The Zero Lower Bound and Endogenous Uncertainty: Online Appendix*

Michael Plante Alexander W. Richter Nathaniel A. Throckmorton

September 10, 2016

ABSTRACT

This appendix provides additional evidence that a stronger correlation between real GDP growth and macroeconomic uncertainty emerged since late 2008. It also contains technical details about how we estimate our structural model and information related to its empirical fit.

*Plante, Federal Reserve Bank of Dallas, 2200 N Pearl St, Dallas, TX 75201 (michael.plante@dal.frb.org); Richter, Federal Reserve Bank of Dallas and Auburn University, 2200 N Pearl St, Dallas, TX 75201 (alex.richter@dal.frb.org); Throckmorton, College of William & Mary, P.O. Box 8795, Williamsburg, VA 23187 (nathrockmorton@wm.edu).

1 RELATIONSHIP BETWEEN UNCERTAINTY AND OUTPUT OUTSIDE THE U.S.

Although our focus is on the U.S., we estimate our VAR using data from other countries that faced constraints on monetary policy to see whether there were similar changes in the correlation. The Euro-zone deposit rate and the bank rate in the U.K. were reduced to 0.25% and 0.5%, respectively, in 2009Q1. Unfortunately, we cannot estimate the model for the Euro area since data is only available since 1995Q1. For the U.K., we use analogous observables to the U.S. data and the same sample, except the pre-ZLB and ZLB samples are split in 2009Q1. The correlation in the pre-ZLB sample exceeds zero in 21% of draws and the median correlation is -0.13. In the ZLB sample, however, the correlation is positive in only 0.3% of draws and the median is -0.62. The difference between the ZLB and pre-ZLB correlations is positive in only 1.5% of draws and the median difference is -0.46, which indicates that the constraint had a similar effect in the U.K.

The central bank has been constrained in Japan for a much longer period than in the U.S and Euro area. In April 1995, the Bank of Japan lowered its discount rate to 1%, and the T-bill rate hit 0.37% in 1995Q3. When we split the sample in 1995Q3, the median correlation is 0.25 in the pre-ZLB sample (1986Q1-1995Q2) and -0.25 in the ZLB sample (1995Q3-2014Q2). The median difference between the correlations is -0.50, and the correlation is positive in only 0.6% of draws, similar to our results with U.S and U.K. data. We also examine the ZLB period before the Great Recession (1995Q3-2007Q4) and since Great Recession (2008Q1-2014Q2). The correlations in those samples are -0.21 and -0.29, respectively. In both cases, the differences from the pre-ZLB sample are positive in less than 2% of draws. Those results provide further evidence that the ZLB is the main source of the stronger correlation, rather than a unique feature of the Great Recession.

We also examine the correlations using survey data from the Euro area and the U.K. The European Central Bank (ECB) has conducted its own SPF since 1999Q1. It asks for forecasts of Euro area real GDP growth. For example, the survey conducted in 1999Q1 requests forecasts for 1999Q3, given the last GDP release is from 1998Q3. Similar to the U.S. SPF, we calculate the forecast dispersion, $\text{ECB SPF FD}_t = \hat{y}^{75}_{t+2|t-2} - \hat{y}^{25}_{t+2|t-2}$, where $\hat{y}^{x}_{t+2|t-2}$ is the xth percentile of the quarter t forecast of real GDP growth in quarter $t+2$, given data in quarter $t-2$ and earlier. The correlation between Euro area real GDP growth and the ECB SPF FD in the pre-ZLB sample (1999Q1-2008Q4) is -0.32, and the correlation in the ZLB sample (2009Q1-2014Q4) is -0.55.

The Bank of England conducts a survey called the Survey of External Forecasters (BOE SEF), which has asked for forecasts of real GDP growth since 1998Q1.[1] Before 2006Q2, the survey asked for projections in quarter 4 of the survey year, quarter 4 1 year ahead, and the same quarter 2 years ahead. For example, the forecast dates in the 2006Q1 survey were 2006Q4, 2007Q4, and 2008Q1. Since 2006Q2, the survey has asked for projections for the same quarter 1, 2, and 3 years ahead. Unfortunately, we cannot calculate correlations in the pre-ZLB sample because the forecast horizons change. In the ZLB sample, the correlation between real GDP growth and the dispersion in forecasts 1 year ahead is -0.63, which is similar to the correlations in the U.S. and Euro area.[2]

2 LONGER-TERM FORECASTS

In the paper, we exclusively focus on survey-based forecasts over a one-quarter horizon. The SPF also asks forecasters to predict various macro variables over longer horizons. We denote the

[1] See Boero et al. (2008) for more information about the BOE SEF and how it compares to other related surveys.

[2] There is also a survey of Japanese professional forecasters, but it began in mid-2004 and does not provide a large enough sample. See Komine et al. (2009) for details about the survey and analysis of the forecasters' performance.

	Real GDP Growth			IP Growth		
	SPF(2)	SPF(3)	SPF(4)	SPF(2)	SPF(3)	SPF(4)
Pre-ZLB Sample (1986Q1-2008Q3)	-0.09	-0.10	-0.02	-0.26***	-0.24**	-0.17*
ZLB Sample (2008Q4-2014Q2)	-0.64***	-0.47**	-0.46**	-0.76***	-0.57***	-0.32*
Difference (ZLB-Pre-ZLB)	-0.55***	-0.37**	-0.45**	-0.50***	-0.34**	-0.15

Table 1: Correlations between the SPF FD j quarters ahead and our two measures of economic activity before and during the ZLB period. An asterisk indicates a correlation in the pre-ZLB or ZLB sample is statistically less than 0 or the difference between the correlations in the two samples is significant at a ***1%, **5%, and *10% level.

inter-quartile range of the quarter t forecasts of real GDP growth in quarter $t+k$ as SPF FD(k). Table 1 plots the correlations between the SPF FD(k) and our two measures of economic activity for $k \in \{2,3,4\}$. All of our qualitative results are robust to longer forecasting horizons. The correlations in the pre-ZLB sample are weak when calculated with real GDP growth and slightly stronger with IP growth. In the ZLB sample, however, the correlations are far more negative, and the differences from the pre-ZLB sample are significant at a 5% level or higher in all but one case.[3]

3 ESTIMATION ALGORITHM

We use a random walk Metropolis-Hastings algorithm to estimate our model with quarterly data from 1986Q1 to 2014Q2. To measure how well the model fits the data, we use the adapted particle filter described in Algorithm 12 in Herbst and Schorfheide (2016), which modifies the filter in Stewart and McCarty (1992) and Gordon et al. (1993) to better account for outliers in the data.

3.1 METROPOLIS-HASTINGS ALGORITHM

The following steps outline the algorithm:

1. Specify the prior distributions, means, variances, and bounds of each element of the vector of N_e estimated parameters, $\theta \equiv \{\varphi, h, \rho_\beta, \rho_g, \rho_i, \sigma_\varepsilon, \sigma_\upsilon, \sigma_\nu, \phi_\pi, \phi_y, \bar{g}, \bar{\pi}\}$.
2. Our observables are the per capita real GDP growth rate, the inflation rate, and the federal fund rate. To match their model implied values, transform the data on per capita real GDP (RGDP/CNP), the GDP Deflator (DEF), and the federal funds rate (FFR) according to:

$$\hat{y}_t^{data} = \log(RGDP_t/CNP_t) - \log(RGDP_{t-1}/CNP_{t-1}),$$
$$\hat{\pi}_t^{data} = \log(DEF_t/DEF_{t-1}),$$
$$\hat{\imath}_t^{data} = \log((1 + FFR_t/100)^{1/4}),$$

   The matrix of observables is $\hat{\mathbf{x}}^{data} \equiv [\hat{y}_t^{data}, \hat{\pi}_t^{data}, \hat{\imath}_t^{data}]_{t=1986Q1}^{2014Q2}$, which has $T \times 3$ elements.
3. Find the posterior mode to initialise the preliminary Metropolis-Hastings step.
   (a) For all $i \in \{1, \ldots, N_m\}$, where $N_m = 10{,}000$, apply the following steps:
      i. Draw $\hat{\theta}_i$ from the joint prior distribution and calculate its density value:

$$\log \ell_i^{prior} = \sum_{j=1}^{N_e} \log p(\hat{\theta}_{i,j} | \mu_j, \sigma_j^2),$$

[3] Andrade et al. (2014) examine forecaster disagreement over longer horizons using Blue Chip Financial Forecasts.

where p is the prior density function of parameter j with mean μ_j and variance σ_j^2.

ii. Given $\hat{\theta}_i$, solve the model according to Appendix A. If the algorithm converges, then compute the stochastic steady state, otherwise repeat step 3(a)i and redraw $\hat{\theta}_i$.

iii. If the stochastic steady state exists, then use the particle filter in section 3.2 to obtain the log-likelihood value for the model, $\log \ell_i^{model}$, otherwise repeat step 3(a)i.

iv. The posterior log-likelihood is $\log \ell_i^{post} = \log \ell_i^{prior} + \log \ell_i^{model}$

(b) Calculate $\max(\log \ell_1^{post}, \ldots, \log \ell_{N_m}^{post})$ and find the corresponding parameter vector, $\hat{\theta}_0$.

4. Approximate the covariance matrix for the joint posterior distribution of the parameters, Σ, which is used to draw candidates during the preliminary Metropolis-Hastings step.

(a) Locate the draws with a likelihood in the $p = 0.9$ percentile. Stack the $N_{m,sub} = (1-p)N_m$ draws in a $N_{m,sub} \times N_e$ matrix, $\hat{\Theta}$, and define $\tilde{\Theta} = \hat{\Theta} - \sum_{i=1}^{N_{m,sub}} \hat{\theta}_{i,j}/N_{m,sub}$.

(b) Calculate $\Sigma = \tilde{\Theta}'\tilde{\Theta}/N_{m,sub}$ and verify it is positive definite, otherwise repeat step 3.

5. Perform an initial run of the random walk Metropolis-Hastings algorithm.

(a) For all $i \in \{0, \ldots, N_d\}$, where $N_d = 25{,}000$, perform the following steps:

i. Draw a candidate vector of parameters, $\hat{\theta}_i^{cand}$, where

$$\hat{\theta}_i{}^{cand} \sim \begin{cases} \mathbb{N}(\hat{\theta}_0, c_0\Sigma) & \text{for } i = 0, \\ \mathbb{N}(\hat{\theta}_{i-1}, c\Sigma) & \text{for } i > 0. \end{cases}$$

We set $c_0 = 0$ and tune c to target an overall acceptance rate of roughly 30%.

ii. Calculate the prior density value, $\log \ell_i^{prior}$, of the candidate draw, $\hat{\theta}_i^{cand}$ as in 3(a)i.

iii. Given $\hat{\theta}_i^{cand}$, solve the model according to Appendix A. If the algorithm converges, compute the stochastic steady state, otherwise repeat 5(a)i and draw a new $\hat{\theta}_i^{cand}$.

iv. If the stochastic steady state exists, then use the particle filter in section 3.2 to obtain the log-likelihood value for the model, $\log \ell_i^{model}$, otherwise repeat 5(a)i.

v. Accept or reject the candidate draw according to

$$(\hat{\theta}_i, \log \ell_i) = \begin{cases} (\hat{\theta}_i^{cand}, \log \ell_i^{cand}) & \text{if } i = 0, \\ (\hat{\theta}_i^{cand}, \log \ell_i^{cand}) & \text{if } \log \ell_i^{cand} - \log \ell_{i-1} > \hat{u}, \\ (\hat{\theta}_{i-1}, \log \ell_{i-1}) & \text{otherwise,} \end{cases}$$

where $\hat{u}$ is a draw from a uniform distribution, $\mathbb{U}[0,1]$, and the posterior log-likelihood associated with the candidate draw is $\log \ell_i^{cand} = \log \ell_i^{prior} + \log \ell_i^{model}$.

(b) Burn the first $N_b = 5000$ draws and use the remaining sample to calculate the mean draw, $\bar{\theta}^{preMH} = \sum_{i=N_b+1}^{N_{\text{preMH}}} \hat{\theta}_i$, and the covariance matrix, Σ^{preMH}. We follow step 4 to calculate Σ^{preMH} but use all $N_d - N_b$ draws instead of just the upper pth percentile.

6. Following the procedure in step 5, perform a final run of the Metropolis-Hastings algorithm, where $\theta_0 = \bar{\theta}^{preMH}$ and $\Sigma = \Sigma^{preMH}$. We set $N_d = 100{,}000$ and keep every 100th draw. The remaining $1{,}000$ draws form a representative sample from the joint posterior distribution.

3.2 ADAPTED PARTICLE FILTER

The following steps outline the filter:

1. Initialise the filter by drawing $\mathbf{e}_{t,p} = \{\varepsilon_{t,p}, \upsilon_{t,p}, \nu_{t,p}\}_{t=-24}^{0}$ for all $p \in \{0, \ldots, N_p\}$ and simulating the model, where N_p is the number of particles. The final state vector, $\mathbf{z}_{0,p}$, represents a draw from the ergodic distribution and is used to initialise the filter. We set $N_p = 40{,}000$.
2. For all $p \in \{1, \ldots, N_p\}$ apply the following steps:
   (a) Draw a vector of shocks from an adapted distribution, $\mathbf{e}_{t,p} \sim \mathbb{N}(\bar{\mathbf{e}}_t, I)$, where $\bar{\mathbf{e}}_t$ is chosen to maximise $p(\xi_t|\mathbf{z}_t)p(\mathbf{z}_t|\mathbf{z}_{t-1})$ and $\mathbf{z}_{t-1} = \sum_{p=1}^{N_p} \mathbf{z}_{t-1,p}/N_p$ is the state vector.
      i. Given $\mathbf{z}_{t-1}$ and a guess for $\bar{\mathbf{e}}_t$, obtain $\mathbf{z}_t$, and the endogenous variables, $\mathbf{w}_t$.
      ii. Transform the predictions for real GDP ($\tilde{y}^{gdp}$), inflation (π), and the policy rate (i) according to $\hat{\mathbf{x}}_t^{model} = \left[\log(g_t \tilde{y}_t^{gdp}/\tilde{y}_{t-1}^{gdp}), \log(\pi_t), \log(i_t)\right]$.
      iii. Calculate the difference between the model predictions and the data, $\xi_t = \hat{\mathbf{x}}_t^{model} - \hat{\mathbf{x}}_t^{data}$, which is assumed to be multivariate normally distributed with density:
      $$p(\xi_t|\mathbf{z}_t) = (2\pi)^{-3/2}|H|^{-1/2}\exp(-\xi_t' H^{-1}\xi_t/2),$$
      where $H \equiv \text{diag}(\sigma_{me,\hat{y}}^2, \sigma_{me,\pi}^2, \sigma_{me,i}^2)$ is the measurement error covariance matrix.
      iv. The probability of observing the current state, $\mathbf{z}_t$, given, $\mathbf{z}_{t-1}$, is given by
      $$p(\mathbf{z}_t|\mathbf{z}_{t-1}) = (2\pi)^{-3/2}\exp(-\bar{\mathbf{e}}_t'\bar{\mathbf{e}}_t/2).$$
      v. Maximise $p(\xi_t|\mathbf{z}_t)p(\mathbf{z}_t|\mathbf{z}_{t-1}) \propto \exp(-\xi_t' H^{-1}\xi_t/2)\exp(-\bar{\mathbf{e}}_t'\bar{\mathbf{e}}_t/2)$ by solving for the optimal $\bar{\mathbf{e}}_t$. We converted MATLAB's `fminsearch` routine to Fortran.
   (b) Obtain $\mathbf{z}_{t,p}$, and the vector of endogenous variables, $\mathbf{w}_{t,p}$, given $\mathbf{z}_{t-1,p}$ and $\mathbf{e}_{t,p}$.
   (c) Calculate, $\xi_{t,p} = \hat{\mathbf{x}}_{t,p}^{model} - \hat{\mathbf{x}}_t^{data}$. The unnormalised weight on particle p is given by
   $$\omega_{t,p} = \frac{p(\xi_t|\mathbf{z}_{t,p})p(\mathbf{z}_{t,p}|\mathbf{z}_{t-1,p})}{g(\mathbf{z}_{t,p}|\mathbf{z}_{t-1,p}, \hat{\mathbf{x}}_t^{data})} \propto \frac{\exp(-\xi_{t,p}' H^{-1}\xi_{t,p}/2)\exp(-\mathbf{e}_{t,p}'\mathbf{e}_{t,p}/2)}{\exp(-(\mathbf{e}_{t,p} - \bar{\mathbf{e}}_t)'(\mathbf{e}_{t,p} - \bar{\mathbf{e}}_t)/2)}.$$
   If there was no adaptation, then $\bar{\mathbf{e}}_t = 0$ and $\omega_{t,p} = p(\xi_t|\mathbf{z}_{t,p})$, as it is in a basic filter. The contribution to the model's likelihood in period t is then $\ell_t^{model} = \sum_{p=1}^{N_p} \omega_{t,p}/N_p$.
   (d) Normalise the weights, $W_{t,p} = \omega_{t,p}/\sum_{p=1}^{N_p}\omega_{t,p}$. Then use systematic resampling with replacement from the swarm of particles as described in Kitagawa (1996) to get a set of particles that represents the filter distribution and reshuffle $\{\mathbf{z}_{t,p}\}_{p=1}^{N_p}$ accordingly.
3. Apply step 2 for all $t \in \{1, \ldots, T\}$. The log-likelihood is then $\log \ell^{model} = \sum_{t=1}^{T} \log \ell_t^{model}$.

3.3 SV ESTIMATION PROCESS

To estimate the parameters of the stochastic volatility process we conduct a mode search over $\lambda = (\rho_\sigma, \sigma_x)$, conditional on the posterior mean, $\bar{\theta}$, of the baseline model. The priors for (ρ_σ, σ_x) are the same as those for $(\rho_g, \sigma_\varepsilon)$ (see table 4 in the main paper).

1. For all $i \in \{1, \ldots, N_m\}$, where $N_m = 500$, apply the following steps:
   (a) Draw $\hat{\lambda}_i$ from the joint prior distribution and calculate its density value:
   $$\log \ell_i^{prior} = \sum_{j=1}^{2} \log p(\hat{\lambda}_{i,j}|\mu_j, \sigma_j^2),$$
   where p is the prior density function of parameter j with mean μ_j and variance σ_j^2.

(b) Given $\hat{\lambda}_i$ and $\bar{\theta}$, solve the model according to Appendix A. If the algorithm converges, then compute the stochastic steady state, otherwise repeat step 3(a)i and redraw $\hat{\lambda}_i$.

(c) If the stochastic steady state exists, then use the particle filter in section 3.2 to obtain the log-likelihood value for the SV model, $\log \ell_i^{SV model}$, otherwise repeat step 3(a)i.

(d) The posterior log-likelihood is $\log \ell_i^{post} = \log \ell_i^{prior} + \log \ell_i^{SV model}$

2. Calculate $\max(\log \ell_1^{post}, \ldots, \log \ell_{N_m}^{post})$ and find the corresponding parameter vector, $\hat{\lambda}$.

We verified that the mode search produced the desired result by interpolating a surface based on the draws, as shown in figure 1. There is a clear global maximum, since the surface is concave.

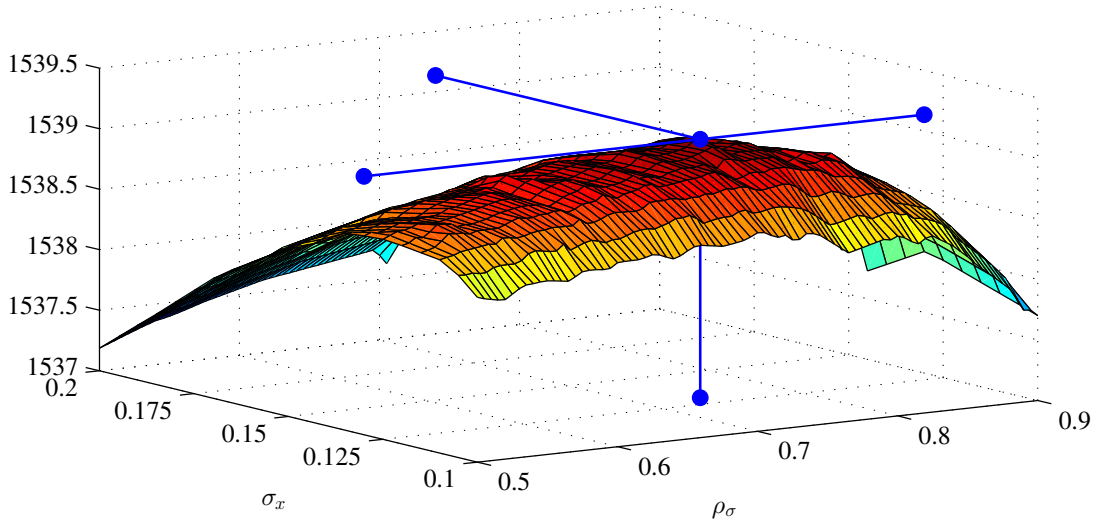

Figure 1: Posterior likelihoods from various (ρ_σ, σ_x) combinations. The solid lines indicate the global maximum.

We conducted a second mode search, conditional on our estimates of the SV parameters and all of the posterior mean parameters except $\bar{\sigma}_\varepsilon$. Given each draw from the prior distribution for $\bar{\sigma}_\varepsilon$, we solve the model and calculate the likelihood. After 100 draws, the likelihood was maximised with $\bar{\sigma}_\varepsilon = 0.00885$, which is only slightly less than our posterior estimate of 0.00968. Moreover, values below 0.007 led to significantly lower likelihoods. These results suggest that estimating the fully nonlinear model with SV would have relatively small impacts on the other posterior estimates.

4 ESTIMATION DIAGNOSTICS

We programmed the estimation procedure in Fortran and ran it on the Hopper computing cluster at Auburn University using Open MPI. We first solved and filtered the model $10{,}000$ times to initialise the Metropolis-Hastings (MH) algorithm at the posterior mode and obtain an estimate of the parameter covariance matrix. We then obtained $25{,}000$ draws from the posterior distribution. Using those draws, we recalculated the covariance matrix and used the posterior mean to initialise a second stage of the MH algorithm. We obtained $100{,}000$ draws and kept every 100th draw without a burn period. The remaining $1{,}000$ draws, which are shown in figure 2, form our posterior distribution. To show that our chains converged to their ergodic distributions, we used the Geweke χ^2 convergence test following Geweke (1992, 1999). The default test is to burn the first half of the chain and then compare the first 20% to the last 50% of the draws. As long as we thin the chains by

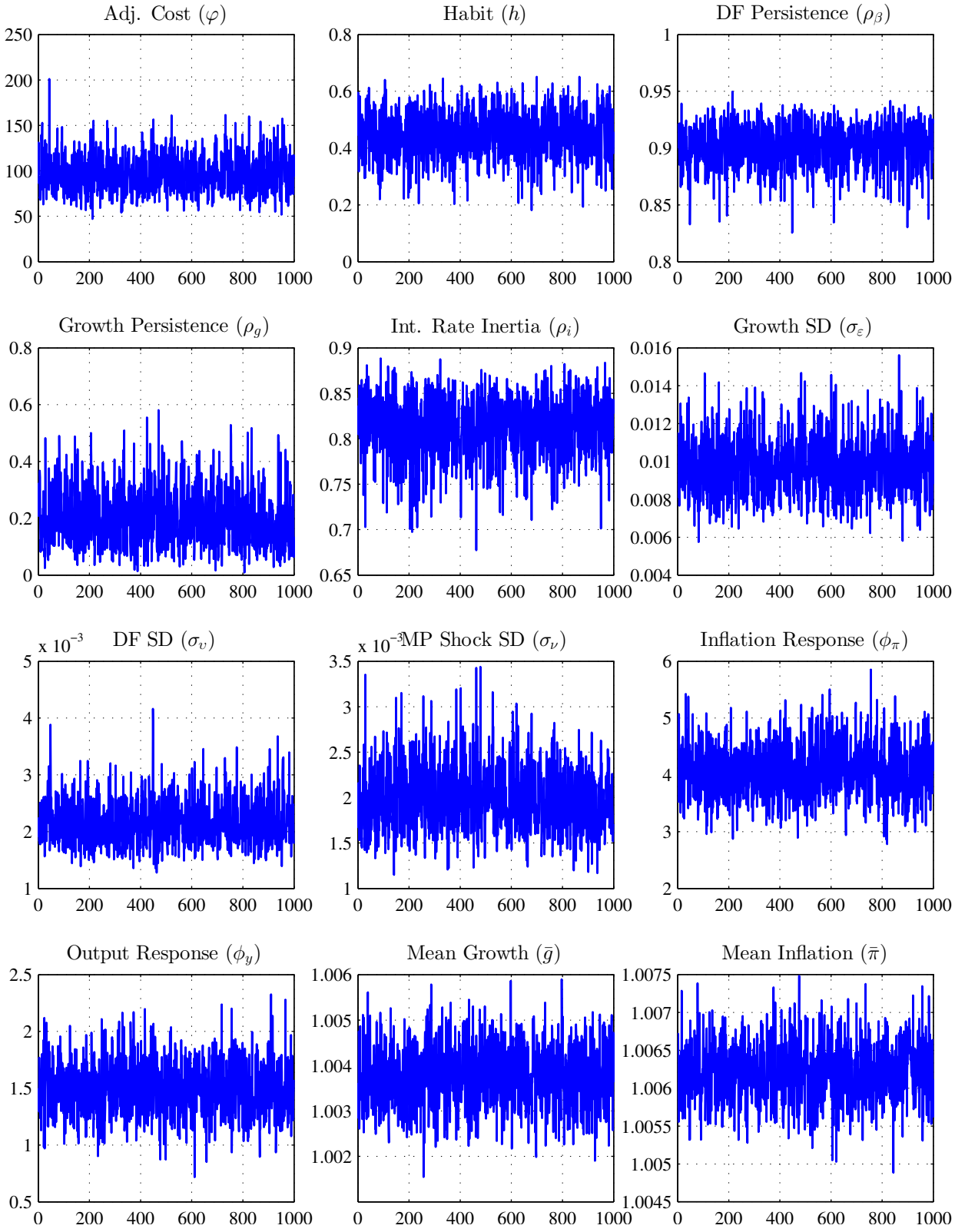

Figure 2: Trace plots. We drew 100,000 draws from each posterior distribution and kept every 100th draw.

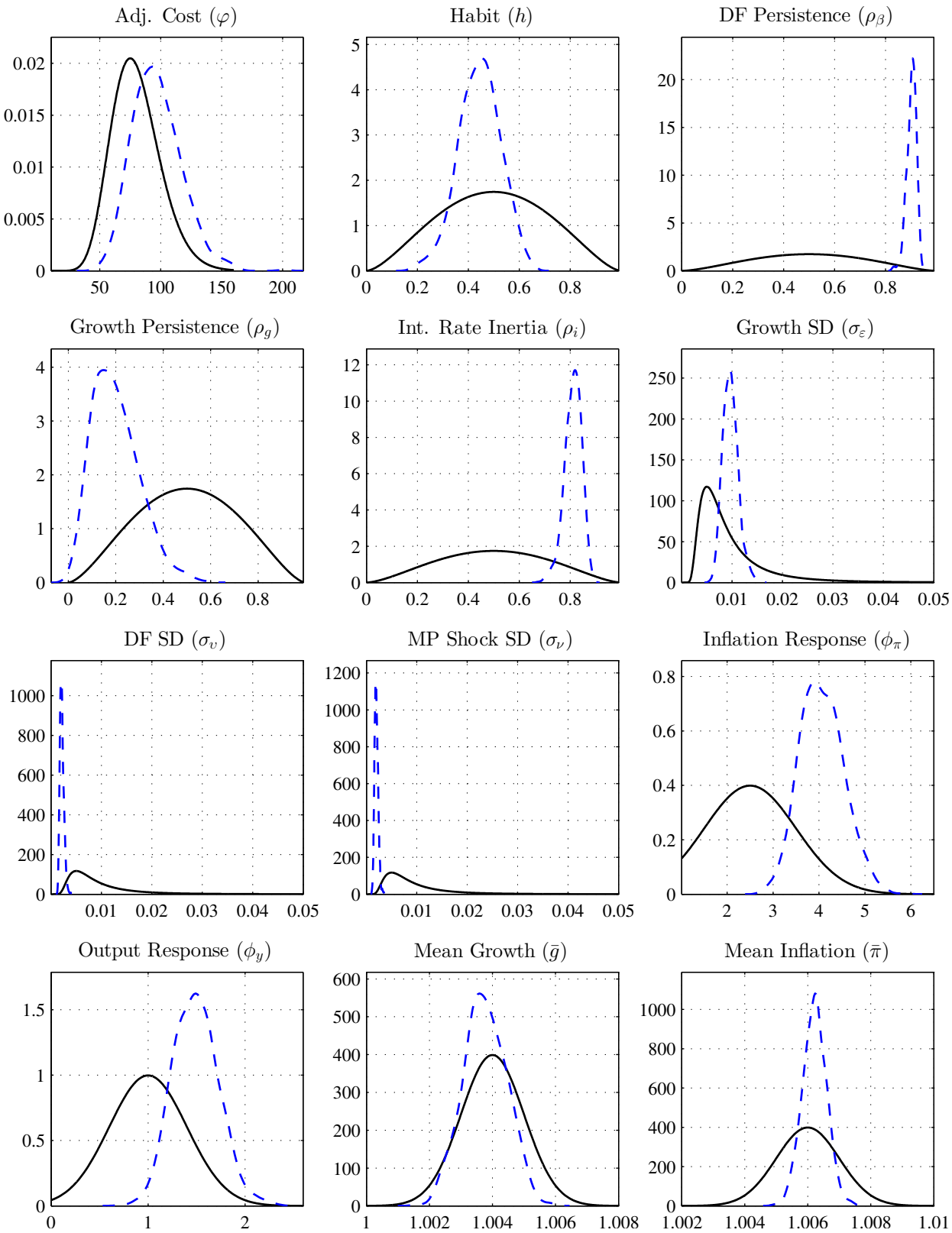

Figure 3: Prior distributions (solid lines) vs. kernel density of the posterior draws (dashed lines).

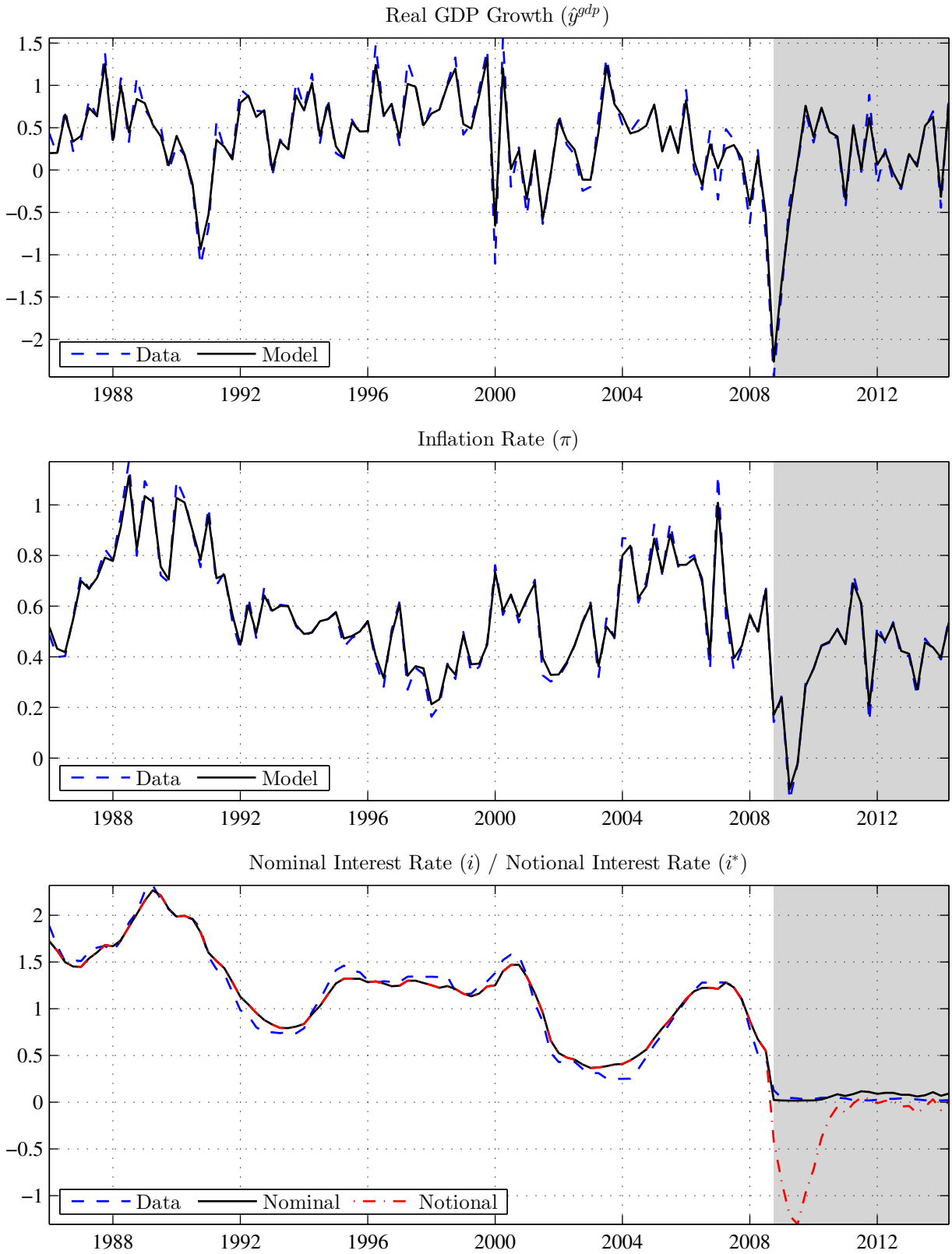

Figure 4: Time paths of the data (dashed line) and the median filtered series from the model (solid line).

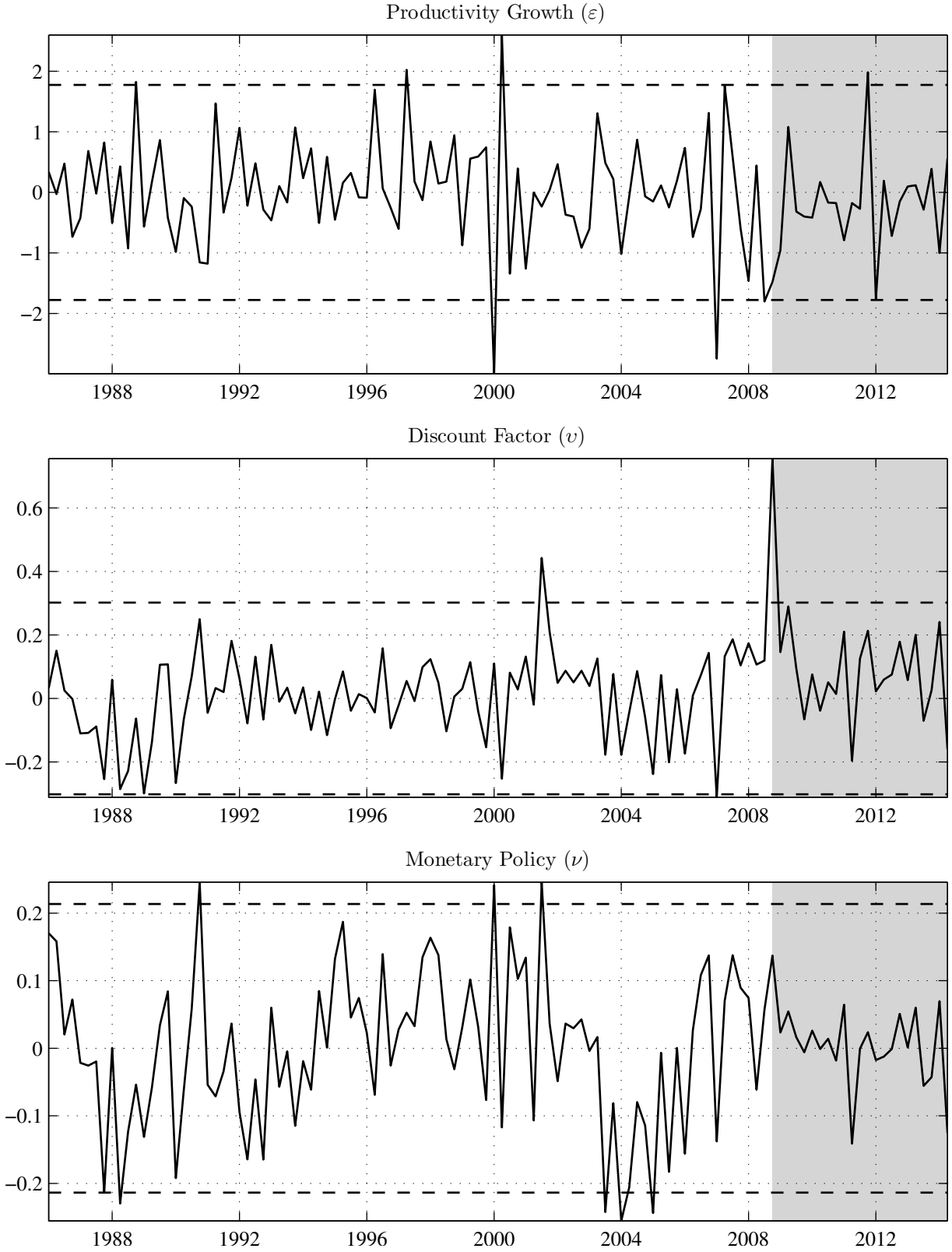

Figure 5: Median paths of the estimated shocks. The horizontal dashed lines denote a two standard deviation shock.

at least 2 draws to reduce the effects of serial correlation, we failed to reject that the sample means at the start and end of the chain were equal at a 5% confidence level for every estimated parameter.

We also applied a kernel density estimator to the draws from our posterior distribution to infer the probability density functions of each parameter. The resulting densities are then compared to the corresponding prior density functions in figure 3. In most cases, the mean of the posterior distribution differs from the prior distribution and the variance is smaller. Both of those results indicate that the parameters are well informed. As further evidence, we estimated the model with the measurement error variances set to 5% and 20% of the data variance. The results were similar.

Figure 4 compares the median paths of real GDP growth, the inflation rate, and the policy rate predicted by the model (solid lines) to their time series in the data (dashed lines). Figure 5 shows the corresponding sequences of productivity growth, discount factor, and monetary policy shocks. To calculate the model predictions and the shocks, we first filter the model and compute the mean path for each draw from our posterior distribution and then use those paths to find the median time series. Our results show that the model fits data well. It matches the regular fluctuations in real GDP growth and inflation, including the sharp declines that occur at the onset of the Great Recession. It also captures the lower-frequency movements in the federal funds rate and, most importantly, the long ZLB period, which is caused by a large discount factor shock in 2008Q4.

5 EMPIRICAL FIT OF THE STRUCTURAL MODEL

	Real GDP Growth ($\hat{y}_t^{gdp}$)		Inflation Rate (π_t)		Interest Rate (i_t)	
	Mean	SD	Mean	SD	Mean	SD
Data	1.44	2.45	2.25	0.97	3.92	2.70
Model	1.54	2.48	2.50	0.96	4.70	1.73
	$(0.66, 2.43)$	$(2.02, 3.04)$	$(1.99, 3.00)$	$(0.73, 1.25)$	$(3.35, 6.09)$	$(1.18, 2.40)$
	Autocorrelations			Cross-Correlations		
	$(\hat{y}_t^{gdp}, \hat{y}_{t-1}^{gdp})$	(π_t, π_{t-1})	(i_t, i_{t-1})	$(\hat{y}_t^{gdp}, \pi_t)$	$(\hat{y}_t^{gdp}, i_t)$	(π_t, i_t)
Data	0.30	0.63	0.99	0.01	0.18	0.47
Model	0.49	0.72	0.89	-0.40	0.06	0.30
	$(0.29, 0.66)$	$(0.60, 0.83)$	$(0.81, 0.95)$	$(-0.63, -0.12)$	$(-0.22, 0.33)$	$(-0.04, 0.60)$

Table 2: Unconditional moments. For each draw from the posterior distribution, we run $10{,}000$ simulations with the same length as the data. To compute the moments, we first calculate time averages and then the means and quantiles across the simulations. The values in parentheses are $(5\%, 95\%)$ credible sets. All values are annualized net rates.

Table 2 compares unconditional moments based on simulations of our model to equivalent statistics in the data. For each posterior draw, we simulate the model $10{,}000$ times for the same number of quarters as our sample. Each simulation is initialised with a state vector drawn from the model's ergodic distribution. The table reports the average and $(5\%, 95\%)$ credible sets of the moments across the posterior draws and the simulations. Of all the moments we report, the mean, standard deviation, and autocorrelation of real GDP growth are the most important to our research question about the relationship between real GDP growth and its uncertainty. For all three statistics, the values in the data are near the mean and within the credible sets predicted by the

model. Also, the mean, standard deviation, and autocorrelation of the inflation rate and the mean of the policy rate are within the model's credible sets. The model does not do as well matching the volatility and autocorrelation of the policy rate. The policy rate is highest in the 1980s and when we remove this period, the volatility is closer to the model's prediction. Also, the federal funds rate is our only observable that has a long period of near zero volatility, which is over-represented in the data relative to the unconditional simulations. The continued variation in real GDP growth and the inflation rate during the ZLB episode helps those variables better match their moments. The cross-correlation in the data between real GDP growth and inflation is outside the model's credible set, but both cross-correlations with the policy rate are near the mean values predicted by model.

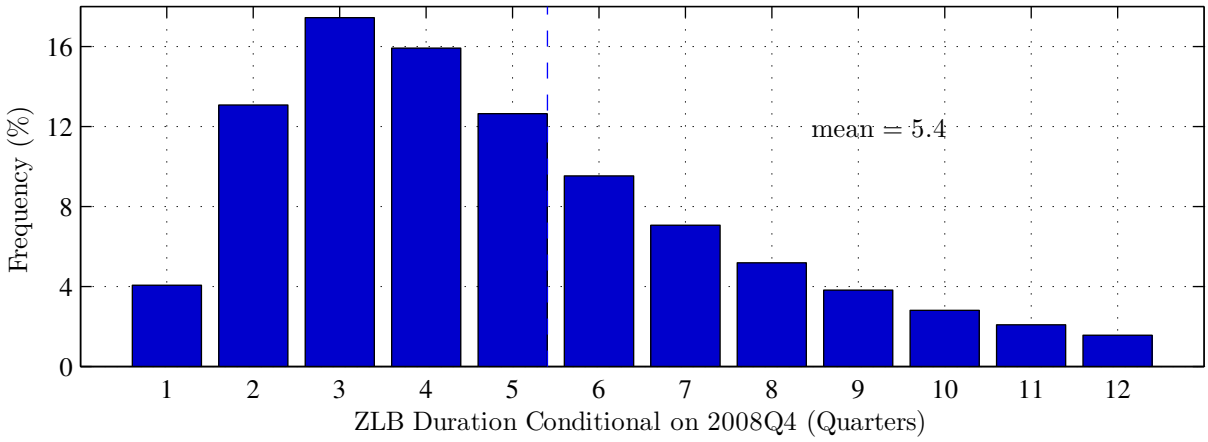

Figure 6: Distribution of ZLB events in our model. The vertical dashed line represents the expected ZLB duration.

Another key test of our model is whether it produces ZLB durations that are consistent with forecasters' expectations. Prior to the FOMC's August 2011 date-based forward guidance, survey data indicated the 3-month T-bill rate was not expected to remain near zero for very long. Blue Chip consensus forecasts between 2008Q4 and 2010Q4 reveal that the 3-month T-bill rate was expected to exceed 0.5% within three or four quarters. Figure 6 plots histograms of the durations of each ZLB event in our model. To compute these histograms, we initialise $10{,}000$ simulations at the filtered state corresponding to 2008Q4, so economic conditions are similar to what forecasters faced. We then count the number of quarters in the first ZLB event for each simulation and report the frequency of the ZLB event durations across the $10{,}000$ simulations. The most likely ZLB duration is 3 quarters and the expected duration is 5.4 quarters, which reflects a long right tail. Those results show that our model produces ZLB events that are in line with survey data. Furthermore, when we filter the data, the model is able to capture the entire ZLB event from 2008Q4 to 2014Q2.

6 GENERALISED IMPULSE RESPONSE FUNCTIONS

The general procedure for calculating GIRFs is described in Koop et al. (1996). The GIRFs are based on the average path from repeated model simulations and generated with the following steps:

1. Initialise each simulation with the desired state vector, $\mathbf{z}_0$. That vector is either the stochastic steady state or the median filtered state vector for a specific quarter (e.g., 2008Q4). To calculate the stochastic steady state, we turn off all shocks and simulate the model. The values

that the simulation converges to is defined as the stochastic steady state, which differs from the deterministic steady state because the policy functions embed higher order moments.

2. Draw random productivity growth, discount factor, and policy shocks, $\{\varepsilon_t, \upsilon_t, \nu_t\}_{t=0}^N$, for each simulation, where N is the simulation length. From the initial state vector, $\mathbf{z}_0$, simulate R equilibrium paths, $\{\mathbf{x}_t^j(\mathbf{z}_0)\}_{t=0}^N$, where $j \in \{1, 2, \ldots, R\}$ and $R = 100{,}000$.
3. Using the same R draws of shocks from step 2, replace the discount factor shock in period one with a 2 SD shock (i.e., set $\upsilon_1 = 2\sigma_\upsilon$ for all $j \in \{1, 2, \ldots, R\}$). Then simulate the model with these alternate sequences of shocks to obtain R equilibrium paths, $\{\mathbf{x}_t^j(\mathbf{z}_0, \upsilon_1)\}_{t=0}^N$.
4. Average across the R simulations from step 2 and step 3 to obtain average paths given by

$$\bar{\mathbf{x}}_t(\mathbf{z}_0) = R^{-1} \textstyle\sum_{j=1}^R \mathbf{x}_t^j(\mathbf{z}_0), \qquad \bar{\mathbf{x}}_t(\mathbf{z}_0, \upsilon_1) = R^{-1} \textstyle\sum_{j=1}^R \mathbf{x}_t^j(\mathbf{z}_0, \upsilon_1).$$

5. The difference between the two average paths is a GIRF.

7 VAR MODEL AND ESTIMATION APPENDIX

7.1 MODEL Following Primiceri (2005), we begin by defining the structural model with k lags:

$$A_t y_t = f_t + F_{1,t} y_{t-1} + \ldots + F_{k,t} y_{t-k} + \Sigma_t \varepsilon_t, \quad t = 1, \ldots, T, \tag{1}$$

where y_t is a $n \times 1$ vector of observed variables, f_t is a $n \times 1$ constant, $F_{i,t}$ are $n \times n$ coefficient matrices, $\Sigma_t = \text{diag}(\sigma_{1,t}, \ldots, \sigma_{n,t})$, and ε_t has a multivariate standard normal distribution. Our recursive identification scheme assumes A_t is lower triangular, meaning for $i, j = 1, \ldots, n$,

$$A_t = \begin{cases} 1 & i = j \\ \alpha_{i,j,t} & i > j \\ 0 & i < j \end{cases}.$$

Premultiplying (1) by A_t^{-1} yields the reduced form VAR model, given by,

$$y_t = b_t + B_{1,t} y_{t-1} + \ldots + B_{k,t} y_{t-k} + A_t^{-1} \Sigma_t \varepsilon_t,$$

where $b_t = A_t^{-1} f_t$ and $B_{i,t} = A_t^{-1} F_{i,t}$. We define the row vector of all right-hand-side coefficients as $\beta_t' = [b_t', \text{vec}(B_{1,t})', \ldots, \text{vec}(B_{1,t})']$ and the corresponding $n \times n(1 + nk)$ matrix of regressors as $X_{t-1} = I_n \otimes (1, y_{t-1}', \ldots, y_{t-k}')$, which simplifies the model to $y_t = X_{t-1}\beta_t + A_t^{-1}\Sigma_t \varepsilon_t$.

We define $\alpha_t = (\alpha_{2,1,t}, \alpha_{3,1,t}, \alpha_{3,2,t}, \alpha_{4,1,t}, \ldots, \alpha_{n,n-1,t})'$ and $h_t = (\log(\sigma_{1,t}), \ldots, \log(\sigma_{n,t}))'$. The time-varying parameters evolve according to first-order random walk processes, given by,

$$\beta_t = \beta_{t-1} + \nu_t, \qquad \alpha_t = \alpha_{t-1} + \zeta_t, \qquad h_t = h_{t-1} + \eta_t,$$

where the disturbances follow a multivariate normal distribution, given by,

$$\begin{bmatrix} \varepsilon_t \\ \nu_t \\ \zeta_t \\ \eta_t \end{bmatrix} \sim \mathbb{N}(0, V), \qquad V \equiv \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & Q & 0 & 0 \\ 0 & 0 & S & 0 \\ 0 & 0 & 0 & W \end{bmatrix}.$$

I is an identity matrix, and S is a block diagonal matrix with a block for each row of matrix A.

7.2 PRIORS The priors for the initial states, β_0, α_0, and h_0, are normally distributed, while the priors for the hyperparameters, Q, S, and W, are distributed as inverse-Wishart. Those distributions are calibrated with OLS point estimates and standard errors from an analogous time-invariant VAR, with a training sample that precedes the one of interest. Specifically, the priors are given by

$$\begin{aligned}
\beta_0 &\sim \mathbb{N}(\hat{\beta}_{OLS}, 4\,\mathrm{var}(\hat{\beta}_{OLS})), & \alpha_0 &\sim \mathbb{N}(\hat{\alpha}_{OLS}, 4\,\mathrm{var}(\hat{\alpha}_{OLS})),\\
h_0 &\sim \mathbb{N}(\log \hat{\sigma}_{OLS}, I_n)), & Q &\sim \mathbb{IW}(40k_Q^2\,\mathrm{var}(\hat{\beta}_{OLS}), 40),\\
W &\sim \mathbb{IW}(4k_W^2 I_n, 4), & S_i &\sim \mathbb{IW}((i+1)k_S^2\,\mathrm{var}(\hat{\alpha}_{i,OLS}), 2),
\end{aligned}$$

where $k_Q = 0.01$, $k_S = 0.1$, $k_W = 0.01$, and i indexes the rows of matrix A.

7.3 ESTIMATION PROCEDURE The model is estimated with a Gibbs sampler, which permits intractable likelihood functions and posterior distributions without an analytic normalizing constant. After initializing β, α, h, Q, W, and S_i, the sampler proceeds through the following steps:

1. Draw W given h.
2. Draw β given α, h, Q, y, which relies on a Bayesian smoother (see Carter and Kohn (1994)).
3. Draw Q given β,
4. Draw α_i given β, h, S_i, y, which also relies on a Bayesian smoother.
5. Draw S_i given α_i.
6. Draw h as follows:
   (a) Draw s from a mixture of normals given $W, \beta, Q, \alpha, S, h, y$ following Kim et al. (1998)
   (b) Draw h given $W, \beta, Q, \alpha, S, y, s$.

Each step makes draws for the entire sample and the extracted parameters are classified as smoothed estimates. See Primiceri (2005), Del Negro and Primiceri (2015), and the code from Koop and Korobilis (2010) for details on how to draw the parameters. The code is available at https://sites.google.com/site/dimitriskorobilis/matlab/code-for-vars and implements the correction in Del Negro and Primiceri (2015). Specifically, α, β, Q, S, W must be drawn conditional on h and the data before the mixture indicators, s, that are used to select the component of the mixture for each variable at each date are drawn, then h is drawn.

8 ADDITIONAL DATA SOURCES

U.K. Real GDP: Chained 2000 national currency units, seasonally adjusted. Source: OECD, Main Economic Indicators. (FRED ID: NAEXKP01GBQ652S).

U.K. GDP Deflator: Quarterly, seasonally adjusted. Source: OECD, Main Economic Indicators. (FRED ID: GBRGDPDEFQISMEI).

U.K. 3-Month Treasury Rate: Quarterly average of monthly values. Source: OECD, Main Economic Indicators. (FRED ID: IR3TTS01GBM156N).

Japan Real GDP: Chained 2005 yen, seasonally adjusted. Source: OECD, Quarterly National Accounts. (FRED ID: JPNRGDPQDSNAQ).

Japan Population: Working Age Population: Aged 15 and Over: All Persons. Source: OECD, Main Economic Indicators. (FRED ID: LFWATTTTJPQ647S).

Japan Consumer Price Index: Index 2005=100, not seasonally adjusted. Source: OECD, Main Economic Indicators Database.

Japan T-bill Rate: Quarterly average of monthly values. Source: International Monetary Fund, International Financial Statistics.

REFERENCES

ANDRADE, P., R. CRUMP, S. EUSEPI, AND E. MOENCH (2014): "Fundamental Disagreement," Banque de France Working Paper 524.

BOERO, G., J. SMITH, AND K. WALLIS (2008): "Uncertainty and Disagreement in Economic Prediction: The Bank of England Survey of External Forecasters," *Economic Journal*, 118, 1107–1127.

CARTER, C. AND R. KOHN (1994): "On Gibbs sampling for state space models," *Biometrika*, 81, 541–553.

DEL NEGRO, M. AND G. E. PRIMICERI (2015): "Time Varing Structural Vector Autoregressions and Monetary Policy: A Corrigendum," *Review of Economic Studies*, 82, 1342–1345.

GEWEKE, J. (1992): "Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments," in *Bayesian Statistics*, ed. by J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, Oxford: Clarendon Press, vol. 4, 641–649.

——— (1999): "Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication," *Econometric Reviews*, 18, 1–73.

GORDON, N. J., D. J. SALMOND, AND A. F. M. SMITH (1993): "Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation," *IEE Proceedings F - Radar and Signal Processing*, 140, 107–113.

HERBST, E. P. AND F. SCHORFHEIDE (2016): *Bayesian Estimation of DSGE Models*, Princeton University Press.

KIM, S., N. SHEPHARD, AND S. CHIB (1998): "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *The Review of Economic Studies*, 65, 361–393.

KITAGAWA, G. (1996): "Monte Carlo Filter and Smoother for Non-Gaussian Nonlinear State Space Models," *Journal of Computational and Graphical Statistics*, 5, pp. 1–25.

KOMINE, T., K. BAN, M. KAWAGOE, AND H. YOSHIDA (2009): "What Have We Learned from a Survey of Japanese Professional Forecasters? Taking Stock of Four Years of ESP Forecast Experience," ESRI Discussion Paper Series No. 214.

KOOP, G. AND D. KOROBILIS (2010): "Bayesian Multivariate Time Series Methods for Empirical Macroeconomics," *Foundations and Trends(R) in Econometrics*, 3, 267–358.

KOOP, G., M. H. PESARAN, AND S. M. POTTER (1996): "Impulse Response Analysis in Nonlinear Multivariate Models," *Journal of Econometrics*, 74, 119–147.

PRIMICERI, G. E. (2005): "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, 72, 821–852.

STEWART, L. AND P. MCCARTY, JR (1992): "Use of Bayesian Belief Networks to Fuse Continuous and Discrete Information for Target Recognition, Tracking, and Situation Assessment," *Proc. SPIE*, 1699, 177–185.