generates forecasts with properties similar to those in the forecast data. But with this model structure, we can compute actual uncertainty – the conditional variance of the forecast of next quarter growth. We find that model uncertainty is a source of large, counter-cyclical uncertainty shocks that covary imperfectly with standard uncertainty proxies.

The model also helps us to understand where uncertainty shocks come from. People use simple models to forecast complex economic processes. They have to. Any model that is as complicated as the economy itself would be intractable and not useful. Recognizing

34

this, it is natural for agents to consider more than one possible model of the world. Of course, forecasters are not endowed with knowledge of the parameters of their model either. They estimate them over time. When we estimate a forecasting model with these three ingredients: state uncertainty, model uncertainty and parameter uncertainty, the resulting uncertainty shocks resemble some of the uncertainty proxies, but are not identical to them.

A next step for this agenda is to explore firm-level earnings forecasts and firm-level shocks and determine whether our model can also explain uncertainty facts at the micro level.

References

Bachmann, R., and C. Bayer (2012a): "Investment Dispersion and the Business Cycle," University of Aachen working paper.

——— (2012b): "Wait-and-See Business Cycles?," University of Aachen working paper.

Bachmann, R., S. Elstner, and E. Sims (2012): "Uncertainty and Economic Activity: Evidence from Business Survey Data," University of Aachen working paper.

Baker, S., N. Bloom, and S. Davis (2012): "Measuring Economic Policy Uncertainty," Stanford University working paper.

Bansal, R., and I. Shaliastovich (2010): "Confidence Risk and Asset Prices," *American Economic Review*, 100(2), 537–41.

Basu, S., and B. Bundick (2012): "Uncertainty Shocks in a Model of Effective Demand," Boston College working paper.

Bianchi, F., C. Ilut, and M. Schneider (2012): "Ambiguous business cycles and uncertainty shocks," Stanford University working paper.

Bloom, N. (2009): "The Impact of Uncertainty Shocks," *Econometrica*, 77(3), 623–685.

Bloom, N., M. Floetotto, N. Jaimovich, I. Sapora-Eksten, and S. Terry (2012): "Really Uncertain Business Cycles," NBER working paper 13385.

Born, B., A. Peter, and J. Pfeifer (2011): "Fiscal News and Macroeconomic Volatility," Bonn econ discussion papers, University of Bonn, Germany.

Bruno, V., and H. Shin (2012): "Capital Flows, Cross-Border Banking and Global Liquidity," Princeton University working paper.

Christiano, L., R. Motto, and M. Rostagno (2012): "Risk Shocks," Northwestern University working paper.

Cogley, T., and T. Sargent (2005): "The Conquest of US Inflation: Learning and Robustness to Model Uncertainty," *Review of Economic Dynamics*, 8, 528–563.

Collin-Dufresne, P., M. Johannes, and L. Lochstoer (2013): "Parameter Learning in General Equilibrium: The Asset Pricing Implications," Columbia University working paper.

Diether, K., C. Malloy, and A. Scherbina (2002): "Differences of Opinion and the Cross-Section of Stock Returns," *Journal of Finance*, 57, 2113–2141.

ELLIOTT, G., AND A. TIMMERMANN (2008): "Economic Forecasting," *Journal of Economic Literature*, 46(1), 3–56.

FERNANDEZ-VILLAVERDE, J., P. GUERRON-QUINTANA, J. F. RUBIO-RAMIREZ, AND M. URIBE (2011): "Risk Matters: The Real Effects of Volatility Shocks," *American Economic Review*, 101(6), 2530–61.

FERNANDEZ-VILLAVERDE, J., K. KUESTER, P. GUERRON, AND J. RUBIO-RAMIREZ (2012): "Fiscal Volatility Shocks and Economic Activity," University of Pennsylvania working paper.

HANSEN, L. (2007): "Beliefs, Doubts and Learning: Valuing Macroeconomic Risk," *American Economic Review*, 97(2), 1–30.

JOHANNES, M., L. LOCHSTOER, AND Y. MOU (2011): "Learning About Consumption Dynamics," Columbia University working paper.

JOHNSON, T. (2004): "Forecast Dispersion and the Cross Section of Expected Returns," *Journal of Finance*, 59, 1957–1978.

JURADO, K., S. LUDVIGSON, AND S. NG (2013): "Measuring Uncertainty," New York University working paper.

NAKAMURA, E., D. SERGEYEV, AND J. STEINSSON (2012): "Growth-Rate and Uncertainty Shocks in Consumption: Cross-Country Evidence," Columbia University working paper.

NIMARK, K. (2012): "Man Bites Dog Business Cycles," Working paper.

PASTOR, L., AND P. VERONESI (2012): "Uncertainty about Government Policy and Stock Prices," *Journal of Finance*, forthcoming.

ROBERTS, G., A. GELMAN, AND W. GILKS (1997): "Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorithms," *Annals of Applied Probability*, 7(1), 110–120.

STOCK, J., AND M. WATSON (2012): "Disentangling the Channels of the 2007-2009 Recession," NBER working paper 18094.

A   Estimating the model

For each of the models with parameter uncertainty (models 2-4), the updating process starts with prior beliefs about each of the model parameters. The complete set of prior means and variances for the model with normally-distributed innovations is in table 9.

	2-state model			2-state model, fat tails	
Par	Mean	Std	Par	Mean	Std
μ_1	5.9	0.7	μ_1	-3.02%	1%
μ_2	-0.3	1	μ_2	3.22%	0.8%
σ^2	$(3.8)^2$	2.5	σ^2	$(2.17\%)^2$	2%
q_{11}	0.87	0.032	q_{11}	0.67	0.16
q_{22}	0.7	0.16	q_{22}	0.95	0.05

Table 9: Prior assumptions for normal and t distributed models for GDP growth. These priors are the maximum likelihood estimates from estimating () or () on data from 1947-1968:Q3.

	2-state			2-state, t, $\nu = 5$	
Parameter	Mean	Stdev	Parameter	Mean	Stdev
μ_1	-1.55	1	μ_1	0.4	0.23
μ_2	1.92	0.4	μ_2	4.8	0.7
σ^2	$(3.56)^2$	1.5	σ^2	$(2.17)^2$	1.6
q_{11}	0.62	0.16	q_{11}	0.86	0.16
q_{22}	0.93	0.06	q_{22}	0.5	0.1

Figure 4: Prior assumptions for normal and t distributed models for TFP growth.

A.1 Computational algorithm

In what follows we show how to use Metropolis-Hastings algorithm to generate samples from $p\left(\theta|y^t\right)$ for each $t = 0, 1, 2, .., T$. [10]

The general idea of MCMC methods is to design a Markov chain whose stationary distribution, π (with $\pi T = \pi$ where T is a transitional kernel), is the distribution p we are seeking to characterize. In particular, the Metropolis-Hastings sampling algorithm constructs an ergodic Markov chain that satisfies a detailed balance property with respect to p and, therefore, produces the respective approximate samples. The transition kernel of that chain, T, is constructed based on sampling from a proposal conditional distribution $q\left(\theta|\theta^{(d)}\right)$ where d denotes the number of the sampling step. Specifically, given the d-step

---

[10]We drop here the dependence on $\mathcal{M}$ hoping that no confusion arises; the algorithm is applied independently to generate the respective samples under each model.

in the random walk $\theta^{(d)}$ the next-step $\theta^{(d+1)}$ is generated as follows

$$\theta^{(d+1)} = \begin{cases} \theta' & \text{with probability } \alpha\left(\theta^{(d)}, \theta'\right) = \min\left(1, \frac{p\left(\theta'|y^t\right)}{p\left(\theta^{(d)}|y^t\right)} \frac{q\left(\theta^{(d)}|\theta'\right)}{q\left(\theta'|\theta^{(d)}\right)}\right) \\ \theta^{(d)} & \text{with probability } 1 - \alpha\left(\theta^{(d)}, \theta'\right) \end{cases}$$

where $\theta' \sim q\left(\theta|\theta^{(d)}\right)$.

In our application, the simulation of the parameters is done through simple random walk proposals. In particular, for the means the proposed move is

$$\mu'_S = \mu_S + \tau_u \varepsilon_S \tag{18}$$

where $S \in \{1, ..., N\}$, $\varepsilon_S \sim \mathcal{N}(0, 1)$.

For the variance σ^2, the proposed move is a multiplicative random walk

$$\log \sigma' = \log \sigma + \tau_\sigma \xi_\sigma \tag{19}$$

where $\xi_\sigma \sim \mathcal{N}(0, 1)$.

In the case of the transition probability matrix, the move is slightly more involved due to the constraint on the sum of rows. We reparameterize each row $(q_{i1}, ..., q_{iN})$ as

$$q_{ij} = \frac{\omega_{ij}}{\sum_j \omega_{ij}}, \quad \omega_{ij} > 0, \quad j \in \{1, ..., N\}$$

so that the summation constraint does not hinder the random walk. The proposed move on ω_{ij} is then given by

$$\log \omega'_{ij} = \log \omega_{ij} + \tau_\omega \xi_\omega \tag{20}$$

where $\xi_\omega \sim \mathcal{N}(0, 1)$. Note that this reparametrization requires that we select a prior distribution on ω_{ij} rather than on q_{ij}.

The parameters τ_u, τ_σ, and τ_ω can be adjusted to optimize the performance of the sampler. Choosing a proposal with small variance would result in relatively high acceptance rates but with strongly correlated consecutive samples. See Roberts, Gelman, and Gilks (1997) for the results on optimal scaling of the random walk Metropolis algorithm.

Since the proposal is symmetric in all the cases (18), (19), and (20) and since ε_S, ξ_σ, and ξ_ω are drawn independently from one another, we have $q\left(\theta|\theta'\right) = q\left(\theta'|\theta\right)$, and the acceptance probability simplifies to $\min\left(1, \frac{p\left(\theta'|y^t\right)}{p\left(\theta^{(d)}|y^t\right)}\right)$. To compute that acceptance ratio, note that the posterior distribution $p\left(\theta|y^t\right)$ is given by

$$p\left(\theta|y^t\right) = \frac{p\left(y^t|\theta\right) p\left(\theta\right)}{p\left(y^t\right)}$$

where $p\left(y^t\right) = \int p\left(y^t|\theta\right) p\left(\theta\right) d\theta$ is the marginal likelihood (or data density).

The means, medians and confidence intervals of the resulting parameter estimates are illustrated in figure 5 for the model with normally-distributed shocks and in figure 6 for the model with t-distributed shocks.

A.2 Estimating the model with heterogeneous signals

With heterogenous signals, the model is the following

$$
\begin{aligned}
y_{t+1} &= \mu_{S_{t+1}} + \sigma e_{t+1} \\
z_{it} &= y_{t+1} + \sigma_\epsilon \epsilon_{it} + \sigma_\eta \eta_t
\end{aligned}
$$

where we assume that η_t, e_t, $\epsilon_{it} \sim N(0,1)$ and that η_t, and e_t are independent from one another and from ϵ_{it} for all t and all i. The forecaster wishes to compute his forecast

$$
E_t\left(y_{t+1}|y^t, z_i^t\right) = \int y_{t+1} p\left(y_{t+1}|y^t, z_i^t\right) dy_{t+1}
$$

where the predictive density is given by

$$
\begin{aligned}
p\left(y_{t+1}|y^t, z_i^t\right) &= \sum_{S_{t+1}} \int p\left(y_{t+1}|\theta, S_{t+1}\right) p\left(\theta, S_{t+1}|y^t, z_i^t\right) d\theta \\
&= \sum_{S_{t+1}} \int p\left(y_{t+1}|\theta, S_{t+1}\right) \Pr\left(S_{t+1}|\theta, y^t, z_i^t\right) p\left(\theta|y^t, z_i^t\right) d\theta
\end{aligned}
$$

In other words, our forecaster is still faced with the joint estimation problem of parameters and states, but now has additional information to condition on. We can still use the MCMC methods to approximate the posterior parameter distribution, $p\left(\theta|y^t, z_i^t\right)$. To that end, we need to show how to determine the data likelihood for a given parameter vector, $p\left(y^t, z_i^t|\theta\right)$, and how the agent updates state beliefs conditional on the observations of GDP growth rates, as well as the signal.

For a given θ, the likelihood $p\left(y^T, z_i^T\right)$ can be determined as follows

$$
\log p\left(y^T, z_i^T\right) = \sum_{t=1}^{T} \log\left[p\left(y_t, z_{it}|y^{t-1}, z_i^{t-1}\right)\right])
$$

In turn,

$$
p\left(y_t, z_{it}|y^{t-1}, z_i^{t-1}\right) = \sum_{k=1}^{2}\sum_{j=1}^{2} p\left(y_t, z_{it}|y_{t-1}, z_{it-1}, S_t = k, S_{t+1} = j\right) \Pr\left(S_t = k|y^{t-1}, z_i^{t-1}\right) \Pr\left(S_{t+1} = j|S_t = k\right)
$$

where $p\left(y_t, z_{it}|y_{t-1}, z_{it-1}, S_t = k, S_{t+1} = j\right)$ is a density of the following multivariate normal distribution

$$
MVN\left(\begin{bmatrix} \frac{\mu_{S_t=k}}{\sigma^2} + \frac{z_{it}}{\sigma_\eta^2 + \sigma_\epsilon^2} \\ \frac{1}{\sigma^2} + \frac{1}{\sigma_\eta^2 + \sigma_\epsilon^2} \\ \mu_{S_{t+1}=j} \end{bmatrix}, \begin{bmatrix} \left(\frac{1}{\sigma^2} + \frac{1}{\sigma_\eta^2 + \sigma_\epsilon^2}\right)^{-1} & 0 \\ 0 & \sigma^2 + \sigma_\eta^2 + \sigma_\epsilon^2 \end{bmatrix}\right)
$$

40

To show how the agent updates his beliefs recursively, assume for a moment that we know $\psi_{k,t} \equiv \Pr\left(S_t = k | g^t, z_i^t\right)$ with $\sum_{k=1}^{2} \psi_{k,t} = 1$. The posterior state belief $\psi_{k,t+1}$ can be updated recursively in the following way

$$\psi_{k,t+1} = \frac{\sum_{m=1}^{2}\sum_{j=1}^{2} p\left(y_{t+1}, z_{it+1} | S_{t+1} = k, S_t = j, S_{t+2} = m, y^t, z_i^t\right) q_{km} q_{jk} \psi_{j,t}}{p\left(y_{t+1}, z_{it+1} | y^t, z_i^t\right)}$$

A.3 Priors based on 1947-68 data

The priors from the results reported so far come from ML estimates on the full sample of data. Another way to estimate priors is to use only data prior to the start of our exercise and estimate ML parameters from that data.

model		state unc	param unc	signals	model unc
		(1)	(2)	(3)	(4)
Mean	U_t	4.73%	3.39%	2.39%	%
Std deviation	U_t	0.17%	0.48%	0.22%	%
Coef of variation	U_t	0.04	0.14	0.09	
Autocorrelation	U_t	0.62	0.86	0.86	
Corr(U_t, GDP_t)		-0.26	0.214	0.11	-

Table 10: **Properties model uncertainty series.** Column 1 labeled 'state unc' uses equation (4) to forecast y_{t+1}. Columns 2, 3 and 4 use equations (7), (10) and (**??**) to forecast y_{t+1}.

A.4 Estimating a continuous state model

The forecasting models we examined in the main text involved two or three discrete state. We could instead, consider a continuous state model, where the state is persistent. When we do, we find very similar results. Consider the following forecasting model:

$$y_t = s_t + \sigma \varepsilon_{y,t}$$
$$s_t = \rho s_{t-1} + \sigma_s \varepsilon_{s,t}$$

where $\varepsilon_{y,t}$ and $\varepsilon_{s,t}$ are standard normal random variables independent of each other. The vector of parameters that we need to estimate in the real time is now given by $\theta \equiv [\rho, \sigma, \sigma_s]'$. We begin with the following priors for these parameters

We will proceed as before and estimate this parameters vector using random-walk Metropolis Hastings algorithm. To that end, recall that the likelihood function of a given

Prior assumptions

parameter	mean	stdev
ρ	0.5	0.08
σ^2	3	0.5
σ_s^2	5	1

vector of parameters, θ, (which we need for the accept/reject step of the algorithm) is given by

$$p\left(y^T|\theta\right) = \prod_{t=0}^{T-1} p\left(y_{t+1}|y^t,\theta\right)$$

In turn, the predictive distribution of the data, $p\left(y_{t+1}|y^t,\theta\right)$ can be obtained as an integral against the filtering distribution

$$p\left(y_{t+1}|y^t,\theta\right) = \int\int p\left(y_{t+1}|s_{t+1},\theta\right) p\left(s_{t+1}|s_t,\theta\right) p\left(s_t|y^t,\theta\right) ds_t ds_{t+1}$$

We apply Kalman filtering techniques to obtain the distribution of the filtered values of the hidden state, $p\left(s_t|y^t,\theta\right)$, which will be characterized by the following moments $\widehat{s}_t \equiv E\left[s_t|y^{t-1},\theta\right]$ and $Var_{s,t} \equiv E\left[(s_t - \widehat{s}_t)(s_t - \widehat{s}_t)'\right]$. The predictive distribution of the data is then given by the following moments $E\left[y_t|y^{t-1},\theta\right] = \widehat{s}_t$ and $Var\left[y_t|y^{t-1},\theta\right] = Var_{s,t} + \sigma^2$.

Results: Moments

moment	$E\left[y_t|y^{t-1}\right]$	U_t	V_t	FE_{t+1}
mean	1.82	3.61	3.91	2.56
stdev	1.83	0.23	0.48	2.51
coeff of variation	1.01	0.06	0.12	0.98
autocorrelation	0.64	0.96	0.95	0.24
corr w/GDP	0.22	0.13	0.14	0.18

B   Estimated ARCH/GARCH process for each macro variable

In this section we present the GARCH models that we estimate to infer volatility proxies for TFP, GDP, government spending and the three month treasury rate. We estimate volatility using an ARMA model with ARCH or GARCH errors. The estimation procedure

is maximum likelihood. We considered several models and chose the AR and MA orders based on the significance of additional variables and their effect on the log-likelihood. Similarly, we considered different lags of linear terms for ϵ_t and variances σ_t^2 in the GARCH specification and used the significance and effect of additional variables on the log-likelihood to inform the specification choice.

GDP We use quarterly GDP growth rate data for 1947:Q2–2012:Q2.[11] The subsample of the GDP data which matches the uncertainty data is 1968:Q4 to 2011:Q4. The results of ADF tests indicate that the series for the full sample and matching subsample are stationary.

Best-fitting homoskedastic model Recall that we defined $y_{t+1} \equiv \ln(gdp_t) - \ln(gdp_{t-1})$. The best-fitting processes is an ARMA(1,0):

$$\Delta gdp_{t+1} = 3.25 + 0.37\Delta gdp_t + \epsilon_{t+1}, \quad \varepsilon_{t+1} \sim \mathcal{N}(0, 14.43)$$

The log-likelihood is -713.27.

Best-fitting heteroskedastic model For heteroskedastic processes, the best specification is ARMA(1,0) for growth and GARCH(1) for variance:

$$\Delta gdp_{t+1} = 3.38 + 0.41\Delta gdp_t + \epsilon_{t+1}, \quad \varepsilon_{t+1} \sim \mathcal{N}(0, \sigma_{t+1}^2)$$

$$\sigma_{t+1}^2 = 0.52 + 0.76\sigma_t^2 + 0.24\epsilon_t^2$$

The log-likelihood is -693.90.

Government spending Our data on federal spending and state and local government spending is for 1968:Q4 to 2012:Q2.[12] So that we have a stationary series we analyze the annualized quarterly growth rate, approximated using the same method as for TFP. The results of ADF tests (Table **??**) are consistent with our growth rate series being stationary.

Best-fitting homoskedastic model The best-fitting model for federal spending is an ARMA(3,3):

$$\Delta g_{t+1} = 1.759 + 1.450\Delta g_t - 1.283\Delta g_{t-1} + 0.726\Delta g_{t-2} + \varepsilon_{t+1} - 1.578\varepsilon_t + 1.575\varepsilon_{t-1} - 0.761\varepsilon_{t-2},$$

$$\varepsilon_{t+1} \sim \mathcal{N}(0, 41.940).$$

---

[11]Data source: Bureau of Economic Analysis (http://www.bea.gov/national/index.htm#gdp). We are using the seasonally adjusted annual rate for the quarterly percentage change in real GDP. Note that this data is for the actual percentage change, not an approximation. The version of the data is 27 July 2012, downloaded 2 August 2012.

[12]Source: National Income and Product Accounts, Table 3.9.3, Real Government Consumption Expenditures and Gross Investment, Quantity Indexes. Available at http://www.bea.gov//national/nipaweb/DownSS2.asp, downloaded 8 August 2012. The data is seasonally adjusted.

For state and local government spending we use an ARMA(1,1):

$$\Delta g_{t+1} = 1.816 + 0.850 \Delta g_t + \varepsilon_{t+1} - 0.485 \varepsilon_t, \quad \varepsilon_{t+1} \sim \mathcal{N}(0, 5.806).$$

All parameters are significant at 1% except the constant for federal spending (not significant) and the constant for state and local spending (significant at 5%). The log-likelihoods are -405.42 and -282.95 respectively. The p-values for the ARCH-LM tests are 0.672 and 0.000 respectively. Therefore homoskedasticity can be rejected for state and local government spending, but cannot be rejected for federal spending.

Best-fitting heteroskedastic model The best-fitting model for federal spending is an ARMA(3,3) for growth and ARCH(1) for variance:

$$\Delta g_{t+1} = 7.031 + 1.513 \Delta g_t - 1.235 \Delta g_{t-1} + 0.715 \Delta g_{t-2} + \varepsilon_{t+1} - 1.746 \varepsilon_t + 1.888 \varepsilon_{t-1} - 0.963 \varepsilon_{t-2},$$

$$\varepsilon_{t+1} \sim \mathcal{N}(0, \sigma^2_{t+1}),$$
$$\sigma^2_{t+1} = 34.335 + 0.012 \varepsilon_t^2.$$

All parameters are significant at 1% except for the ARCH parameter (insignificant). The log-likelihood is -392.77. For state and local government spending, we use an ARMA (1,1) for growth and ARCH(1) for variance (an AR(2) with an ARCH(1) process fits equally well):

$$\Delta g_{t+1} = 1.487 + 0.863 \Delta g_t + \varepsilon_{t+1} - 0.450 \varepsilon_t, \quad \varepsilon_{t+1} \sim \mathcal{N}(0, \sigma^2_{t+1}),$$

$$\sigma^2_{t+1} = 3.483 + 0.419 \varepsilon_t^2.$$

All parameters are significant at 1% except for the constant for the growth process (significant at 10%) and the ARCH parameter (significant at 5%). The log-likelihood is -276.49.

Interest rates The interest rate that we use is the three month treasury bill rate.[13] Ideally we would like to use the policy rate (the federal funds rate), but no forecast data is available for this. We have therefore chosen the interest rate for which forecast data is available whose time horizon is closest to the federal funds rate. We have data on this rate for 1954:Q1–2012:Q2. But we use only the data starting in 1968q4 so that the volaility sample and the forecast error sample are the same.

So that our data is stationary we first difference it. The results of ADF tests (Table **??**) sample support this: at the 5% significance level we cannot reject the null hypothesis that the interest rate level data has a unit root and we can reject this null hypothesis for the first-difference of the data. However the outcome of the test for the levels data is sensitive to the order of the AR model that we fit to the data. If we assume that the levels data

---

[13]Data source: Board of Governors (http://research.stlouisfed.org/fred2/series/DTB3/). We use the 3-Month Treasury Bill: Secondary Market Rate. The data has a daily frequency and we use the average for each quarter.

is modeled by an AR(1) process instead of an AR(2) process then the null hypothesis is rejected at the 5% level.

Best-fitting homoskedastic model The best-fitting process is an ARMA(1,0) with no constant:

$$\Delta r_t = 0.313\Delta r_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, 0.3606).$$

The AR(1) coefficient is significant at the 5% level, the variance estimate is significant at the 5% level. The log-likelihood is -111.86. The p-value for the ARCH-LM test is 0.045 so homoskedasticity can be rejected at 5%.

Best-fitting heteroskedastic model The best-fitting process is an ARMA(1,0) with no constant and a GARCH(1,1) for variance:

$$\Delta r_t = 0.663\Delta r_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \sigma_t^2),$$

$$\sigma_t^2 = 0.016 + 0.737\sigma_{t-1}^2 + 0.162\varepsilon_{t-1}^2$$

All coefficients are significant at the 1% level. The log-likelihood is -81.22.

For both the homoskedastic and hetereskedastic processes for the matching subsample we estimated models with higher order AR and MA components and found that the extra coefficients were not significant.

Checking stationarity For each sample period of each variable we use Augmented Dickey-Fuller (ADF) tests to check that the series is stationary. This test is based on estimating an AR model for the data. For each sample we use the best-fitting AR model as the basis for the test.[14]

As discussed in Section 2, we estimate homoskedastic and heteroskedastic models for each series. For the homoskedastic models we use robust standard errors. We use the errors from the homoskedastic models to test the null hypothesis of no heteroskedasticity using an ARCH-LM test. This test requires fitting an AR model to the squared errors of the homoskedastic model. For each sample we set the AR order equal to the order of the AR component of the ARCH/GARCH model that we use.

---

[14]Note that the number of lags used for the test is one less than the order of the AR process that we model the data with. If we are using an AR(p) process to model the variable x_t then the regression underlying the ADF test is

$$\Delta x_t = \phi x_{t-1} + \sum_{j=1}^{p-1} \alpha_j \Delta x_{t-j} + u_t, \tag{21}$$

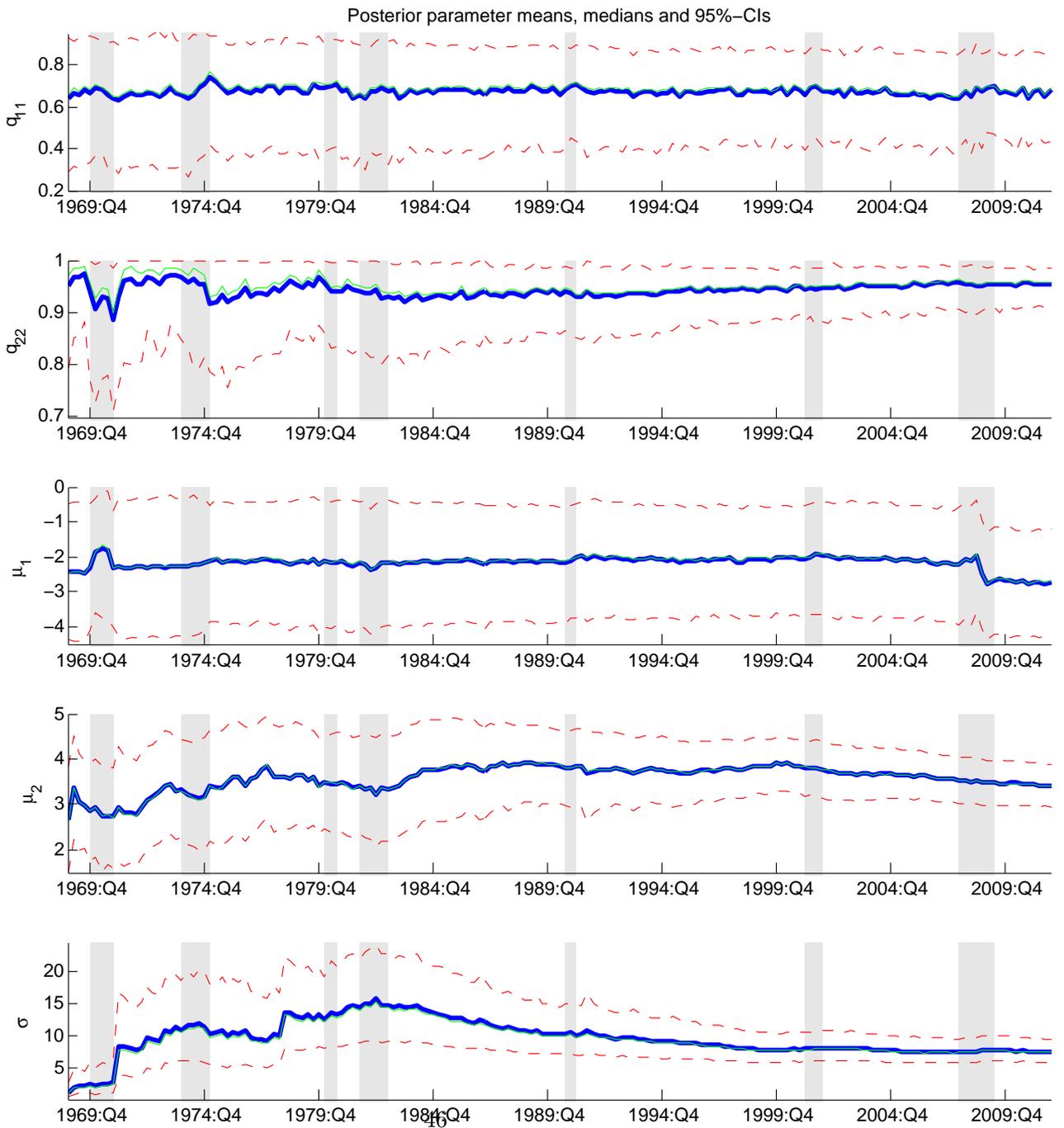where u_t is the error term for an OLS regression.

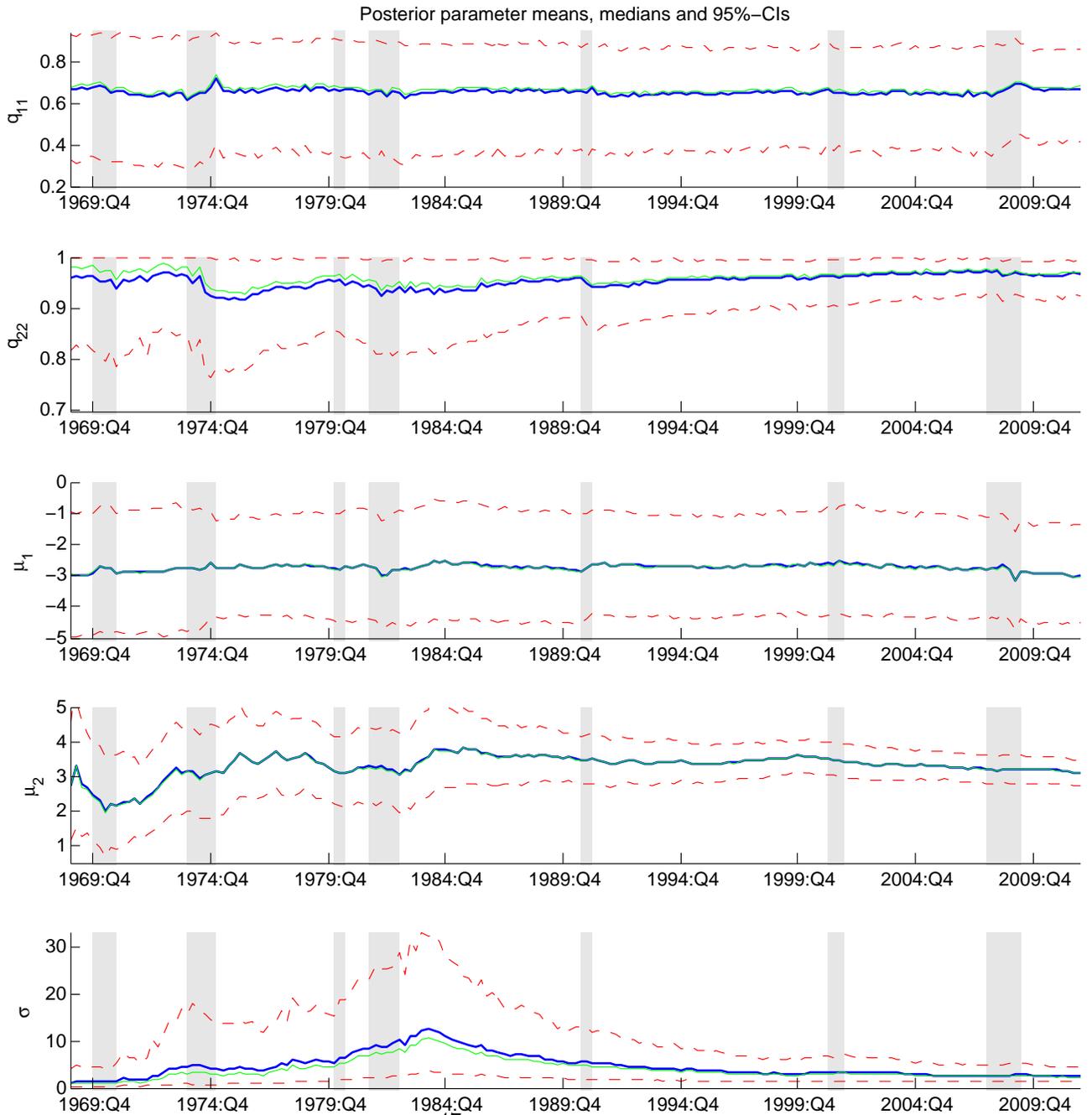Figure 5: Parameter estimates in model with normal shocks.

Figure 6: Parameter estimates in model with *t*-distributed shocks.