Appendix for
FEEDBACKS: FINANCIAL MARKETS AND ECONOMIC ACTIVITY

For Online Publication

Markus K. Brunnermeier, Darius Palia, Karthik A. Sastry, and Christopher A. Sims

Contents

1. Choice of Prior Distributions

As described in the main text, the model with Gaussian errors has n^2 free parameters in A_0, $(M-1)n$ free parameters in the Λ_m, and $n^2 p$ free parameters in the A_j, where n is the number of observed variables, M the number of variance regimes, T the number of observations excluding initial conditions, and p the number of lags. The mixture of normal models (including the baseline t model) have an additional nT parameters.

1.1. Shock Impacts A_0 and Relative Variances Λ_i.

The Gaussian prior for A_0, treating each element as independent with standard deviation 200 and means of zero off the diagonal and 100 on the diagonal, implicitly scales each equation to an expected prior standard deviation of 0.01. This is a standard assumption for log macro time series, and potentially a slight over-estimate for interest rates (expressed in decimal terms, with 0.01 as 1 percent).

The Dirichlet(2) prior on $\lambda_{\cdot,i}/M$, a multivariate generalization of a Beta distribution, centers the model around equal variance in each period and enforces the normalization that the average variance equal one.

1.2. Reduced Form Coefficients B_J.

For priors on the reduced form coefficients, we implement dummy observations following Sims and Zha (1996).

First, a set of "Minnesota prior" dummy observations centers belief around independent random walks in each equation, with prior precision around a mean of zero increasing for further lags. In the notation of the reference, we use a tightness of 3 and decay of 0.5.

A second set of dummy observations, a "unit root prior," expresses belief that the variables will stay at a "mean level." We estimate the prior mean as the sample mean from the lagged observations, which are not used on the left-hand-side of estimation. One observation expresses belief that *all* variables stay at the level, and another n observations express the belief that each *independently* stays at the level. Again, in the notation of the reference, we specify this with tightness 5 and persistence 1.

We turn these priors on the reduced form coefficients of B_j into priors on the structural coefficients A_j by multiplying the dummy data, along with the actual data, by A_0.

1.3. "Fat-tail shocks" $\xi_{i,t}$.

We assume the "fat-tail shocks"[1] in the normal mixture versions of the model are i.i.d. across equation and time. The former restriction is natural because, *a priori*, we do not distinguish the order of equations. The latter we claim is reasonable since the regime-changing heteroskedasticity controls low-frequency volatility movements.

[1] The main article discusses the sense in which we could call these "parameters" or "shocks" — in a Bayesian framework, they are treated symmetrically as random variables.

Step	Distribution	Method
(1)	$\mathbb{P}[\theta_1 \mid \theta_3, Y]$	Metropolis-Hastings step
(2)	$\mathbb{P}[\theta_2 \mid \theta_1, \theta_3, Y]$	Exact conditional
(3)	$\mathbb{P}[\theta_3 \mid \theta_1, \theta_2, Y] = \mathbb{P}[\theta_3 \mid \theta_2, Y]$	Exact conditional

TABLE 1. Overview of Gibbs Sampling algorithm.

In the normal mixture case, we assume three normal distributions (based on fitting residuals from the standard model) with prior probabilities $\alpha_1, \alpha_2, \alpha_3$. In the t-distributed case, we assume a prior distribution of inverse-gamma(scale = $k/2$, rate = $2/k$) to make the errors t with unit scale and k degrees of freedom.

To select parameters, we estimated a maximum likelihood fit based on residuals from the Gaussian error model. We picked standard deviations .68, 1.17, and 2.55 with prior probabilities .59, .39 and .02 in the first case. The second case has degrees of freedom 5.7. On the residuals, the difference in fit between the two models was only 0.22; this was a 179 log point improvement over a model of independent, unit-scale normals.

2. POSTERIOR SAMPLING METHOD

This section describes in detail the method we use to sample the posterior in the models with normal mixture errors.[2] The procedure for sampling the model with Normal errors is nested within. We separate the parameters into the following three blocks:

(1) $\theta_1 = \left\{A_0, (\Lambda_m)_{m=1}^{M}\right\}$. These are the contemporaneous coefficients and their scalings.
(2) $\theta_2 = \left\{(A_i)_{i=1}^{p}, (\varepsilon_{i,t})_{i,t=1}^{n,T}\right\}$. These are the structural form coefficients and the structural shocks.
(3) $\theta_3 = \left\{(\xi_{i,t})_{i,t=1}^{n,T}\right\}$. These are the variance adjustments (or high-frequency volatility shocks) for each equation in each time period. We can think of them as uniformly 1 (point mass prior) if errors are Gaussian.

and proceed with a Gibbs sampling algorithm, a special case of the Metropolis-Hastings posterior sampling algorithm that samples from the joint posterior $\mathbb{P}[\theta_1, \theta_2, \theta_3 \mid Y]$ by iteratively taking samples from conditional distributions.

Our approach, outlined in Table 1, is a somewhat non-standard Metropolis-in-Gibbs approach because the first step samples from a distribution integrated over the parameter block θ_2. This poses no additional technical challenges for proving the validity of Gibbs sampling for delivering a random sample from the joint posterior $\mathbb{P}[\theta_1, \theta_2, \theta_3 \mid Y]$. But it does greatly enhance the efficiency of the Metropolis-Hastings step by eliminating $n^2 p$

[2]The next section (C) describes in more detail the computatuional implementation of the method and its numerical convergence.

highly dependent parameters from the updating.[3] Step (3) also technically uses a marginal posterior, but that is also an exact conditional (put differently the ε_{it} are sufficient to define that conditional distrubtion).

The remainder of this section describes each block in further detail.

2.1. **First Block** ($\mathbb{P}[\theta_1 \mid \theta_3, Y]$)**.** This step requires a Metropolis update, and is the most computationally intensive of the three.

Recall the structure of the model

$$A_0 y_t = \sum_{j=1}^{p} A_j y_{t-j} + C + \varepsilon_t \tag{1}$$

with

$$\varepsilon_{i,t} \sim \text{Normal}\left(0, \lambda_{i,t}\xi_{i,t}\right) \tag{2}$$

A useful notation for each equation i of this system is

$$Y((A_0)_{i,\cdot})' = X(A_+)_{i,\cdot} + E_i \tag{3}$$

where we stack all the left-hand-side data into a $T \times n$ matrix Y, we select the ith row of A_0 (and then transpose it into a column vector), we collect all right-hand-side variables into the matrix X, stack the ith rows of the A_i and constant C into the vector $(A_+)_{i,\cdot}$, and stack all the $\varepsilon_{i,t}$ into E_i.

The first step is to augment the observed data in X and Y with "dummy observations" that implement the prior. These observations are described in more detail in the first Appendix section. This gives the augmented system of equations

$$\tilde{Y}((A_0)_{i,\cdot})' = \tilde{X}(A_+)_{i,\cdot} + E_i \tag{4}$$

Next we need to turn E_t into Gaussian errors of unit variance. One way to represent this is to construct a matrix Ξ_i with diagonal elements $\sqrt{\xi_{i,t}}$ and a matrix L_i with diagonal elements $\sqrt{\lambda_{i,t}}$ and then work with

$$\Lambda_i \Xi_i \tilde{Y}((A_0)_{i,\cdot})' = \Lambda_i \Xi_i \tilde{X}(A_+)_{i,\cdot} + \Lambda_i \Xi_i E_i \tag{5}$$

Now it is straightforward to find the posterior mode, the posterior covariance matrix for the coefficients of each equation (in $(A_+)_{i,\cdot}$), and the marginal likelihood

$$\mathbb{P}\left[Y_t \mid Y_0, A_+, \theta_1, \theta_3\right]$$

The goal for our sampling is to evaluate $\mathbb{P}\left[Y_t \mid Y_0, A_+, \theta_1, \theta_3\right]\mathbb{P}[\theta_1]$, an un-normalized posterior, and apply a standard Random Walk Metropolis (RWM) algorithm. In this case, starting from some draw $\theta_{1,j}$, we generate a proposal $\theta_1^{(k)}$ from a Gaussian distribution

[3] As will be illustrated below, the nT parameters of ε_{it} are linear functions of Y and the A_i, so they do not need to be drawn separately

with meal θ_1' and specified variance V_{θ_1}, customarily estimated from an optimization to find the posterior mode (in this case, an optimization of the model with Gaussian errors) times a scale factor. Scale factors of 0.05-0.10 were succesful in generating reasonable acceptance rates of about 20 to 30 percent.[4]

Given the proposal, we set $\theta_1^{(k+1)} = \theta_1'$ with probability

$$\min\left\{1, \frac{\mathbb{P}\left[Y_t \mid Y_0, A_+, \theta_1', \theta_3\right]\mathbb{P}[\theta_1']}{\mathbb{P}\left[Y_t \mid Y_0, A_+, \theta_1^{(j)}, \theta_3\right]\mathbb{P}[\theta_1^{(j)}]}\right\} \tag{6}$$

else we set $\theta_1^{(k+1)} = \theta_1^{(k)}$.

2.2. **Second Block** ($\mathbb{P}[\theta_2 \mid \theta_1, \theta_3, Y]$)**.** The previous section already suggested the conditional distribution of A_+, which is Gaussian with a block diagonal covariance structure. More explicitly, if $\hat{X} := \Lambda_i \Xi_i \tilde{X}$ and $\hat{Y} = \Lambda_i \Xi_i \tilde{Y}((A_0)_{i,\cdot})'$,

$$\mathbb{P}\left[(A_+)_{i,\cdot} \mid A_0, \Lambda\right] \sim \text{Normal}\left(\left(\hat{X}'\hat{X}\right)^{-1}\hat{X}'\hat{Y}, \left(\hat{X}'\hat{X}\right)^{-1}\right) \tag{7}$$

Note that if the third block is degenerate (e.g., if we want the model with Normal shocks), we can delay this block until after completing a full sample from the first. This is computationally convenient, as we may only want to compute impulse responses for some subset (often subsampling) of the full sample of (A_0, Λ).

In the general case, we also want to form the structural errors. In fact the object that is more useful to us will be $u_{i,t} := \varepsilon_{i,t}/\lambda_{i,t}$, and that is formed as

$$U_i = (\Xi_i)^{-1}\left(\hat{Y} - \hat{X}(A_+)_{i,\cdot}\right) \tag{8}$$

The implementation of this is to take the residuals from the linear regression (which is evaluated at the conditional posterior mode of A_+), subtract the correction from the deviation of the A_+ draw from that posterior mode, and then divide by the ξ_{it} element-by-element.

2.3. **Third Block** ($\mathbb{P}[\theta_3 \mid \theta_1, \theta_2, Y]$)**.** In the two cases we consider, there are simple ways to sample the posterior distributions of the $\xi_{i,t}$ with closed-form distributions. The important object, as stated previously, is $u_{i,t} := \varepsilon_{i,t}/\lambda_{i,t}$, the residual before transformation by $\xi_{i,t}$. We can observe that $\mathbb{P}[\xi_{i,t} \mid \theta_1, \theta_2] = \mathbb{P}[\xi_{i,t} \mid u_{i,t}]$, and the updates can be done in parallel for each equation and each time period.

[4] Gelman et al. (2014) suggests that RWM jump rules for a d-dimensional posterior are optimally efficient with covariance matrix $c^2\Sigma$ where $c = 2.4/\sqrt{d}$ and Σ is an estimate of the posterior covariance matrix. For $d = 150$, this formula suggests $c^2 = 0.0384$, which is close to what we use. In this context, effiency measures the number of MCMC draws to direct draws from the posterior to achieve a target variance. A more efficient sampler needs fewer draws to give us the same "performance" with respect to this (important) moment.

If $\xi_{i,t}$ has an inverse-gamma distribution with shape α and rate β, the posterior distribution conditional on $u_{i,t}$ is $\text{IG}(\alpha + 1/2, \beta + u_{i,t}^2/2)$.

In the multinomial normal mixture case, which is discussed in more detail in Appendix 7, the posterior distribution places posterior probability mass

$$p_i = \frac{\alpha_i \varphi(u_{i,t}/\sqrt{x_i})/\sqrt{x_i}}{\sum_{i=1}^{N} \alpha_i \varphi(u_{i,t}/\sqrt{x_i})/\sqrt{x_i}}$$

on each of the N possible values x_i, where α_i are the prior weights and $\varphi(\cdot)$ evaluates the Gaussian pdf.

2.4. **A note about normalization.** Our prior, because it puts positive prior means on the diagonal elements of A_0, is not invariant to permutations and scale changes of equation coefficients. As a result, we find no indication that our posterior sampling scheme is distorting results by not eliminating draws that are permutations or sign-switches of each other. We have verified that there are local peaks in the posterior at points in the parameter space that permute the ordering of variables but otherwise match the parameter values at the global peak. The prior is smaller by more than 100 log points at these permutations of the global posterior model, though, and this is enough to prevent even our MCMC samples of hundreds of thousands of draws from wandering into the neighborhoods of these local peaks. Nonetheless these methods, if applied on data for which identification did not emerge as strongly, might need to test for and eliminate permuted or sign-switched models.[5]

3. Computational Details and Algorithm Convergence

This Appendix section describes, in some detail, the implementation of the posterior sampling meethod described in Appendix B. A code supplement provides software to replicate all of this paper's empirical analysis in R.[6]

3.1. **Execution and Timing.** Major calculations were executed on two machines, (1) a 16-core, 2.60GHz Intel Xeon workstation and (2) a 4-core, 4.00 GHz Intel i7-4790K desktop. Evaluating a function proportional to the posterior of A_0 and Λ conditional on the $\xi_{i,t}$ and integrated over the coefficients A_+ (in the previous section's notation, $\mathbb{P}[\theta_1, \theta_3 \mid Y]$) took approximately 0.20 seconds for the full model (10 variables, 6 variance regimes) and the full set of data. Each posterior draw required two likelihood evalulations (for the Metropolis-Hastings step), so we could simulate about 215,000 draws per computer per 24 hours.

[5] This is a special case of the kind of normalization issue discussed by Hamilton, Waggoner and Zha (2007).

[6] Documentation within the code supplement describes, in some more detail, how to use all the programs and recreate specific figures and tables.

Obtaining a reliable sample from the model posterior(s) proved challenging. Our main results are based off of pooling draws among several dispersed MCMC chains. We believe this approach should give conservative estimates of parameter variance and allay concerns that our algorithm sampled only a small subset of the parameter space (e.g., around one thin peak but ignoring other high density areas).

For each model, we ran two stages of MCMC with an adaptive step in the middle. Before any runs, we ran an optimization algorithm on the posterior of the normal errors model and obtained posterior mode x_0 and estimated inverse Hessian matrix C_0. Our first stage of MCMC consisted of several parallel chains started from dispersed starting points $x_0 + N(0, C_0)$, where $N(0, C_0)$ was a Gaussian random vector with zero main and covariance C_0. Each chain used random-walk Metropolis transition density $N(0, kC_0)$, where k was adjusted to obtain about 20% acceptance on the Metropolis step (set around 0.04). In the second stage, we calculated a covariance matrix C_1 from the pooled draws of the first stage, picked an x_1 from the highest (conditional on ξ_{it}) density draw from the first stage, and set starting points $x_1 + N(0, C_1)$ and RWM transition densities $N(0, kC_1)$ tuned in the same way.

3.2. **Burn-in.** We based our estimate of appropriate "burn-in" (discarded initial draws before sufficient convergence to the posterior) on the stability of trace plots for the log posterior density (from the Metropolis step, so conditional on $(\xi_{it})_{i,t}$ and integrated over the $(A_i)_{i\geq 1}$) and the parameters. Figure 1 provides some example trace plots from selected parameters of one representative MCMC chain. The initial draw indexed 0 is more than 300,000 draws from the initialization. We take the lack of low-frequency trend in the log posterior density (and parameter) trace plots as some evidence that the sampler has stabilized around the highest density area of the posterior.

3.3. **Convergence and Effective Sample Size.** There are two general areas of concern for an iterative simulation posterior sample:

(1) Within-sequence auto-correlation reduces the "effective" amount of independent draws.
(2) The simulation sequences do not represent the target (posterior) distribution. Symptoms of this can include disagreement among chains with dispersed starting points, or significant non-stationary "drift" in certain parmeters over draws.

The simplest, within-chain measure of "effective sample size" is the appropriate n_e such that $\text{Var}[\hat{\mu}] = \text{Var}[\mu]/n_e$, where $\hat{\mu}$ is the empirical mean and σ^2 is the (consistently estimated) variance of each scalar parameter. The first three columns of Table 2 ("Within chain") gives quantiles of these ESS statistics for each parameter and for each chain (i.e., so the quantiles are over (number of chains) x (number of parameters) objects). The ESS for

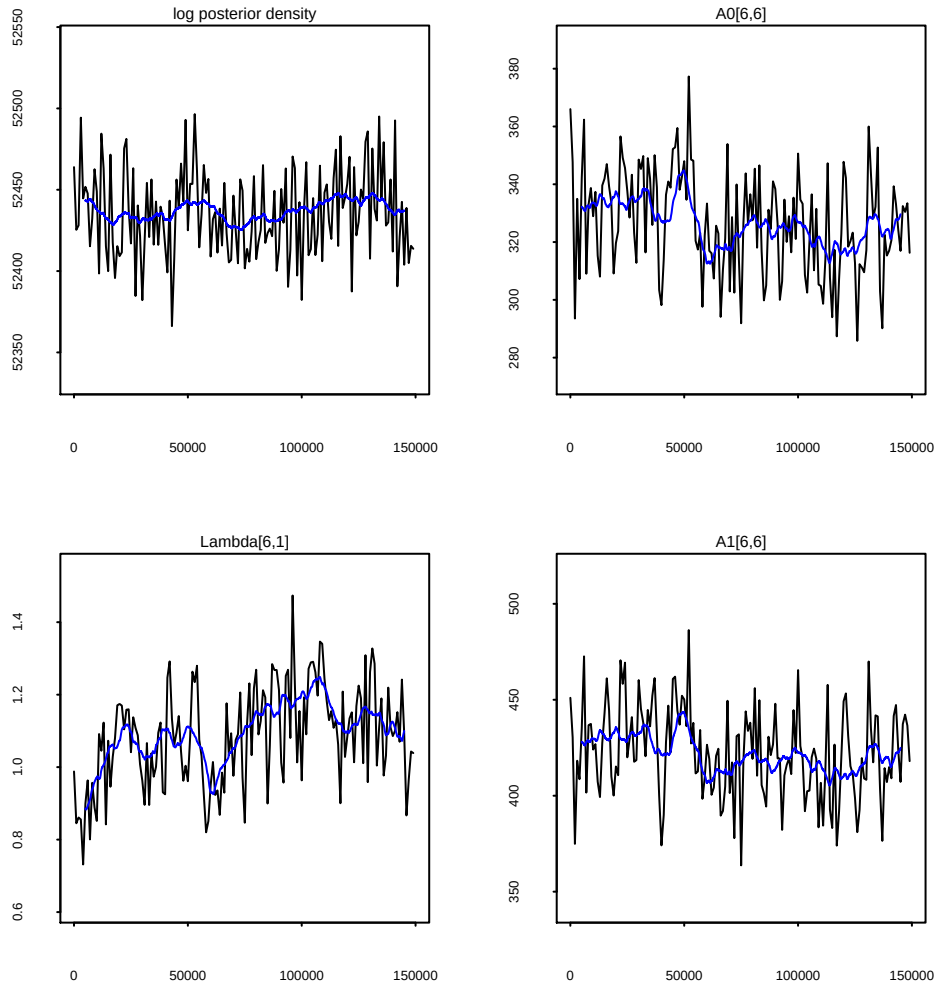

FIGURE 1. Trace plots for log posterior density (top left); the 6th diagonal entry of A_0, corresponding to the monetary policy equation (top right); the variance of the 6th (monetary policy) shock in the first period (bottom left); and the first auto-regressive coefficient for the 6th shock ($A_1[6,6]$, where the square brackets index the matrix). The black line plots every 1,000th draw, and the blue line is a two-sided moving average of 1,000 draws subsampled from every 10,000.

the impulse response parameters[7] are relatively healthy, and tend to increase linearly (if slowly) with sample size (based on incremental ESS calculations which are available upon request). Less confidence about the sampling of certain parameters of A_0 and Λ, which

[7]Hear, the impulse response parameters are the "average" response of each variable to each shock at the 1, 12, and 48 month horizons.

might also reflect relatively poor identification, is less important for the main qualitative results of our paper, so we find it less concerning.

The next three columns of Table 2 gives an "across chain" measure of ESS based on the individual chain variances of each parameter relative to the pooled variances. In this sense, it jointly addresses concerns (1) and (2). Across chain ESS is considerably lower for all parameters, which motivates our using pooled MCMC draws for major calculations. But ESS for the impulse responses seems relatively reasonable.

A different way of presenting the across=chain ESS measures is given in Table 3, which reports the the Gelman and Rubin (1992) "potential scale-reduction factor" defined as $\text{PSRF} = \sqrt{1 + m/\text{ESS}}$, where m is the number of MCMC chains. The PSRF can be interpreted as the scale by which the parameter distribution could be reduced if simulations were continued to infinite length. The reference Gelman et al. (2014) suggests a "safe" threshold value of $PSRF < 1.1$, which is mostly (but not completely) met by our impulse response parameters. In particular, for the t-distributed, full data model, 85% of impulse response parameters corresponding to shocks 3 (HH credit expansion), 6 (monetary policy), 9 (GZ spread), and 10 (interbank spread) are less than 1.1.

We consider these results sufficiently converged because the impulse response functions are ultimately the parameters of greatest interest for our analysis. Furthermore, pooling across several chains produces a conservative estimate of posterior variance, and the additional posterior variance that is attributed (in our sample) to cross-chain differences is, in the scale of economic interpretation and relative to the additional uncertainty in sampling the $(A_i)_{i\geq 1}$ coefficients, not so large.

3.4. **Marginal Likelihood Estimates.** We numerically estimate the marginal data density (MDD) of various models using the modified harmonic mean formula implemented by Geweke (1999). This approach uses the following identity for some probability density g (with conditioning on F omitted):

$$\begin{aligned}\mathbb{E}\left[\frac{g(\theta)}{\mathbb{P}[Y \mid \theta]}\right] &= \int \frac{g(\theta)}{\mathbb{P}[Y \mid \theta]\mathbb{P}[\theta]} P[\theta \mid Y] d\theta \\ &= \left(\int \mathbb{P}[Y \mid \theta'] \, d\theta'\right)^{-1} \end{aligned} \tag{9}$$

This is the inverse of the marginal data density that we are targeting As a Monte-Carlo approximation to the above we take the following:

$$\hat{\mathbb{E}}\left[\frac{g(\theta)}{\mathbb{P}[Y \mid \theta]}\right] = G^{-1} \sum_{i=1}^{N} \frac{g(\theta^{(i)})}{\mathbb{P}[Y \mid \theta^{(i)}]\mathbb{P}[\theta^{(i)}]} \tag{10}$$

where $i = \{1 \dots N\}$ index individual draws.

		Within chain			Across chain		
		A_0	Λ	IR	A_0	Λ	IR
t, full data	25%	82.99	57.82	1225.97	3.32	0.99	76.65
	50%	115.75	86.10	1661.85	8.42	2.76	176.22
	75%	140.62	119.66	2517.96	21.36	5.60	515.65
multinomial, full data	25%	93.76	63.89	1230.26	4.61	0.86	67.05
	50%	130.36	90.50	1614.72	10.55	3.00	192.10
	75%	161.48	118.10	2326.62	27.14	10.40	750.59
Gaussian, full data	25%	120.26	74.01	1086.44	8.34	4.09	539.94
	50%	174.72	153.15	1457.54	18.47	10.42	1120.87
	75%	211.89	205.17	2257.27	49.65	22.53	2337.71
t, data to 12/07	25%	40.87	33.97	69.11	5.57	2.26	17.59
	50%	55.30	45.97	286.92	10.96	4.58	65.67
	75%	70.06	60.06	918.15	24.84	8.79	254.72

TABLE 2. Distribution of effective sample size (ESS) measures of posterior simluation convergence across variables. "Within chain" means ESS is calculated within each chain, but quantiles are across chains (i.e., quantiles of a list of (n chains) $\times$ (n parameters) values). "Across chain" ESS is calculated using relative variance within chain and across chains. The columns are: 1) only elements of contemporaneous coefficents matrix A_0; 2) only state variances; and 3) the impulse response of each variable to each shock at the 1 month, 12 month, and 48 month horizons.

In our specific case, we rely on the Monte Carlo method to numerically integrate over $\tilde{\theta} := \{\theta_1, \theta_3\} = \{A_0, \{\Lambda_m\}_{m=1}^{M}, \{\xi_{i,t}\}_{i,t=1}^{n,T}\}$, since we have access to the posterior density already integrated over θ_2. We set $g(\tilde{\theta}) = f(\theta_1)g(\theta_3)$, where $f(\theta_1)$ is a Gaussian approximation of the distribution of θ_1, truncated to $100\rho\%$ of its probability mass to control the effects of outliers, and $g(\theta_3)$ is the appropriate prior distribution for the $\{\xi_{i,t}\}_{i,t=1}^{n,T}$. We make separate estimates for each of several MCMC chains, and use the comparison between chains to (very roughly) estimate the variance in our estimators (which, in our experiences, tends to be underestimated by what asymptotic theory in the number of draws and chains suggestions).

Table 4 gives some sense of the dispersion in our marginal data density estimates for $\rho = 0.95$. Results for other $\rho \in [0.90, 0.99]$ were very similar (i.e., within 1 log point or less). In absolute terms, the consistency between chains is low (a 5 log point difference implies a substantial $\exp(5) \approx 150$ odds ratio between two models). But the precision is enough to distinguish among our models, in the sense that no rankings are flipped when we use the minimum estimate for the higher density model and the maximum estimate for the lower density model. The last line of Table 4 shows MDD results for the model with time variation in A_0 described below in section 7.1.

		A_0	Λ	IR
t, full data	25%	1.13	1.44	1.01
	50%	1.31	1.78	1.02
	75%	1.68	2.66	1.04
	% < 1.1	24%	4%	94%
multinomial, full data	25%	1.05	1.14	1.00
	50%	1.13	1.42	1.01
	75%	1.29	2.12	1.02
	% < 1.1	38%	20%	98%
Gaussian, full data	25%	1.04	1.09	1.00
	50%	1.10	1.18	1.00
	75%	1.22	1.41	1.00
	% < 1.1	48%	36%	100%
t, data to 12/07	25%	1.07	1.27	1.01
	50%	1.21	1.48	1.04
	75%	1.46	2.02	1.15
	% < 1.1	29%	14%	68%

TABLE 3. Distribution of potential scale factor reduction (PSRF) measures of posterior simluation convergence across variables. The columns are: 1) only elements of contemporaneous coefficents matrix A_0; 2) only state variances; and 3) the impulse response of each variable to each shock at the 1 month, 12 month, and 48 month horizons. Numbers are based on 6000 draws per chain subsampled from 60,000 (first three models) and 4000 draws subsampled from 40,000 (t up to 12/07 model).

Model	Chains	Median	Min.	Max.
t, full data	6	21434	21420	21448
multinomial, full data	3	21344	21344	21345
Gaussian, full data	4	20712	20705	20714
Gaussian, full data with 3-year growth	3	20195	20170	20204
Gaussian, time-varying A_0	4	19511	19500	19520

TABLE 4. Marginal data density estimates for different models and different truncations of the Gaussian weighting distribution.

4. COMPARISON OF SHOCK IDENTIFICATION

Table 5 shows the correlation of several leading monetary instruments with the structural shocks estimated in our model (taken from the posterior mode of our t-errors model). Formally, this is related to the regression one would run to rotate the matrix of shocks in a "Proxy SVAR" model. The shocks are surprises to the current month and three-month-ahead Fed Funds futures (FF1 and FF4), the same shocks with the gap between Greenbook and private-sector forecasts projected out (the same shocks with "(GB)"), and a version

VAR Shock	External Shock				
	FF1	FF4	FF1 (GB)	FF4 (GB)	RR
IP	-0.041	-0.007	-0.092	-0.102	-0.024
P	-0.079	-0.016	-0.058	-0.006	0.050
HHC	-0.040	-0.117	-0.037	-0.119	-0.125
BC	-0.084	-0.060	-0.107	-0.096	0.093
M1	-0.024	0.024	-0.075	0.023	0.032
Mon. Pol.	0.536	0.479	0.334	0.244	0.315
PCM	0.029	-0.005	0.025	-0.012	0.069
TS	0.042	0.032	0.006	0.007	0.053
GZ	0.172	0.106	0.194	0.141	0.082
ES	-0.187	-0.237	-0.042	-0.068	-0.095
Start date	2/1990	2/1990	2/1990	2/1990	1/1983
End date	12/2007	12/2007	12/2007	12/2007	12/2007

TABLE 5. Correlation of external monetary instruments (as reported by Ramey, 2016) with the posterior mode structural shocks from the t-errors model. Green color denotes the largest correlations in absolute value.

of the Romer and Romer (2004) narrative shock constructed by Johannes Weiland. All data are as constructed and compiled by Ramey (2016). All versions of the instrument correlate most with the monetary shock, though there is some persistent ambiguity about the correlation with spread shocks.

5. CONSTRUCTION OF CREDIT DATA

This section provides more details about our selection and treatment of aggregate credit data and a comparison with the quarterly frequency data available in the U.S. Flow of Funds. Later, in Appendix 9, we investigate how switching between these two data sources changes our main results.

For our main VAR analysis, we use total loans data from the Federal Reserve's H.8 "Assets and Liabilities of Commercial Banks in the United States" dataset. These data are compiled from a sample of approximately 875 domestically chartered banks and "foreign-related institutions" as of the most current release.[8] These constitute, to our knowledge, the most complete monthly or higher frequency source of historical credit data in the United States. There are conceptual barriers to extrapolating conclusions with these data to conclusions about household credit and firm credit more properly defined. Our measure of "household" credit includes both houshoeld and commercial real estate, and our measure of "firm" credit includes only commercial and industrial loans. Additionally the

[8]A full documentation can be found at https://www.federalreserve.gov/releases/h8/About.htm.

Date	Consumer	RE	C & I	Reason
Oct. 2006	-3.1	-144.6	-3.1	Acquisition from nonbank institution
Mar. 2007	0.0	80.6	0.0	Conversion to thrift bank
Oct. 2008	-15.5	-192.6	-1.6	Acquisition from nonbank institution
Nov. 2009	0.0	-89.0	-1.8	Acquisition from nonbank institution
Mar. 2010	-90.9	-7.2	-7.7	Adoption of FAS 166 Accounting Standards (1)
Apr. 2010	-272.7	-21.5	-23.2	Adoption of FAS 166 Accounting Standards (2)

TABLE 6. Changes made to commercial bank credit data, in consumer (Cons.), real estate (RE), and commercial & industrial (C & I) categories, based on release notes. Units are millions of dollars. "Household" in the dataset is the sum of consumer and real estate; "Business" is C & I.

subset of credit issued by commercial banks may not be representative of all credit in the economy. In spite of these obstacles, we argue that a commercial bank data series with some corrections for large "measurement breaks" (including bank mergers, accounting standards changes, and other "non-fundamental" sources of variation) closely matches the dynamics of more precisely defined, lower-frequency measures. Thus we expect (and find) that results would be very similar using a different, quarterly-frequency dataset.[9]

5.1. **Correction for Measurement Breaks.** At several points in the raw credit data, we found big outliers in growth rates that corresponded to special notes in the H.8 data releases. Under the assumption that these observations do not represent relevant variation in credit conditions, we manually "corrected" the data by adding or subtracting the quantities attributed to these events in the documentation (Table 6).[10] Note that these changes affect the entire time series after the break date in terms of levels.

Figures 2 and 3 show these changes in log difference units for household and business credit measures respectively. It is worth noting that our correction does not completely smooth over the accounting change in April 2010, presumably because there is some change in balance sheets in anticipation in the previous period. We suspect that the model with normal mixture errors would not be very affected by one relatively small outlier, but an alternate approach that we have not yet implemented is to correct around a linear trend for that three month stretch.

[9]An alternative option we do not pursue is interpolating the quarterly data with these monthly estimates. The Federal Reserve's G.19 "Consumer Credit" series, for instance, does something similar. We have (standard) concerns about endogeneity issues in time-series regressions.

[10]The corrections for March and April relate to the same event, which ocurred in the last week of March. Our monthly data are simple averages of these weekly data. So our solution was to make 1/4 of the correction in March (i.e., to correct for 1/4 of the averaged data which was affected) and 3/4 of the correction in April. Figures 2 and 3 show that the resulting data series look reasonably smooth.

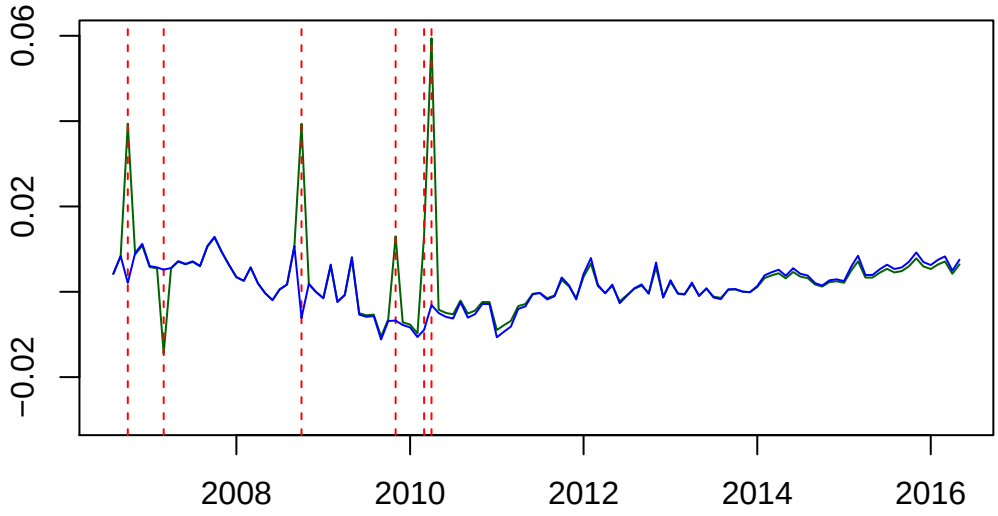

FIGURE 2. Log differences of household credit series before (green) and after (blue) corrections. Vertical red lines denote the dates of changes.

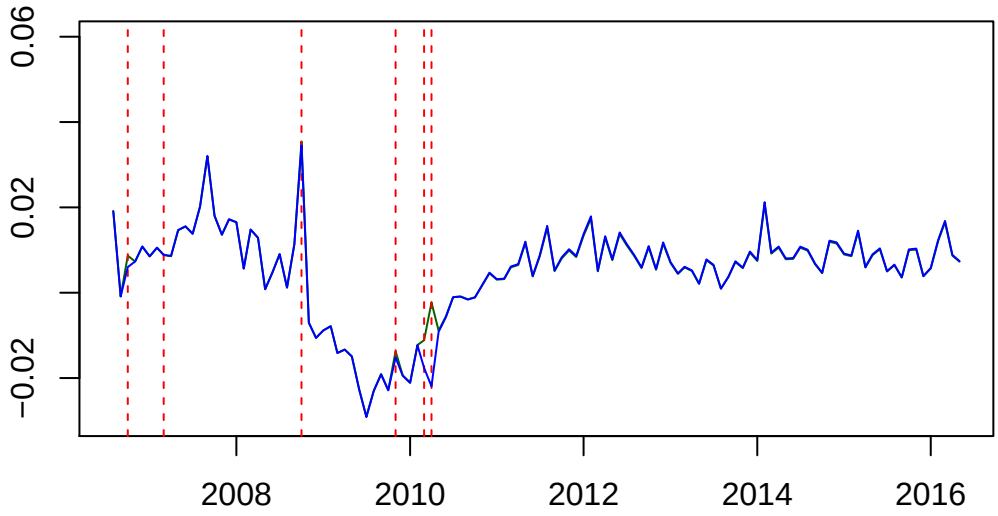

FIGURE 3. Log differences of business credit series before (green) and after (blue) corrections. Vertical red lines denote the dates of changes.

In a model with Gaussian errors and heteroskedasticity, the pre-correction data series generated outliers up to 7 standard deviations and consequently had a significant effect on model fit (entirely removing our result of a weakly significant negative output response to a household credit shock).

5.2. **Comparison with Quarterly Data.** Many other studies have used the quarterly credit series published in the Federal Reserve's Z.1 "Financial Accounts of the United States." These data are separate for the balance sheets of "Households and nonprofit organizations," "Nonfinancial corporate businesses," "Nonfinancial noncorporate businesses," "State and local governments," the "Federal government", "Domestic financial sectors," and the "Rest of the world." The Bank for International Settlements (BIS), in its multi-country panel dataset, reports versions of the "Households and nonprofit organizations" and "Nonfinancial corporate business" series adjusted for breaks.

Figure 4 compares measures of household credit in our dataset versus the single series in the quarterly dataset. In general, we find that the sum of consumer and real estate credi closely approximates the dynamics of the quarterly series, though perhaps with an extra large contraction after the 2008 financial crisis. A series constructed with just consumer credit (in blue), while it does avoid the issue of including real estate loans to businesses, seems to grow at a signficantly slower rate and include different medium-frequency (3-4 year) cyclical patterns.

Figure 5 plots a similar comparison for the business credit series, but with multiple options at the quarterly frequency. The monthly data (in red) does not so clearly approximate the blue series of nonfinancial corporate credit, which is favored by the BIS (and other panel studies which use that dataset). It does, however, capture some of the same long-term trends (including accelerations and slow-downs around business cycles), and much more closely approximates the nonfinancial noncorporate series (in green) and the average between the two quarterly series (in orange).

6. Large Residuals in the Main Model

Table 7 displays the 10 largest residuals unscaled by λ, on the left, and the 10 largest scaled (i.e. "surprise units") residuals on the right. The two 9/11 residuals are more prominent as surprises.

Five of the ten biggest surprises are in 2007-8. This concentration in time suggests that our specification of variance regime periods could be improved.

The far right column of the table shows that these residuals, though far out in the tails of the $t(5.7)$ distribution we have assumed for them, are not totally unexpected. However, with the approximately 5000 residual values from our model, the $t(5.7)$ distribution makes the expected number of surprise residuals as big as 7.52 only 1.8, whereas we have

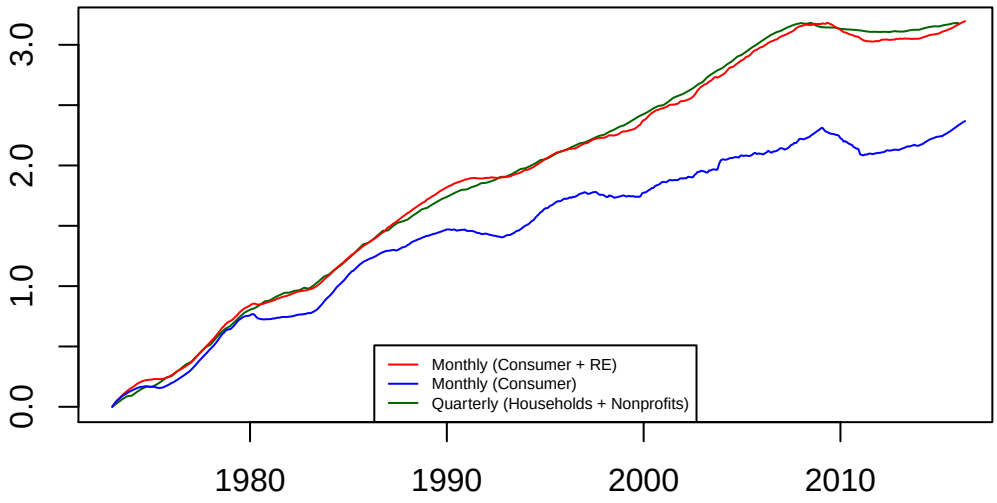

FIGURE 4. Comparison of household credit series, in log levels normalized to zero at the beginning of 1973. The data series we use in our main analysis is in red.

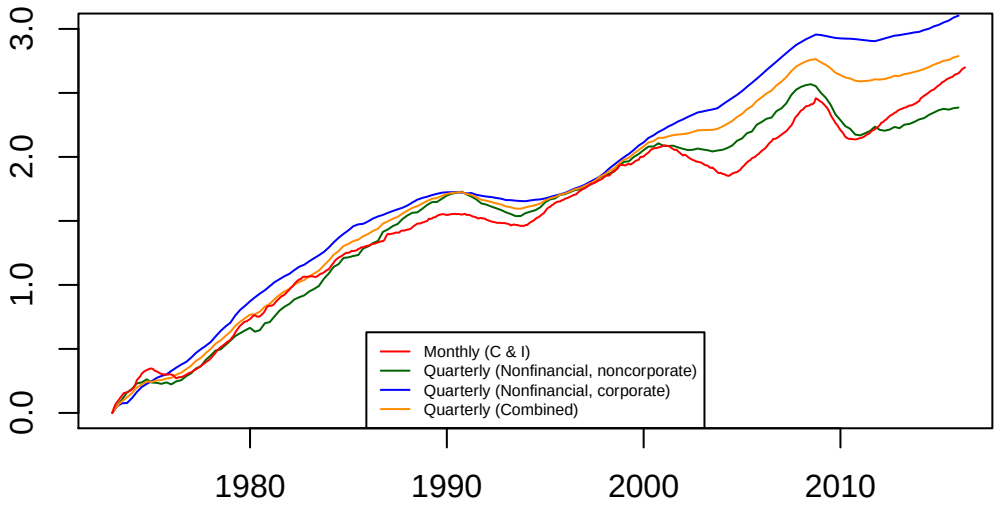

FIGURE 5. Comparison of firm credit series, in log levels normalized to zero at the beginning of 1973. The data series we use in our main analysis is in red.

Sorted by ε_{it}				Sorted by $\varepsilon_{it}/\sqrt{\lambda_{it}}$			
Month	Shock	ε_{it}	$dy_{i,t}$	Month	Shock	$\varepsilon_{it}/\sqrt{\lambda_{it}}$	Expected count
5/1980	6	-22.627	-0.067	10/2001	5	-12.667	0.109
10/2008	9	18.275	0.023	5/1980	6	-11.134	0.221
10/2008	10	10.840	0.019	9/2001	5	10.897	0.248
3/1980	6	10.409	0.031	10/2008	9	10.875	0.251
2/1981	6	-10.135	-0.030	2/2008	6	-9.082	0.668
9/2008	1	-9.875	-0.043	9/2008	1	-8.919	0.736
10/2001	5	-9.766	-0.047	10/2008	10	8.297	1.083
5/1981	6	9.312	0.028	7/2002	9	8.014	1.302
7/1974	10	8.950	0.016	8/2007	10	7.627	1.691
8/2011	5	8.507	0.041	8/1974	8	-7.518	1.824

TABLE 7. The ten largest "shocks" in the model (evaluated at the posterior median), as defined in two different ways. The left panel ranks shocks by the size of the structural residual (the "economic size" of the shock), and gives the impact on the same numbered variable as $dy_{i,t}$. The right panel ranks shocks by the size of the residual scaled by the regime variance. These should be draws from a $t(\text{df} = 5.7)$ random variable, so the final column gives the expected count of a draw this large in absolute value over 5000 observations.

estimated 10 of this size. This discrepancy suggests another direction in which our model specification could be improved.

7. Robustness of the Main Model

Here we discuss small perturbations of the paper's main VAR specification. These include changes to the distribution of structural shocks, the sample period, and the method of identification, the assumed linear specification of credit growth rates

7.1. **Models with more, or less, time variation.** The first and last lines of the main paper's Table 3 display MDD's for Gaussian-error, reduced form VAR models. The first line is for a model that keeps both Σ and $A(L)$ constant over the whole sample — i.e., fits a reduced form VAR with constant coefficients. It uses the same type of Minnesota prior as our main model, but we experimented with the prior parameter values to see if they would improve fit. The MDD reported is the best we found in these experiments. The difference between the reported MDD and what we found with the prior parameters the same as in the main model was 14 log points. As note in the text, comparing the MDD for this model with that for the main model's Gaussian-error version provides a measure of how important it is to model fit to allow for time-variation in variances. Table 4 shows the uncertainty in the MDD calculations arising out of the MCMC computations, as discussed in appendix section 3.4.

The last line of Table 3 is for a model that allows completely different reduced form VAR's in each of our regime periods. We use our Minnesota prior for each of these periods. Because the periods are relatively short, the parameter estimates are uncertain and strongly affected by the prior. Nonetheless, if there were sharp variations in $A(L)$ across regimes, this model would be expected to fit better than our main model, which it does not. Here again we experimented with parameters of the prior and display the best MDD we found, which differs from that obtained with the prior parameters in our main model by 43 log points.

We also looked at a Gaussian-errors model that relaxed our main model's restriction that A_0 remains constant across regimes, while preserving the assumption that other coefficients in $A(L)$ are constant. This model relaxes the main model's constraint that Σ_t, the reduced form residual covariance matrix, must have the same eigenvectors in each regime. It does, though, imply constraints on the nature of variation in the reduced form coefficients and in the impulse responses, and these constraints have no simple interpretation. Because the model leaves A_0 unidentified, we constrained it to take triangular form (which puts no constraint on Σ). The reasons for working with this model were that it allowed a form of time variation in $A(L)$ without adding nearly as many parameters as did the completely unrestricted model, and that it could be estimated with minor modifications of our existing code. The model did not fit well. Its MDD values are shown in appendix Table 4.

7.2. **Different Error Distributions.** Figures 6 and 7 present the full full impulse responses for the models with t and normal error distributions, respectively. Most of the impulse responses are similar, and in most cases the error bands are wider for the normal model than for the t model. This is what would be expected if the t model were correct, since in that case the normal model would still give consistent estimates, but would provide less efficient estimates.

There are a few notable differences between results from the models with Gaussian and t-distributed shocks. The differences between the two models can be seen in Figure 8, which compares the median impulse responses for IP, HHC, BC, R, PCM, GZ and ES to shocks 3 (HHC), 4 (BC), 6 (R, or monetary policy), 7 (PCM), 9 (GZ) and 10 (ES), with each shock scaled so that its largest initial component is the same size for both the t and Gaussian models. The differences between them are mainly within the error bands, with four major exceptions. The normal errors model puts more posterior probability on a nonzero output effect for the household credit shock and the interbank spread shock. The output response of the former, with normal errors, is comfortably significant (less than zero) with 68% bands and barely significant with 90% bands (which are not pictured).

The output response of the latter is comfortably significant, from the first 18 months of response, at both levels.

The Gaussian errors model shows a substantial negative response of output to shock 7 (PCM), whose largest component is a persistent increase in commodity prices. The shock also induces a modest, but significantly positive, upward movement in BC. Since this shock moves PCM immediately and BC with a delay, it is best interpreted as a commodity price shock, not an autonomous credit shock. It nonetheless implies a component of variation with a negative correlation between business credit and later output growth.

The normal model places considerably less posterior probability on a negative output response for a business credit shock, though in this case the t model's response is well within the normal model's error bands.

Note that these models are not simply alternative views of the data. The t-distributed errors model fits better than the normal model by a wide margin. To the extent that the Gaussian and t models differ, this is likely due to the Gaussian model being distorted by large outliers that are downweighted in the t model.

We also experimented with a mixture-of-normal errors model, which is not discussed in the main paper and is included in this appendix mainly as a robustness check. It sets the structural residual $\varepsilon_{i,t} \sim N(0, \lambda_{i,t}\xi_{i,t})$ with

$$\xi_{i,t} = \beta_i \text{ with probability } \alpha_i \text{ for } i \in \{1 \dots k\} \tag{11}$$

We choose $k = 3$ and set $\beta_1 < \beta_2 < \beta_3$, which provides an intuition of "low, medium, and high" variance options for each observation. Based on maximum likelihood estimation with residuals from the normal-errors model, we set $\beta_i = \{.46, 1.37, 6.50\}$ and $\alpha_i = \{.59, .39, .02\}$. The normal-mixture model produced impulse responses and fit in between the Gaussian and t models for the most part (Figure 9).

7.3. **Shortened Sample Period.** Figure 10 shows the full impulse responses from the t model estimated with data through 12/07. The impulse responses are quite similar to those from the full sample, with no differences that merit singling out. This suggests that the impact of spread shocks on macro variables is not special to the 2008-9 crash and was estimable from earlier data.

7.4. **Triangular Normalization.** Figures 11 and 12 show the full impulse responses of two Gaussian-errors models with triangular restrictions on short-run responses. The former uses these restrictions for identification (i.e., is a standard "Cholesky" recursive VAR with no heteroskedasticity). The latter includes heteroskedasticity and imposes short-run restrictions as over-identifying.

The Cholesky model includes a sizable long-term output contraction (about 0.4% over 48 months, explaining 13.0% of variance at that horizon) in response to a business loans

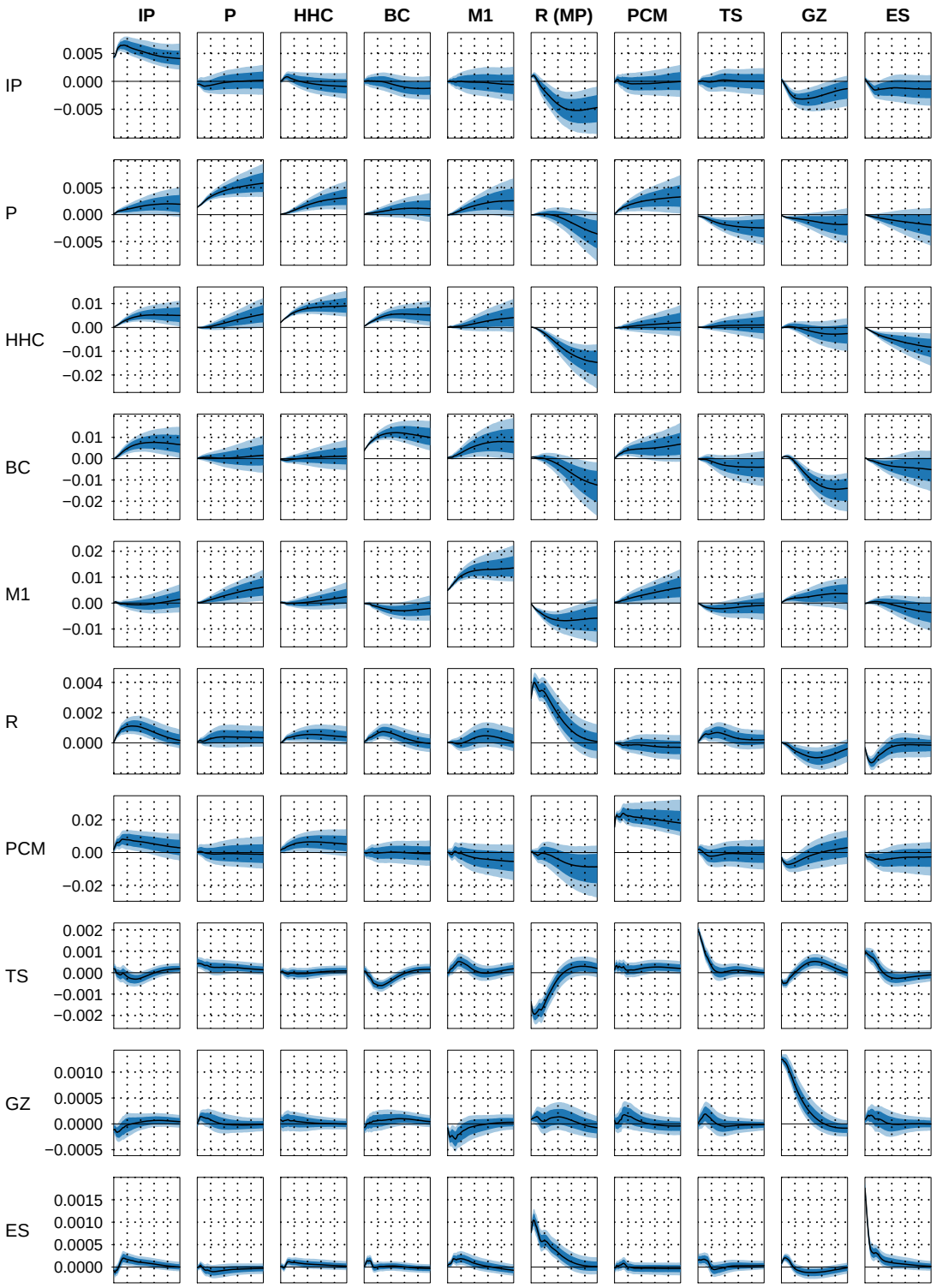

FIGURE 6. Impulse responses to the ten orthogonal structural shocks in the t errors model over 60 months, with 68% (dark blue) and 90% (light blue) posterior uncertainty regions. Scaled to an "average" period with unit scale.

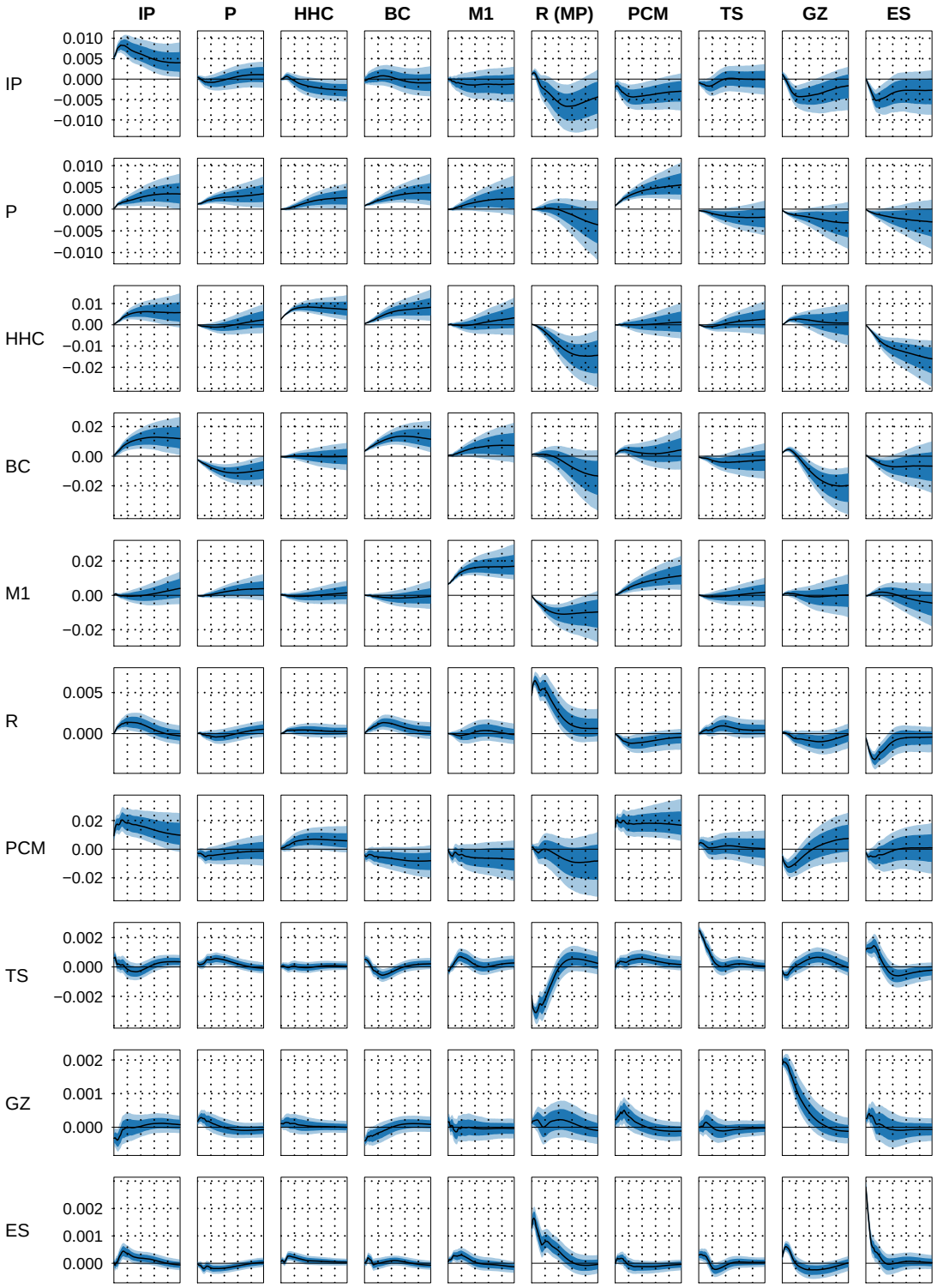

FIGURE 7. Impulse responses to the ten orthogonal structural shocks in the Gaussian errors model over 60 months, with 68% (dark blue) and 90% (light blue) posterior uncertainty regions. Scaled to an "average" period with unit variances.

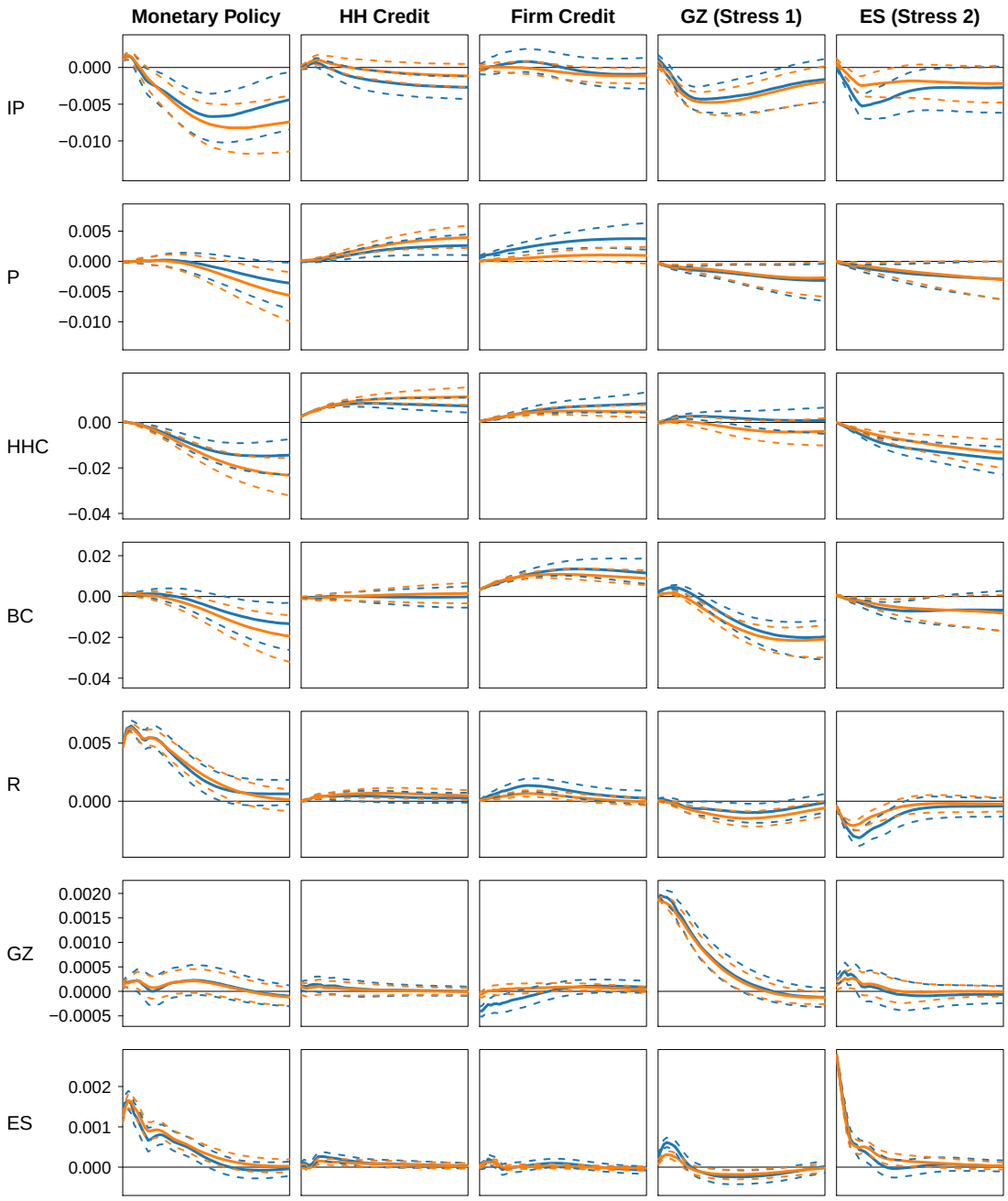

FIGURE 8. Impulse responses for shocks 3, 4, 6, 7, 9 and 10, Gaussian model in blue, t model in red, scaled to have identical sized initial shocks in HHC, BC, R, PCM, GZ and ES, respectively. The solid lines are from the posterior mode and median respectively, and dotted lines are 68% bands.

shock. This shock, like the corresponding shock in the baseline (no restrictions, with heteroskedasticity) model, includes some monetary tightening in the response to this fourth shock. The modest tightening of monetary policy and credit spreads suggests that this

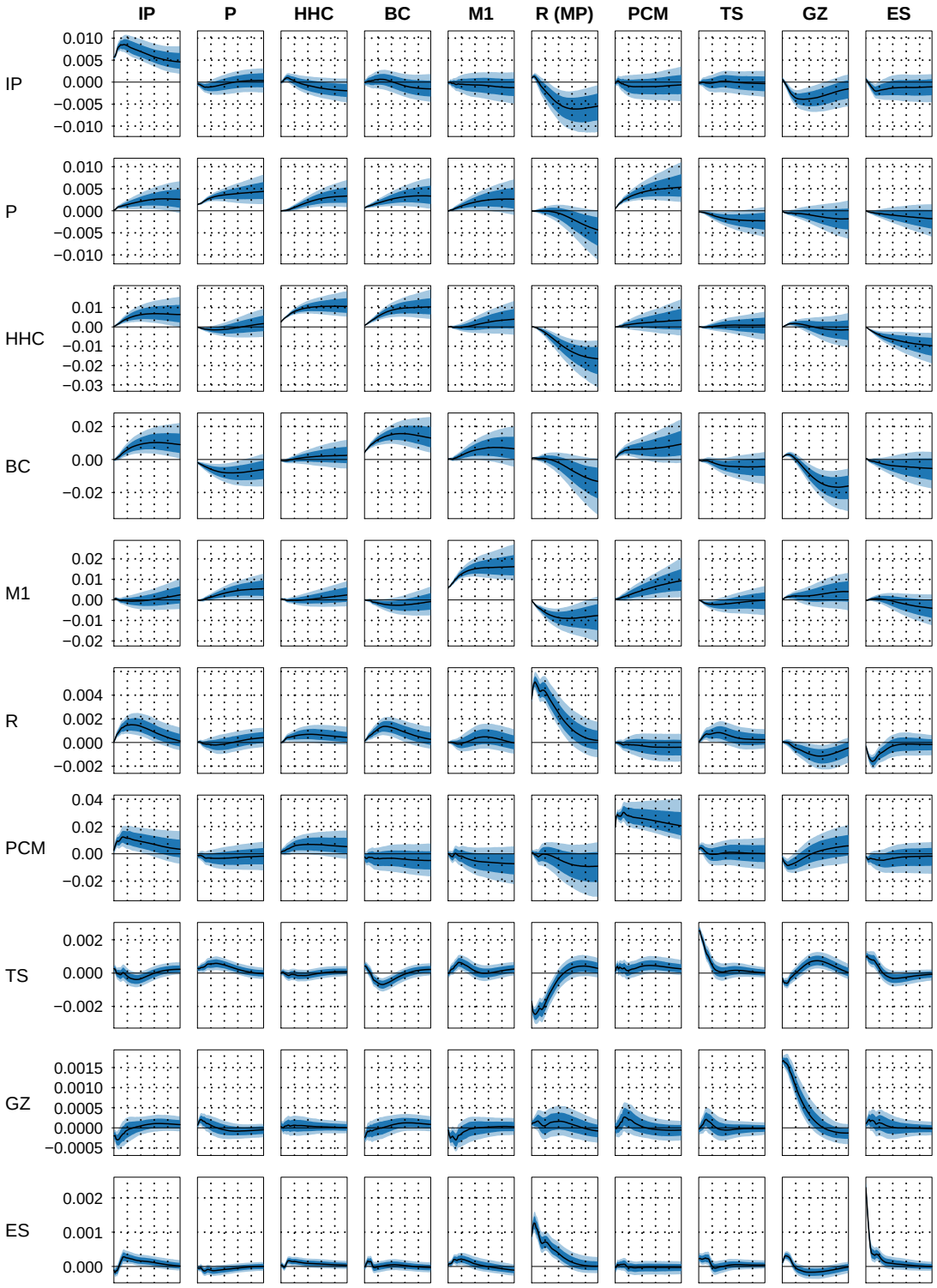

FIGURE 9. Impulse responses to the ten orthogonal structural shocks in the model with discrete normal mixture distributed errors over 60 months, with 68% (dark blue) and 90% (light blue) posterior uncertainty regions. Scaled to an "average" period with unit variances.

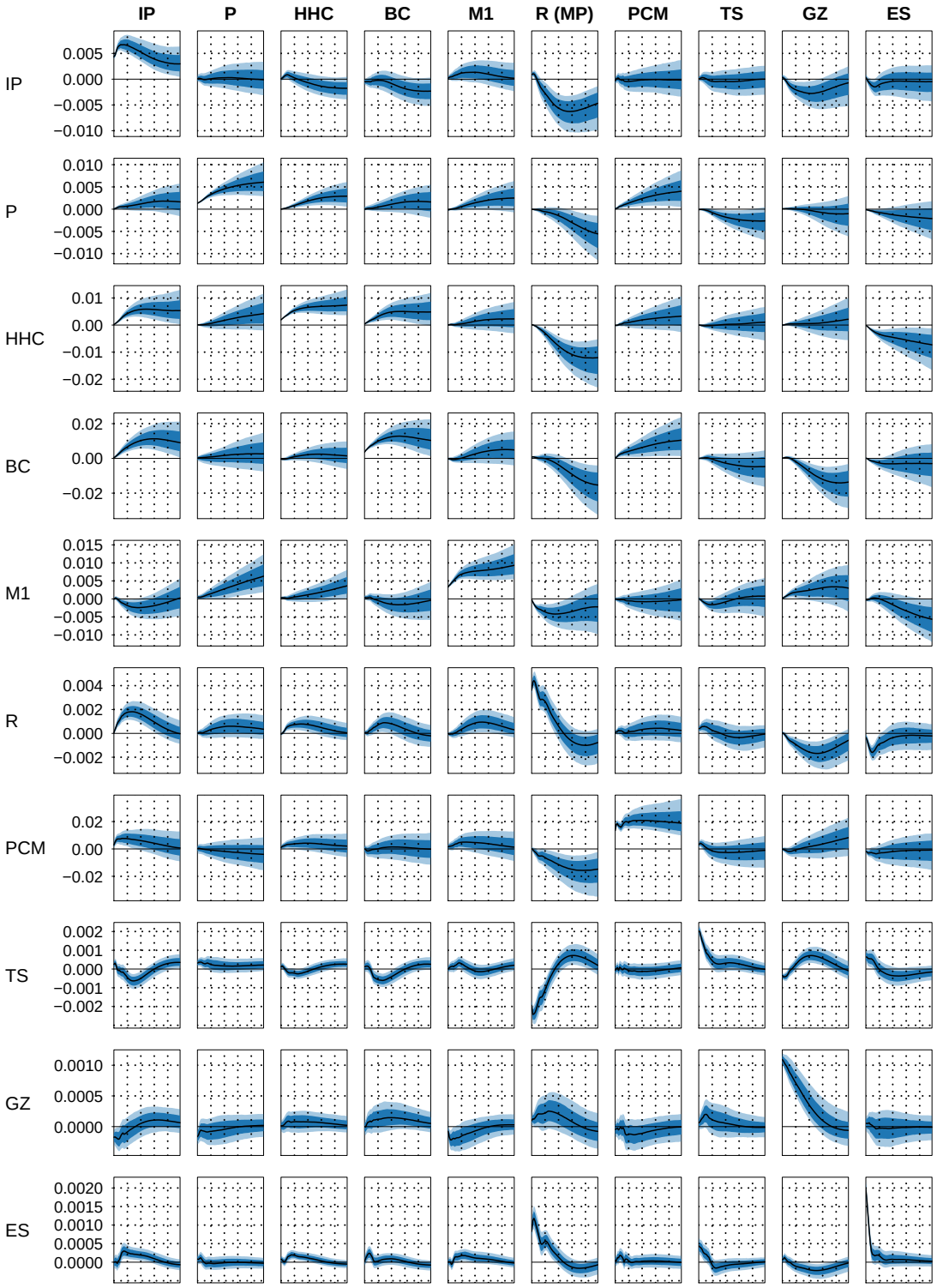

FIGURE 10. Impulse responses to the ten orthogonal structural shocks in the t errors model, estimated with data up to December 2007, over 60 months, with 68% (dark blue) and 90% (light blue) posterior uncertainty regions. Scaled to an "average" period with unit scale.

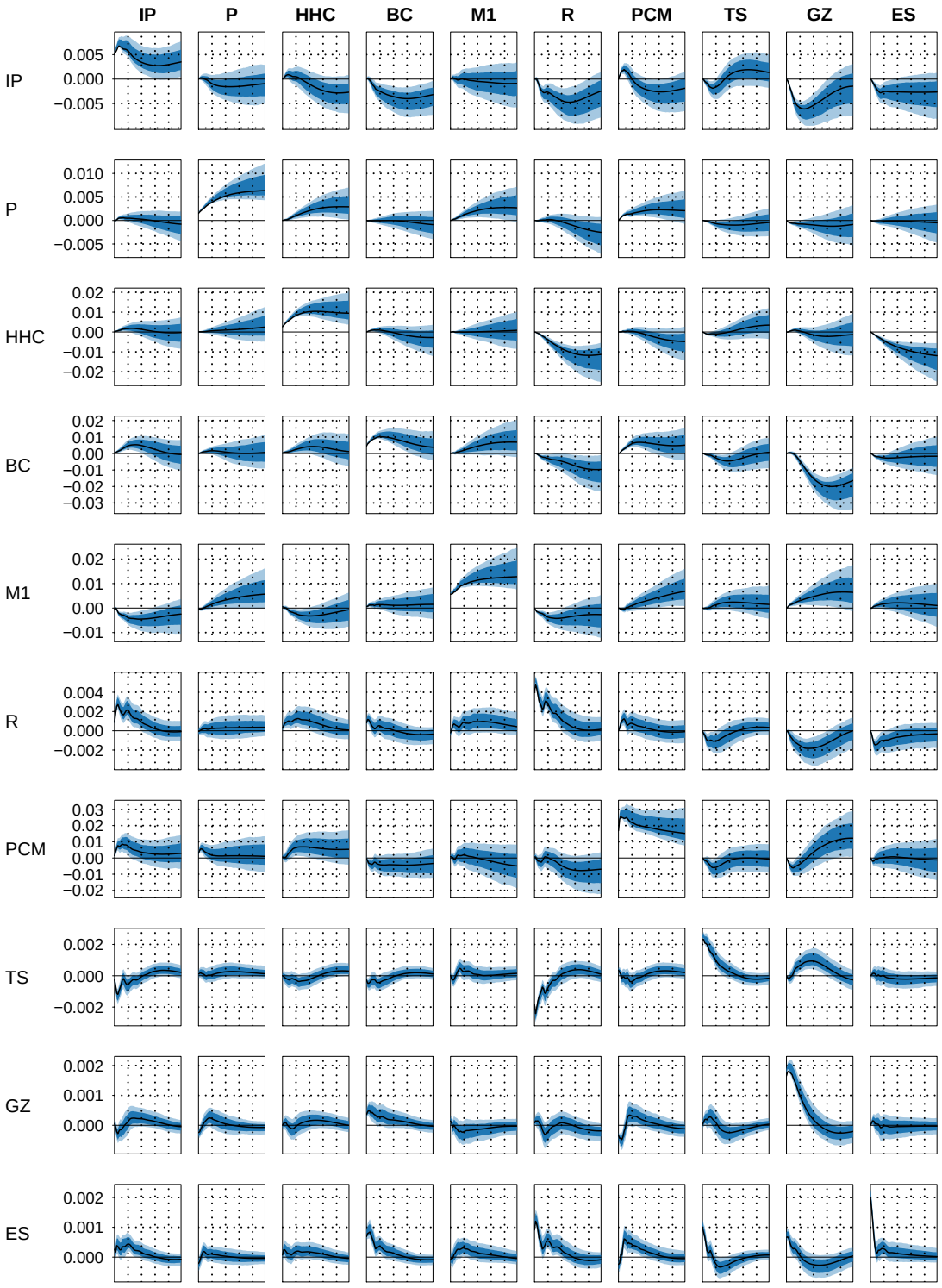

FIGURE 11. Impulse responses to 10 structural shocks in a Cholesky-identified VAR with constant structural variances over 60 months, with 68% (dark blue) and 90% (light blue) posterior uncertainty regions.

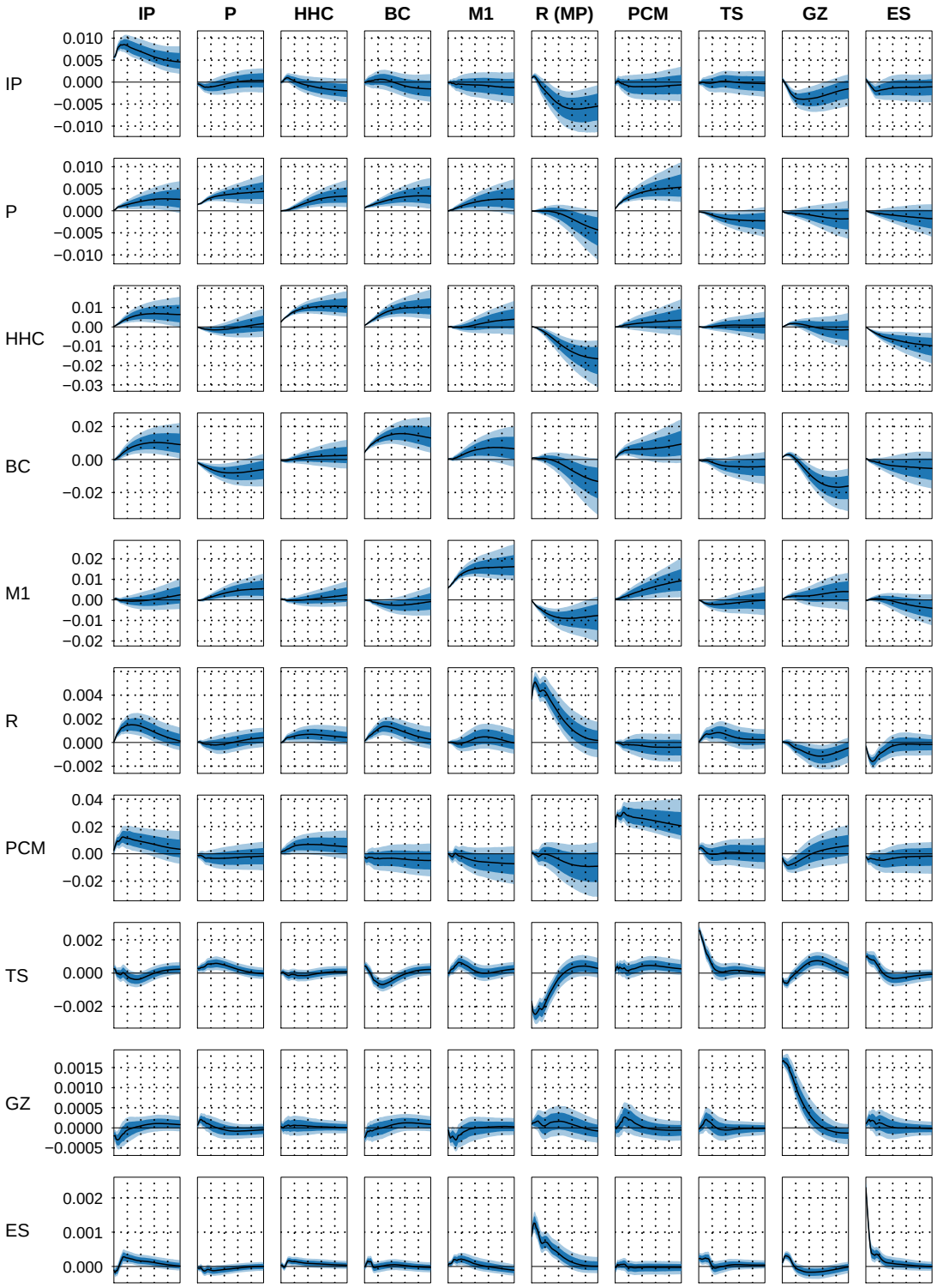

FIGURE 12. Impulse responses to 10 structural shocks in a model with triangular restrictions and heteroskedasticity over 60 months, with 68% (dark blue) and 90% (light blue) posterior uncertainty regions.

		Heteroskedasticity Model			
		4 (BC)	6 (R/MP)	9 (GZ)	10 (ES)
Cholesky Model	4 (BC)	0.996	0.159	0.148	0.135
	6 (R/MP)	-0.008	0.883	-0.016	0.114
	9 (GZ)	0.005	0.002	0.943	0.165
	10 (ES)	-0.009	-0.003	0.017	0.916

TABLE 8. Correlation among shocks (indexed by number, named in the columns) in two triangularized models, without regime heteroskedasticity (rows) and with regime heteroskedasticity (columns).

shock at least partially represents business (but not household) borrowing in the midst of both monetary stringency and financial stress. On the other hand, the overidentified (triangular model with regime-switching variance) model produces impulse responses that look much like their counterparts in the unrestricted Gaussian-errors model.

Table 8 sheds some light on the discrepancy by reporting the correlations between shocks 4 (BC, business credit), 6 (R, monetary policy), 9 (GZ, corporate bond spread), and 10 (ES, inter-bank (Eurodollar) spread) in the two models. The business credit shock in the plain, Cholesky-identified VAR (rows) "absorbs" small amounts of each of the three interest rate shocks in the model with regime switching heteroskedasticity. Breaking this down by regime illustrates the pattern more closely. In the period of aggressive Federal Reserve action (October 1979 to December 1982), the Cholesky VAR's business loans shock is very closely correlated with the other model's monetary policy shock (0.467). In the financial crisis and Great Recession (January 2008 to December 2010), the Cholesky VAR's business loans shock is positively correlated with the corporate bond and inter-bank shocks (0.644 and 0.620 respectively) and negatively correlated with the monetary policy shock (-0.206) to partially offset the effect. In this context, identification through heteroskedasticity delivers a different, and perhaps *a priori* more reasonable, separation of policy, financial stress, and credit expansion effects than a simple triangular orthogonalization.

7.5. **Long-term Growth Rates.** Our estimated system uses 10 months of lagged values, or with quarterly data four lags. Of course an autoregressive system with 10 lags can capture dynamics that operate over much longer spans of time than 10 months, but certain kinds of dynamics, e.g. pure delays of over 10 months, cannot be well approximated by a 10-lag autoregressive system. As a check, we try estimating our main (Gaussian errors) specification with three-year growth rates of household and business credit instead of levels. Figure 13 displays the impulse responses of output, prices, and the differenced credit aggregates in this model's two "credit aggregate" shocks. The household credit shock in this model explains a large proportion of long-term forecast errors in the growth

rate of credit (29% at 5 years), but still only a modest amount (6%) of GDP variation at the same horizon.

We can directly and easily compare the fit of this model and the original models through posterior odds conditional on initial conditions.[11] The differences model fits considerably worse than the levels model (see Table 4). So even if the model did imply a very strong connection between credit growth and recessions, we would have strong evidence to favor the original model.

7.6. **Non-linear Transformation.** Potentially our method is failing to capture a positive long-term relationship because it considers only linear effects. This implies that small credit movements have effects in proportion to large ones, and that negative shocks have an equal and opposite effect as positive shocks.[12]

While extensive exploration of possible nonlinearities in the model would have to be a new research project, we did try applying a smooth non-linear transformation to the 3-year growth rates of credit that allowed increased weight on large positive growth rates. The idea was to explore the hypothesis, put forward in other research, that modest credit expansion has no negative effects, while unusually rapid credit expansion does create future problems. We considered transforming credit growth according to the function

$$f(c_t) = \begin{cases} c_t & \text{if } c_t < a \\ \alpha_1 + \alpha_2 c_t + \alpha_3 c_t^2 & \text{if } a \leq c_t < b \\ \beta c_t & \text{if } c_t \geq b \end{cases}$$

for c_t as credit growth rates, $b > a > 0$, and with the coefficients α_i and β calibrated to make slopes and levels continuous at a and b. This allowed the "extra weight" $f'(c_t)$ to get larger without getting unboundedly big. We looked for a posterior mode optimizing over a, b, and β, as well as the other parameters of the model.[13] In all optimization exercises to find the posterior mode of such a model, the data favored models without any nonlinear transformation.[14]

8. Models with Fewer Variables

In this section, we estimate a VAR model with only the following variables: industrial production, the price level, and the two credit variables. It is possible to obtain results

[11] The Jacobian of any linear transformation like this is one. But now, if the model is with three three-year growth rates, we are implicitly conditioning on the values of $\{c_0 - c_{-36}, c_{-1} - c_{-37}, \ldots\}$, where c_t is the credit variable. This is different but not higher dimensional information.

[12] **?**, for instance, focus on upper n percentile credit growth events.

[13] Wwe used a coarse grid over a and b (based on matching quantiles of the observed distribution) and searched continuously over β, with an exponential prior with parameter 2

[14] In this specification, $\beta_1 = 1, \alpha_1 = \alpha_3 = 0, \alpha_2 = 1$.

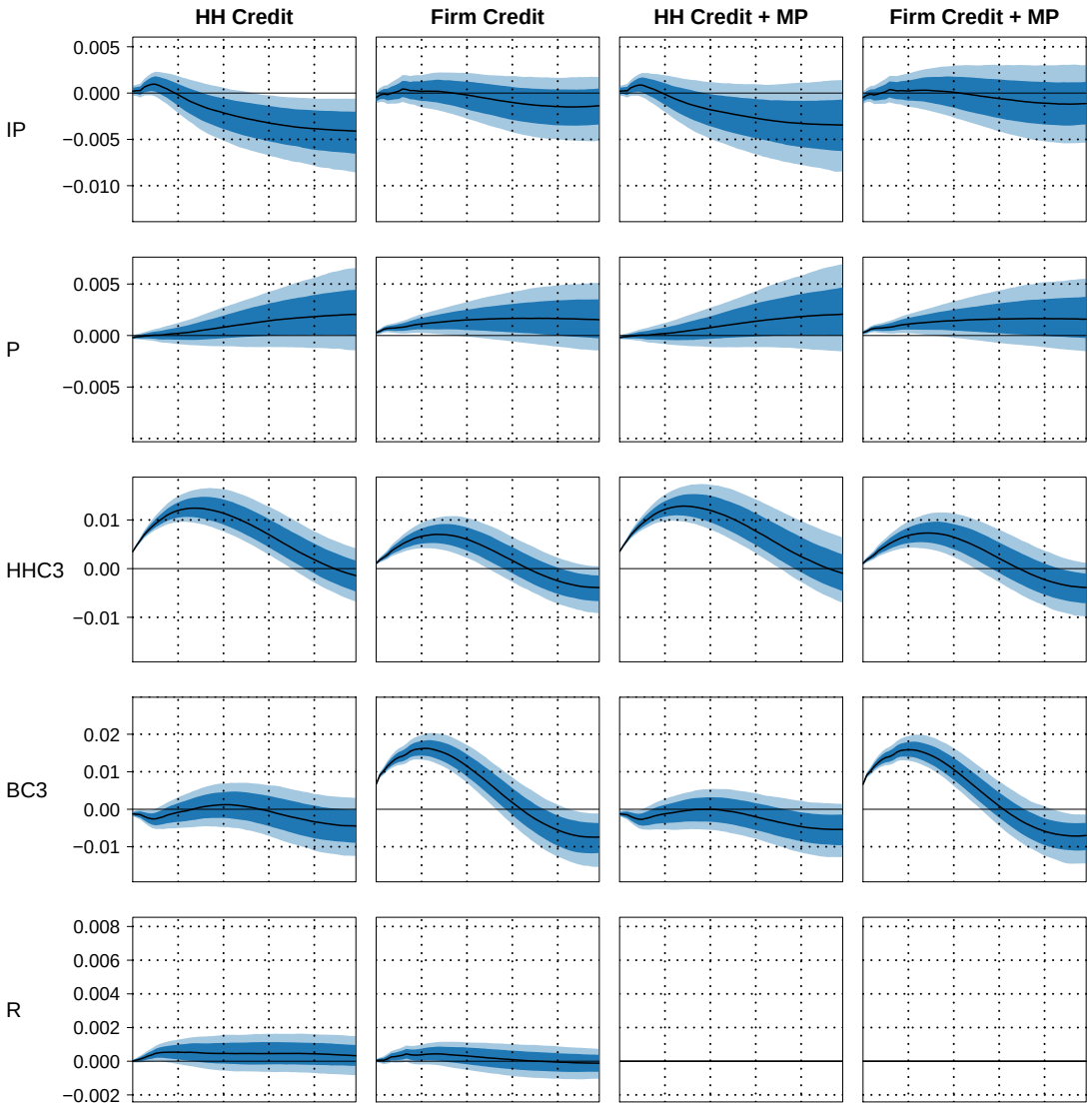

FIGURE 13. Impulse responses to candidate credit growth shocks in a (Gaussian errors) model with 3 year growth rates of credit aggregates.

suggesting that credit shocks lead to long-term output contractions. We argue based on Monte Carlo simulations (in analogy with our analysis of single-equation models in Section IV) that the small-system results are consistent with our main, larger model.

8.1. **A Small VAR Model.** Figure 14 presents the impulse response of the aforementioned four-variable model, identified with heteroskedasticity. In such a model, a shock associated a persistent increase in business loans is associated with a greater than half a percentage persistent reduction in output. This shock (at posterior mode parameters) accounts for 58% of business loan variation and 7% of output forecast error variation in one-quarter-ahead forecasts, and 22% and 8% respectively in five-year-ahead forecasts.

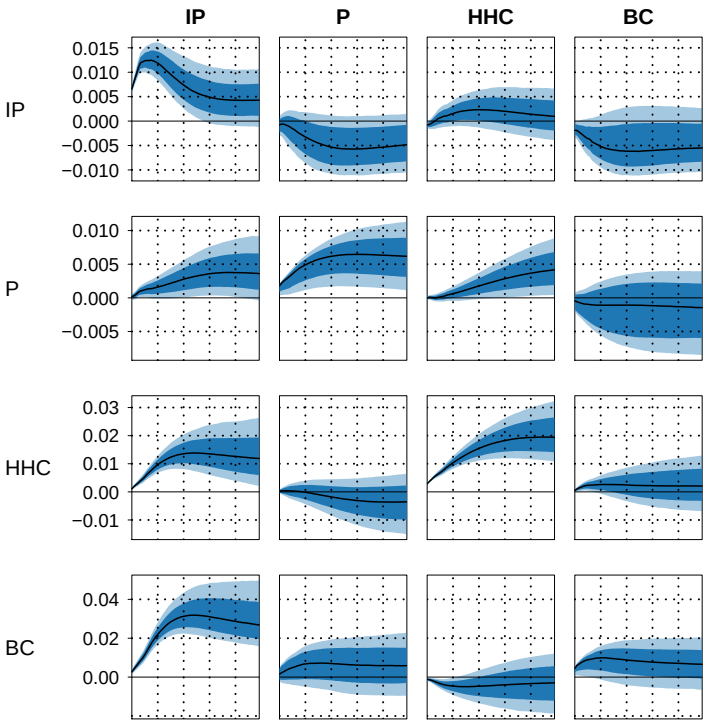

FIGURE 14. Impulse responses to four structural shocks in a "small model" over 60 months, with 68% (dark blue) and 90% (light blue) posterior uncertainty regions. Scaled to an "average" period with unit variances.

This is considerable relative to the fraction of output variation produced by the corresponding shock in the 10-variable model.[15]

If one estimates a four-variable VAR without regime-changing heteroskedasticity (and identifies structural shocks with Cholesky ordering), the results are qualitatively similar to those in Figure 14 but with tighter error bands.

8.2. **Monte Carlo Exercise.** We can shed some light on this by simulating data from the posterior mode point estimates of the large model, then estimating the same small models on the simulated data. Like in Section IV, we simulated 6,000 samples from the posterior mode of the *t*-errors model. and, for each artificial time series, estimated a four-variable Cholesky-identified VAR (including output (IP), the price level (P), household credit (HHC), and business credit (BC)).[16] We obtained negative and "large" (greater than a 0.5% output contraction) response to a one-standard deviation credit shock with the

[15]In the larger model, two shocks together explain approximately the same first period variance. One is the 4th shock (which explains 38.8% of variance at the posterior median) and the other is the 2nd shock (which explains 26.2%). Neither produces a negative correlation between business credit and output.

[16]These are the point estimate impulse response from lower triangular (Cholesky) identified VARs, with "standard" coefficient shrinkage, unit root, and covariance priors.

Proportion $\hat{\beta} < 0$		
	HH Credit	Bus. Credit
monthly	0.21	0.75
quarterly	0.29	0.75
annual	0.50	0.75
Proportion $\hat{\beta} < -.005$		
	HH Credit	Bus. Credit
monthly	0.00	0.11
quarterly	0.02	0.30
annual	0.16	0.45

TABLE 9. Probability, in simulated draws from the posterior distribution of the t-model for the data, of a negative or "economically significant" (more than 0.5% negative) 3-year response of IP to a positive credit shock in a 4-variable VAR. Based on 6,000 simulated data series. Parameter draws that implied explosive behavior were treated as not having $\hat{\beta} < 0$, so these proportions are lower bounds.

probabilities shown in Table 9. These calculations are done with data at different frequencies, aggregated by taking averages of the monthly data our model generates. Observing negative point estimates (suggesting a negative impulse response of output to credit shocks) is not so uncommon, particularly as the frequency of data is reduced to annual.

The posterior simulations show that pattern of results observed in the real data is not unlikely, with 11% of samples showing a strong negative effect of BC shocks on output and none showing an effect of HHC shocks on output. With more highly time aggregated data, as has been used in some previous studies, our model implies that the probability of finding BC shocks predict substantially negative future output growth is stronger.

9. MODELS WITH QUARTERLY DATA

Figure 15 displays the results from estimating our model with a normal error assumption on quarterly averaged data. With quarterly data, the monetary policy shock (MP) and the interbank spread shock (ES) are not well separated. The monthly data show interest rates responding with some delay, but within a couple of months, to the ES shock. The quarterly data show the interest rate response to thes ES shock as larger and within-quarter. The monthly data show the interbank spread responding to the MP shock, but not as strongly as to the ES shock. In the quarterly data, the MP shock accounts for most of the variance of the interbank spread (ES) variable. Because the quarterly data seem to

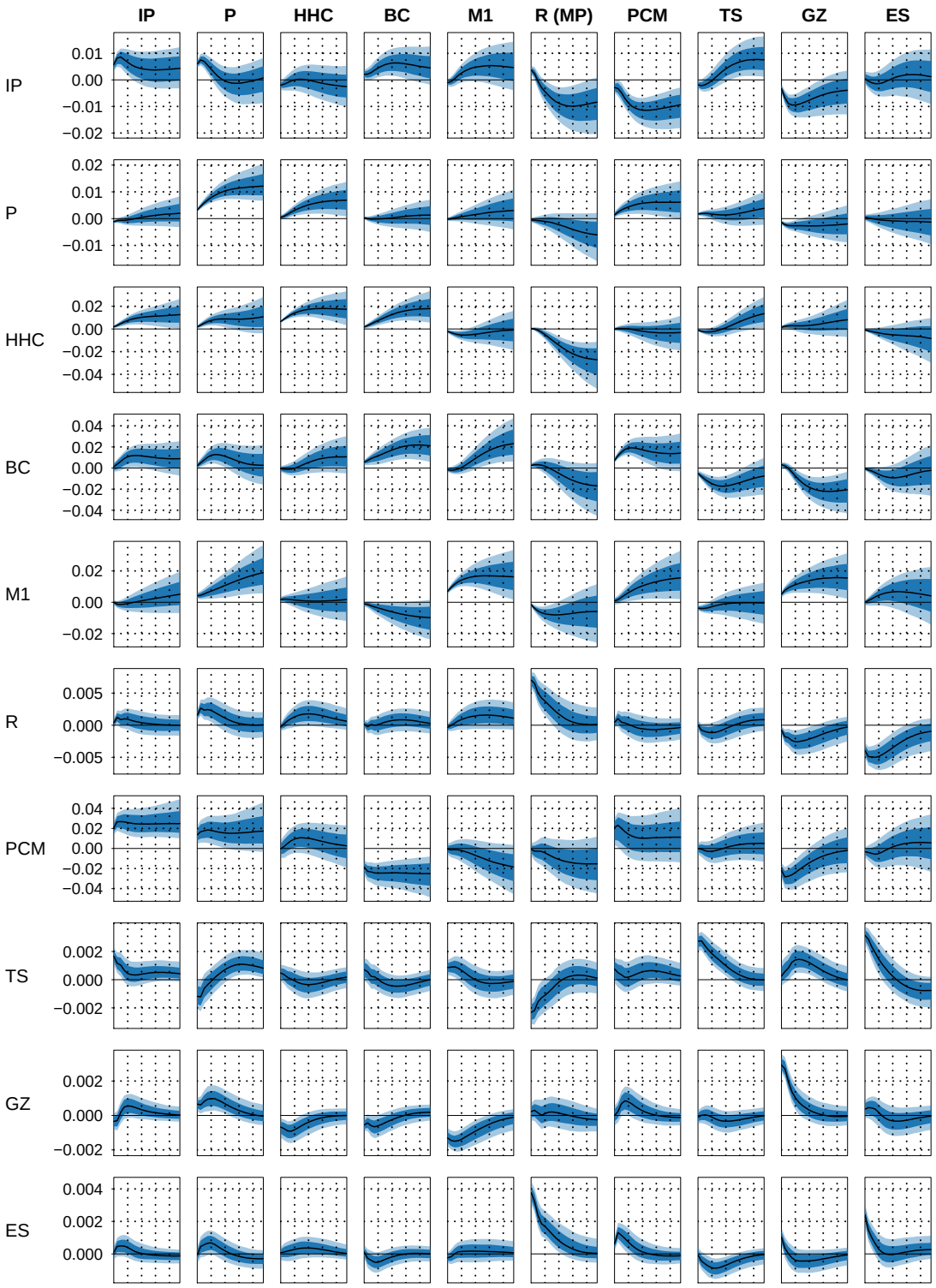

FIGURE 15. Impulse responses with quarterly-averaged data. Heteroskedasticity with change dates as in the main monthly model. Normal errors assumption.

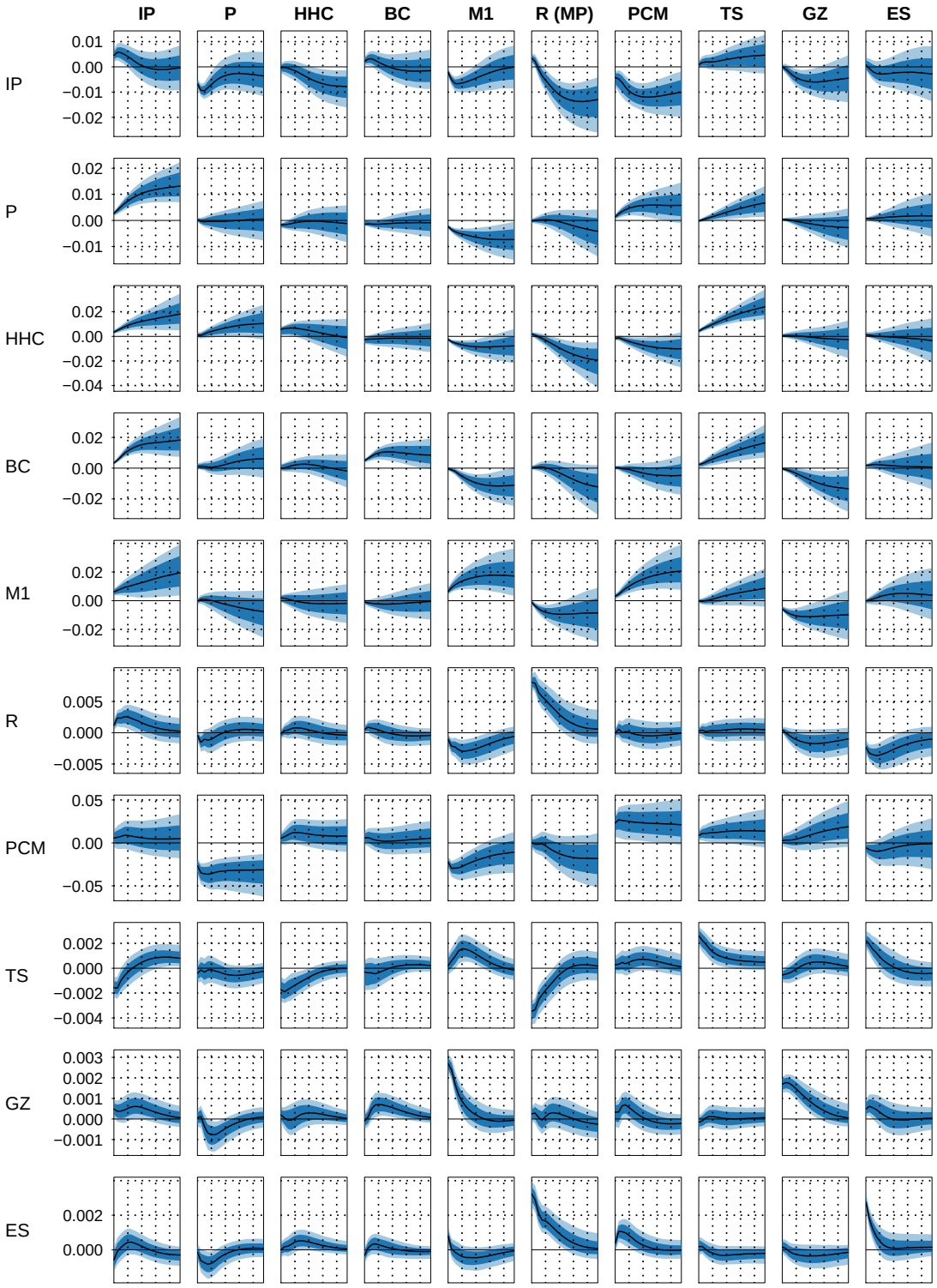

FIGURE 16. Impulse responses with quarterly-averaged data and Flow of Funds credit series. Heteroskedasticity with change dates as in the main monthly model. Normal errors assumption.

Shock	Q1 1973 – Q4 1979	Q1 1979 – Q4 1982	Q1 1983 – Q4 1989	Q1 1990 – Q4 2007	Q1 2008 – Q4 2010	Q1 2011 – Q1 2015
HHC	0.73	2.89	1.03	0.77	0.27	0.29
BC	0.19	2.16	1.30	0.90	1.07	0.36
R (MP)	1.81	3.13	0.46	0.06	0.51	0.01

TABLE 10. Selected relative variances in a model estimated with quarterly credit data from the Federal Reserve/BIS.

partially confound the monetary policy shock with the spread shock, we omit detailed discussion of the rest of the impulse responses.

Nonetheless it is worth noting that in Figure 16, where we have substituted the Flow of Funds versions of our credit aggregates for our own measures, there is evidence that the definition of the credit aggregates does matter. In the flow of funds results, the HHC shock moves the price level down by a statistically significant, but small, amount, and immediately lowers the term spread without much immediate effect on R — i.e. it lowers long rates. The shock is followed with a delay by negative output growth. With our data quarterly averaged, neither the HHC or BC shocks (which in the monthly data we have interpreted as "credit aggregate shocks") is followed by substantial negative output growth, and the simultaneous jumps in P and TS in response to the HHC shock that appear in the flow of funds data results are not present.

These results make clear that the Flow of Funds aggregates and the data we are using behave somewhat differently. It is possible, though, to reconcile these results with some more qualitative aspects of our main model's story. The credit shocks in the model estimated with Flow of Funds data explain only small fractions of credit and output variance. At the 20 quarter (5 year) horizon, this HHC shock explains only 2.7% of forecast error variance in HHC itself and only 7.6% of forecast error variance in IP. Moreover, both credit shocks (to HHC and BC) spike in relative variance during the variance regime of the extreme monetary policy (Q1 1980 to Q4 1982) rather than the Great Moderation, in build-up to the Great Recession (Table 10).

REFERENCES

Gelman, Andrew, and Donald B. Rubin. 1992. "Inference from Iterative Simluation Using Multiple Sequences." *Statistical Science*, 7(4): 457–511.

Gelman, Andrew, John B. Carlin, Hal S. Stern, and Donald B. Rubin. 2014. *Bayesian Data Analysis.* . 3 ed., Boca Raton, FL, USA:Chapman & Hall.

Geweke, John. 1999. "Using simulation methods for bayesian econometric models: inference, development,and communication." *Econometric Reviews*, 18(1): 1–73.

Hamilton, James D., Daniel F. Waggoner, and Tao Zha. 2007. "Normalization in Econometrics." *Econometric Reviews*, 26(2-4): 221–252.

Ramey, V.A. 2016. "Macroeconomic Shocks and Their Propagation." In . Vol. 2 of *Handbook of Macroeconomics*, , ed. John B. Taylor and Harald Uhlig, 71 – 162. Elsevier.

Romer, Christina D., and David H. Romer. 2004. "A New Measure of Monetary Shocks: Derivation and Implications." *American Economic Review*, 94(4): 1055–1084.

Sims, Christopher A., and Tao Zha. 1996. "Bayesian methods for dynamic multivariate models." Federal Reserve Bank of Atlanta FRB Atlanta Working Paper 96-13.