

FEEDBACKS: FINANCIAL MARKETS AND ECONOMIC ACTIVITY

MARKUS BRUNNERMEIER, DARIUS PALIA, KARTHIK A. SASTRY, AND CHRISTOPHER A. SIMS

ABSTRACT. We examine the relation among measures of credit expansion, measures of financial market stress, and standard macroeconomic aggregates. We use a structural model that is “lightly” identified — a form of structural VAR — and that uses monthly data on up to 10 variables. The model explains observed variation as driven by 10 mutually independent structural disturbances, one of which emerges as representing monetary policy. There is more than one financial stress shock, suggesting that attempts to create a one-dimensional index of financial stress may be misguided. In pseudo-out-of-sample forecasting tests, neither bond spreads, interbank spreads, nor credit aggregates had much predictive value far in advance of the 2008-9 downturn, though spreads (but not credit aggregates) were helpful in recognizing the downturn once it had begun. No strong pattern of credit expansion preceding output declines emerges. Some of these results are in apparent conflict with previous empirical work in this area, and we show that our model can explain the previous results.

I. INTRODUCTION

In the long run, credit aggregates tend to expand with GDP, and indeed expand faster than GDP, so that the ratio of credit to GDP is larger in rich countries and tends to grow over time. Figure 1 illustrates this with a plot of the log of real household credit against the log of real GDP, which (except for a couple of small loops) is a nearly straight line with slope well above one. In studies of economic development, the ratio of credit to GDP is sometimes used as a measure of “financial depth,” which is thought to contribute positively to economic growth.¹ On the other hand a number of recent studies, among them Mian, Sufi and Verner (2015), Schularick and Taylor (2012), and Jordà, Schularick and Taylor (2014), claim to have demonstrated a predictive relation between rapid growth of credit and future low GDP growth or higher likelihood of crisis.

Date: September 1, 2016.

This document is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. <http://creativecommons.org/licenses/by-nc-sa/3.0/>.

¹Some earlier studies on this topic include Shaw (1973), McKinnon (1973), and Goldsmith (1969).

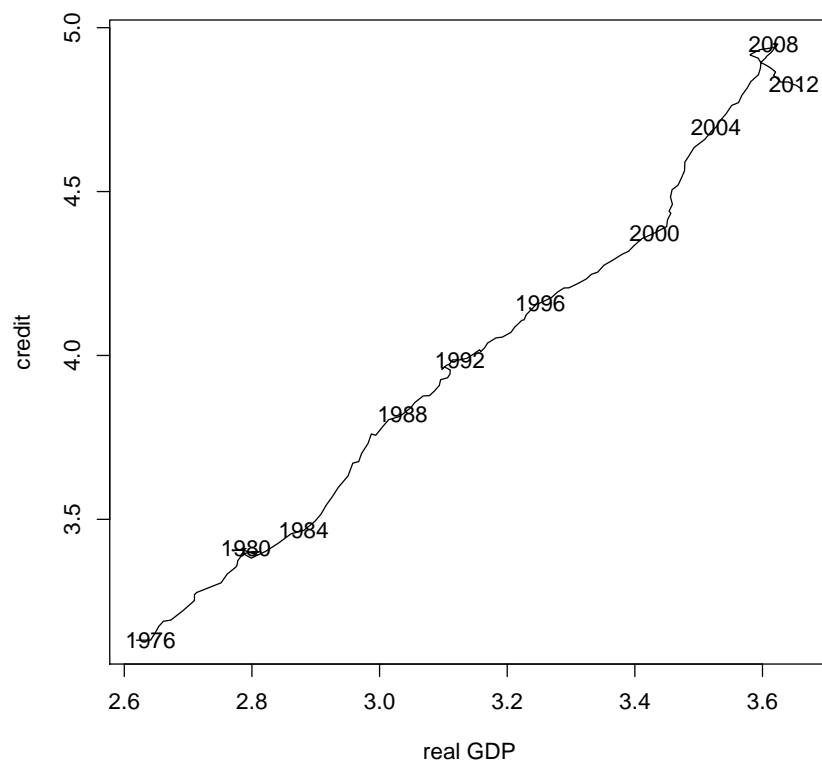


FIGURE 1. US real GDP and household credit (BIS quarterly aha series)

Monetary policy has strong effects on GDP growth and also, unsurprisingly, strong effects on credit growth and on spread variables that measure financial stress. Monetary policy also plausibly responds to rapid credit growth or contraction and to changes in spreads. To understand the policy implications of correlations or predictive regressions relating financial variables to GDP growth, it is essential that we understand the extent to which these correlations are generated by, or mediated by, monetary policy itself.

The structural VAR literature on monetary policy effects succeeded in separating two channels of relation between inflation and interest rates — policy-generated changes in interest rates tend to reduce inflation, while interest rates on average endogenously rise with inflation to compensate investors for inflation-generated losses. For the reasons we have listed here, it seems likely that there are multiple causal channels connecting

spreads, credit aggregates and business activity, and that some of these channels operate with opposite signs. It therefore seems appropriate to estimate a multiple equation model connecting these variables and to imitate if possible the structural VAR literature's approach to unraveling feedbacks in the data.

Most of the existing empirical literature in this area has used short lists of variables and has not attempted to separate monetary policy from other channels of connection between financial variables and the macroeconomy. Studies of the predictive power of credit growth have primarily used single-equation projection methods (e.g., Mian, Sufi and Verner (2015), Jordà, Schularick and Taylor (2014), and Jordà, Schularick and Taylor (2015)) or binary outcome (i.e., crisis or no crisis) predictive models (e.g., Schularick and Taylor (2012) and Drehmann and Juselius (2014)).² Studies focused on the information in credit spreads have looked extensively at single-equation models (e.g., Lopez-Salido, Stein and Zakrajsek (2015), and Krishnamurthy and Muir (2016)) and reduced form multi-equation models (Gilchrist, Yankov and Zakrajsek (2009) and Gilchrist and Zakrajsek (2012)). Gertler and Karadi (2015) and Caldara and Herbst (2016) have introduced credit spread variables into structurally identified, multiple-equation frameworks with monetary policy. But these authors do not focus on what including monetary policy does to the rest of the system (including possible "credit distress" channels) and ignore the role of credit aggregates.³

There have been other studies in this area based on fully interpreted structural dynamic stochastic general equilibrium models, which of course have included estimated effects

²Mian, Sufi and Verner (2015) is unique among these for using data outside of identified "crisis episodes." It also contains a small-scale multivariate example, with three variables (real GDP, household credit to GDP ratio, and business credit to GDP ratio), but does not endogenize interest rate dynamics or separately identify monetary policy.

³Krishnamurthy and Muir (2016) does look at both aggregates and spreads, but they look mainly at interactions of credit growth and credit spreads, instead of both as jointly endogenous parts of a system.

of monetary policy.⁴ These DSGE models, though, have not considered as many financial variables jointly as we consider here and have imposed more, and more arguable, identifying restrictions than we impose here.

Our model uses monthly data on industrial production (IP), the personal consumption expenditure deflator (P) household credit (HHC) business loan credit (BC), money supply (M1), the federal funds rate (R), a commodity price index (PCM), the 10 year over 3-month Treasury term spread (TS), the Gilchrist and Zakrajšek (2012) bond spread (GZ) and the 3-month Eurodollar over Treasury spread (ES).⁵ The sample period runs from January 1973 to June 2015.

We use the identification-through-heteroskedasticity approach pioneered in economics by Rigobon (2003). This approach assumes that the pattern by which disturbances feed through the economy is stable across time, but that the relative sizes of the independent sources of structural disturbance in the system varies across historical periods. We began modeling time variation in disturbance variance because it is so clearly needed to accurately describe financial variables and some macroeconomic variables. We discovered as we proceeded that we obtained stable, interpretable results from this assumption alone, without the need for the usual short or long run restrictions on dynamics usually applied to structural VAR's. Details of the model specification are in Section II below.

Here is a qualitative summary of our results. Details and quantitative results are in Section III below.

One of our 10 shocks, indeed the one with the most widespread effects across variables in the system, we interpret as a monetary policy shock. Responses to it match what is

⁴Christiano, Motto and Rostagno (2014), for instance, estimate a monetary DSGE model based on the contract enforcement friction of Bernanke, Gertler and Gilchrist (1999) and find that "risk shocks" which can be measured in observed credit spreads drive a significant portion of U.S. business cycle dynamics. The model uses data on credit spreads (BAA-AAA) and firm credit in addition to "standard" macro aggregates. Del Negro and Schorfheide (2013) provide a detailed comparison of the forecasting performance of this model, a standard Smets and Wouters (2007) DSGE model, and various reduced-form models.

⁵Details of the data and their sources are laid out in Section III below.

usually assumed about the responses to a monetary policy shock in SVAR models.⁶ This shock produces a sustained and non-trivial increase in the interbank spread (ES) variable.

There are shocks, distinct from the monetary policy shock, that move the GZ and ES spread variables, and then later move IP in the opposite direction. These fit the idea that disturbances that originate in financial markets can have macroeconomic effects.

Several shocks generate substantial movement in the credit variables HHC and BC, and all but one of them move IP in the same direction as the credit aggregates. This fits the idea that most movements in credit aggregates accompany expansion of activity and do not predict future slowdowns. There is a disturbance that moves HHC up, and then with a delay moves IP down. but the downward movement in IP is small and barely statistically significant. There may be periods where this shock is important, so that the credit expansion does predict future contraction in business activity, but a quantitative model that can identify such “bad” credit expansions and thereby allow a policy response would have to be multivariate, to separate this component of credit growth.

In Section IV we conduct pseudo-out-of-sample forecasting experiments to see what predictive value arises from including the spread and credit aggregate variables in the system. We find that the model gives little advance warning of the 2008-9 crash, whether or not the financial variables are included, but that the model tracks the course of the recession considerably better when they are included⁷. Most of the improved tracking of the crisis period comes from including the spread variables, not the credit aggregates.

The limited predictive value for credit aggregates in our system may appear difficult to reconcile with the results of previous studies with smaller models that have found substantial predictive value for credit aggregates in forecasting future business activity or future crises. In Section V we try to account for which aspects of the differences in

⁶Though we pick the shock we label “monetary policy” by looking at the sign and shape of its impulse responses, this is not the same as the frequently applied “sign restriction” approach to SVAR identification. Sign restriction identification does not lead, even asymptotically, to point identification of responses, whereas our approach, if its assumptions are correct, does provide point identification.

⁷This is similar to the conclusion of Del Negro and Schorfheide (2013), who compare a New Keynesian DSGE model with and without financial frictions of the form in Christiano, Motto and Rostagno (2014) and Bernanke, Gertler and Gilchrist (1999).

specification account for the apparent difference in results. We conclude that the results others have obtained with smaller models would not be unlikely if our estimated model represented the true structure of the economy.

II. MODELING FRAMEWORK

This section describes our empirical approach. The first two parts introduce the multivariate time series models we consider and the third part describes our Bayesian estimation method.

II.1. The Basic Model. We specify structural vector auto-regressive (SVAR) models with variances changing at exogenously specified dates. They can be described by the system of dynamic stochastic equations

$$A_0 y_t = \sum_{j=1}^p A_j y_{t-j} + C + \epsilon_t \quad (1)$$

where y_t is an $n \times 1$ vector of observed variables, A_0 is an $n \times n$ matrix which determines simultaneous relationships, the A_j are $n \times n$ matrices of coefficients at each lag j , C is an $n \times 1$ vector of constants, and ϵ_t is a vector of independent shocks. In the base model, these are Gaussian (normally distributed).

We exogenously separate the time span $\{1 \dots T\}$ into M subperiods and set

$$\mathbb{E} [\epsilon_t \epsilon_t'] = \Lambda_m \quad \text{if } t \text{ is in period } m \in \{1 \dots M\} \quad (2)$$

where Λ_m is a diagonal matrix. Thus the variance of the structural shocks changes across periods, but the dynamic relationship among the variables, as determined by A_0 and the A_j , remain fixed. In different terms, the impulse responses to structural shocks will have the same shape across variance periods, but their scales will vary. Our choice of variance regimes in estimated models (discussed in Section III and displayed in Table 2) is motivated by the variation of the time series and outside knowledge about policy changes.

We could fairly easily have allowed for regime changes to evolve as a Markov-switching stochastic process, as in Sims and Zha (2006). However, so long as the regimes are persistent, few in number, and well-determined by the data, inference about the model's dynamics is not likely to be strongly affected by conditioning on the regime switch dates as if known. Of course it is plausible that the variance regime switches are not only random, but endogenously determined. Allowing for that would greatly complicate the model and, since the regime switches are few in the data, might leave the nature of the regime switch endogeneity ill-determined by the data. We leave this to future research.

Our set-up can also be illustrated in the reduced form,

$$y_t = \sum_{j=1}^p B_j y_{t-j} + D + u_t \quad (3)$$

with

$$\mathbb{E} [u_t u_t'] = A_0^{-1} \Lambda_m (A_0^{-1})' \quad \text{if } t \text{ is in period } m \in \{1 \dots M\} \quad (4)$$

Some normalization is required, as we could multiply the rows of A_0 and Λ by scale factors without changing the implied behavior of the data. We impose the restriction

$$\frac{1}{m} \sum_{m=1}^M \lambda_{m,i} = 1 \quad \forall i \in \{1 \dots n\}$$

where $\lambda_{m,i}$ is the i th diagonal element of Λ_m . This makes the cross-period average structural variance 1 in each equation. It can be shown that, given such a normalization and the technical condition that each pair of equations differs in variance in at least one period, we can uniquely identify all n^2 parameters of A_0 (up to flipping the sign of an entire row, or permuting the order of rows).⁸ Thus the variance switching eliminates the need for short-term timing restrictions (i.e., zero elements in A_0) or linear restrictions on A_i for

⁸The intuition is that if Σ_j is the reduced form residual covariance matrix for period j , the expression

$$\Sigma_i^{-1} \Sigma_j = A_0' \Lambda_i^{-1} \Lambda_j (A_0^{-1})'$$

has the form of an eigenvalue decomposition, with the columns of A_0' the eigenvectors. As long as the eigenvalues, the diagonal elements of $\Lambda_i^{-1} \Lambda_j$, are unique (i.e., there is no k, l such that $\lambda_{j,k} / \lambda_{i,k} = \lambda_{j,l} / \lambda_{i,l}$), the rows of A_0 are therefore uniquely determined up to scale once we know Σ_i and Σ_j .

A more formal proof of this can be found, for instance, in Lanne, Lütkepohl and Maciejowska (2010).

$i > 0$ (as in identification through long run restrictions). The fully identified model can allow for every variable to respond to every structural shock in the initial period. *A priori*, this seems quite reasonable at monthly and lower frequencies.

While under the model's assumptions the impulse responses of the system will be consistently estimated in large enough samples, the model does not give names to the shocks that drive it. We pick out one of our 10 estimated shocks as a monetary policy shock, and two others as reflecting disturbances originating in financial markets. Our choices of names for these shocks reflect *a priori* assumptions about what shocks with these names should look like, in terms of the responses they generate. This is similar in spirit to the "sign restriction" approach to SVAR identification. However sign restrictions on impulse responses by themselves do not provide point identification, whereas our approach does allow point identification.

From a purely econometric standpoint, then, our model has two major advantages over any "plain" SVAR model without regime-switching variances. The procedure will be more efficient for the same reasons as standard generalized least squares (GLS): it prevents periods of large shocks from inefficiently dominating the likelihood. And it allows more modeling flexibility, since with it we avoid making arguable linear restrictions on the $(A_i)_{i \geq 0}$ matrices.

II.2. Further Correction for Extreme Events. The previous section's correction for heteroskedasticity will work best if volatilities mainly change at low frequencies. But it does not allow for the possibility of a few isolated large disturbances or outliers. For instance, the bankruptcy of Lehman Brothers in October 2008 and the 650 basis point drop in the Federal Funds rate from April to May 1980 generate outliers of around 6 standard deviations that do not disappear when we allow variance-regime switches. To guard against such large shocks distorting inference, we consider alternative specifications in which

structural errors ϵ_t have normal-mixture distributions, or Gaussian distributions with random variances.⁹

In the model notation, we can introduce random parameters $\tilde{\zeta}_{i,t}$ such that

$$\epsilon_{i,t} \sim \text{Normal}(0, \lambda_{i,t} \tilde{\zeta}_{i,t}) \quad (5)$$

We can also think of these objects as shocks which capture, in a simple way, a high-frequency component of volatility. This structure allows the data to decide whether to downweight certain residuals for “idiosyncratic” (i.e., independent of equation or time period) reasons.

Our goal is to generate a shock distribution with “fat tails” (i.e., a higher chance of high deviation events). This can be achieved by setting several possible distributions for the $\tilde{\zeta}_{i,t}$. The first case we consider is setting

$$\tilde{\zeta}_{i,t} \sim \text{Inverse Gamma}(\text{shape} = \alpha/2, \text{rate} = 2/\alpha). \quad (6)$$

This implies that each $\epsilon_{i,t}$ has an independent Student’s T distribution with α degrees of freedom and unit scale.

We also look at an alternative case with $\sqrt{\tilde{\zeta}_{i,t}}$ as independent k -multinomial with

$$\tilde{\zeta}_{i,t} = \beta_i \text{ with probability } \alpha_i \text{ for } i \in \{1 \dots k\} \quad (7)$$

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.

In both cases, we tune hyperparameters of these specifications to match the distribution of estimated residuals from the Gaussian-errors model. We chose parameters for each distribution based on residuals for the original model. The t -distributed errors model

⁹Models with mixture-of-normals disturbances have appeared previously in the empirical macroeconomics literature. Lanne and Lütkepohl (2010) introduce a maximum likelihood approach to estimating a discrete normal mixture SVAR model, and Chiu, Mumtaz and Pinter (2015) describe a Bayesian Gibbs sampling algorithm with an application to a model with stochastic volatility for U.S. data. Chib and Ramamurthy (2014) present a Gibbs sampling method for estimating a DSGE model with t -distributed shocks and Cúrdia, Del Negro and Greenwald (2014) find that the assumption improves the fit of a New Keynesian DSGE model that already includes low-frequency volatility changes.

has degrees of freedom 5.7 and unit scale. The multinomial normal mixture has possible standard deviations .68, 1.17, and 2.55 with prior probabilities .59, .39 and .02.¹⁰

II.3. Econometric Methodology. Equations (1) and (2), combined with the normalization of variances, describe a model with n^2 free parameters in A_0 , $(M - 1)n$ free parameters in the Λ_m , and n^2p free parameters in the A_j . Each of the “extreme events” variations has another nT parameters.¹¹ We use Bayesian methods to update beliefs about the parameters conditional on observed data $\{y_1 \dots y_T\}$ and initial conditions $\{y_{-p-1} \dots y_0\}$.

On A_0 we specify independent Gaussian priors on all elements, centered around 100 times the identity matrix, with standard deviation 200. For $\lambda_{.i} = \{\lambda_{1,i} \dots \lambda_{M,i}\}$, the vector of variances in each equation i , we put a Dirichlet prior (with $\alpha = 2$) on $\lambda_{.i}/M$. This has a compact support on $[0, 1]^M$ and centers belief around variances being equal in all periods. We use a variation of the “Minnesota prior” described in Sims and Zha (1996) on the reduced form parameters in the matrices B_j and D of Equation (3). These priors, described in more detail in Appendix 1, center belief loosely around independent random walks in each variable. They also imply that constant terms should not interact with near unit roots to imply rapid trend growth and that, if the dynamics are stationary, initial conditions should not be too far from the model’s implied unconditional means. Conditional on A_0 , these priors imply Gaussian priors on the A_j matrices.

In the basic case, without outlier correction, we sample from the posterior distribution for these parameters in a two-step process that exploits the fact that, conditional on knowing A_0 and $\{\Lambda_1 \dots \Lambda_M\}$, we can treat (1) as a system of n independent linear regressions which can be estimated with weighted least squares. In the first step of the sampling process, we use a Random Walk Metropolis algorithm to sample the elements of A_0 and $\{\Lambda_1 \dots \Lambda_M\}$ using the likelihood integrated over the A_j (which is available analytically).

¹⁰All that matters to the likelihood is the shape of these distributions, not their scale, since A_0 can absorb differences in scale. However because our prior on A_0 is not scale invariant, results might have been slightly different if we had used the fitted scale for the t -distributed shocks (.78).

¹¹As mentioned previously, the ξ_{it} can equally well be called parameters or shocks. The fact that there are so many of them does not mean they cause difficulties in estimation, because each has a specified distribution. This is a special case of the general point that Bayesian inference treats parameters and shocks symmetrically. They are all unknown objects with distributions.

Then, after drawing a large MCMC sample from the marginal posterior distribution of A_0 and $\{\Lambda_1 \dots \Lambda_M\}$ we can, for each of the A_0, Λ draws, sample from the coefficients in the A_j which have a known conditional normal distribution. Appendix 3 reports more details about the convergence of the Monte Carlo portion of the method in multiple simulation chains started from dispersed points in the parameter space.

The outlier correction requires adding a Gibbs sampling step to the algorithm. The first step, a Monte Carlo update of A_0 and $\{\Lambda_1 \dots \Lambda_M\}$ conditional on the $\xi_{i,t}$, has the same form as the first step in the Gaussian model. But now we need to draw A_j for $j > 0$ after each draw of A_0, Λ , so that we can form the implied normalized residuals $u_{i,t}$. The $\xi_{i,t}$, conditional on the $u_{i,t}$ are distributed independently across i and t , allowing us to draw directly from their exact conditional posterior distribution.

For a given set of model parameters, we could change the sign of the coefficients in an equation (a row of (A_0, A_1, \dots)) or change the order of the equations (permute the rows of (A_0, A_1, \dots)), without changing the implied distribution of the data. The likelihood maximum therefore recurs through the parameter space at every permutation or sign change of the parameters. This means that a complete MCMC sampling of the posterior distribution would show identical impulse response distributions for all shocks, all centered at zero response — but only if the prior itself were invariant to permutations of the orderings or signs of the equations.

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. Nonetheless these methods, if applied on data where identification did not emerge as strongly, might need to test for and eliminate permuted or sign-switched models.¹²

This probability model implies prior and posterior distributions for all (potentially non-linear) transformations of the coefficients, including the reduced form coefficients B_j and

¹²This is a special case of the kind of normalization issue discussed by Hamilton, Waggoner and Zha (2007).

Abbv.	Description
IP	Industrial production
P	Personal consumption expenditures price index
HHC	Sum of commercial bank real estate and consumer loans
BC	Commercial bank commercial & industrial loans
M1	M1 money supply
R	Federal funds rate
PCM	CRB/BLS spot (commodity) price index
TS	Term spread of 10 year over 3 month Treasuries
GZ	Gilchrist and Zakrajšek (2012) bond spread
ES	“TED spread” of 3-month Eurodollars over 3 month Treasuries

TABLE 1. Data series used in model estimation.

the impulse response functions for variable i to each shock j . In all reported results, following Sims and Zha (1999), we report horizon-by-horizon 68% and 90% posterior density regions as “error bands.”

III. DATA AND RESULTS

Our main specification uses monthly data on 10 time series (listed in Table 1) from January 1973 to June 2015. We include data through the 1970s because they are a valuable source of variation in the time series and because correction for time-varying variance can account for what otherwise might be interpreted as regime change in monetary policy (e.g., as in Sims and Zha (2006)).

Our measures of “household” and “business” credit are based on the Federal Reserve’s weekly surveys of U.S. commercial banks.¹³ These data are different from the quarterly and annual series, based on a more comprehensive survey of lenders and categorized based on the borrower type (including “households and non-profits,” “nonfinancial noncorporate business,” and “nonfinancial corporate business”), used in some other research.¹⁴ Appendix F includes a more detailed discussion of the differences. Although our “household credit” series includes commercial real estate loans (which cannot be separately identified for the entire sample in the data) and our “business credit” data seems

¹³These are published in the H.8 “Assets and Liabilities of Commercial Banks in the United States” release.

¹⁴In particular, the cross-country database, assembled by the BIS (and described in Dembiermont, Drehmann and Muksakunratana (2013)) uses these quarterly data

	Start	End	Description
1	Jan 1973	Sep 1979	Oil crisis, stagflation, and Burns Fed
2	Oct 1979	Dec 1982	Volcker disinflation
3	Jan 1983	Dec 1989	Recovery from Reagan recession
4	Jan 1990	Dec 2007	Great Moderation, Greenspan Fed
5	Jan 2008	Dec 2010	Great Recession
6	Jan 2011	Jun 2015	Zero Lower Bound, Recovery from Great Recession

TABLE 2. Dates for variance regimes in full model specification.

to have more high-frequency variation than the corresponding quarterly series, we believe these data capture the majority of the low-frequency behaviors that are critical for existing empirical evidence of their forecasting power.¹⁵

The inclusion of three credit spreads (of interest rates over short-term Treasuries) is meant to capture several possible dimensions of credit market stress: the term spread captures uncertainty about future movements in fundamentals, the bond spread captures tightness in business financing, and the TED spread captures tightness in bank financing. The first was also expected to, along with the Federal Funds rate, M1, and commodity prices, provide a sharper identification of a monetary policy shock.

We separate the full sample into six variance regimes described in Table 2. From the standpoint of estimation efficiency, we expect the separate treatment of the Volcker disinflation and Great Recession to discourage overfitting of high monetary policy and financial stress variations respectively by allowing the model to “down-weight” these periods’ residuals.

Table 4 displays estimates of the marginal data densities of the base model, the model with a mixture of three normals for structural shocks, and the model with Student’s t distributed shocks. These are reported in a log points scale, so a difference of over 100 is extremely strong evidence (i.e., more than an e^{100} odds ratio) in favor of the better model.

¹⁵One practical complication is dealing with breaks in the credit series introduced by changes in accounting standards and the transfer of assets among classes of bank (e.g., the conversion of a thrift bank to a commercial bank). Our specific calculations for eliminating these breaks, which are particularly large in the real estate credit series, are detailed in the Appendix F.

	Jan 1973 – Sep 1979	Oct 1979 – Dec 1982	Jan 1983 – Dec 1989	Jan 1990 – Dec 2007	Jan 2008 – Dec 2010	Jan 2011 – Jun 2015
1	1.404	1.335	0.694	0.553	1.563	0.450
2	0.675	1.242	1.176	1.161	0.892	0.854
3	0.332	0.334	0.273	1.193	3.293	0.575
4	1.408	0.557	1.176	1.092	1.126	0.641
5	0.207	0.534	0.378	0.820	2.361	1.701
6	0.568	4.850	0.400	0.062	0.118	0.002
7	1.361	0.802	0.698	0.745	1.811	0.583
8	0.969	2.547	0.762	0.414	0.980	0.328
9	0.631	0.334	0.315	0.595	3.717	0.408
10	1.541	1.790	0.460	0.209	1.997	0.004

TABLE 3. Relative variances for each of ten shocks in six periods. Point estimates are posterior medians and numbers in parentheses are 90% bands of posterior draws.

Model	Marginal data density
No heteroskedasticity, Gaussian errors	—
Heteroskedasticity, Gaussian errors	51530.5
Heteroskedasticity, 3-Gaussian mix errors	52281.4
Heteroskedasticity, t errors	52293.2

TABLE 4. Marginal data densities for four VAR models for our 10 time series. (The first row is not yet computed, and the other three come from Monte Carlo runs that are probably not yet fully converged. The next draft of the paper will have more complete results.)

The improvement in fit from the normal mixture or t model is large, and those two models have similar fit.¹⁶

Figure 2 shows the impulse response over five years of all 10 variables to the model's orthogonal structural shocks, scaled to unit draws from a standard t distribution with 5.7 degrees of freedom. Since the diagonal of Λ_i is normalized to sum to one across regimes, these responses are a kind of average across regimes. The model, despite the lack of any identifying restrictions, fits identifiable monetary policy (number 6) and credit spread (9-10) shocks with significant long-term real consequences. Shock 6 is the only one that has

¹⁶The numbers shown in the table are preliminary, as we are not sure the Monte Carlo calculations used to compute them are fully converged, and we do not yet have the value for the model without heteroskedasticity. The improvement in fit from the fat tailed residual distributions seems likely to be robust, however.

an immediate positive R response, a delayed negative IP response, a negative (though ill-determined) long run P response, negative responses to M1 and the two credit aggregates, and a negative response of the term spread (as would be expected if the shock raises current interest rates and lowers expectations of future inflation).

The ninth and tenth shocks are the most important sources of variation in the GZ spread and the ES spread, respectively. The two spreads do not tend to move together in response to these shocks, and two have different patterns of effects on other variables. Both depress IP. Both depress P, though in the case of shock 10 this effect is statistically weak. Shock 9, which immediately impacts the GZ spread, has a strong delayed effect in depressing BC, but modest and indeterminate-signed effect on HHC, while shock 10, which immediately impacts ES, strongly depresses HHC with ill-determined effect on BC. Shock 10 produces an expansionary movement in R, while shock 9 does not. These patterns seem to fit an interpretation that distinguishes a banking credit shock (10) from a non-bank financial disturbance (shock 9). All the effects of these shocks on other variables are delayed, while their effects on the spread variables are immediate. This all fits an interpretation that they reflect disturbances originating in financial markets, with monetary policy at most (with shock 10) trying to partially offset their effects.

The third shock, which starts with an impulse to household credit (net of inflation) and leads to a persistent long-term decline in output, seems to match the “excessive credit growth” story demonstrated empirically by Mian, Sufi and Verner (2015), Schularick and Taylor (2012), and others. There are two major caveats, however. First, the output response with 90% confidence bands is barely significantly less than zero at the five-year or smaller horizon (and the 95% band would not be significant). Second, this shock is one of many in the system, and most of the others would create positive covariances in output and credit over short and long horizons.

The seventh shock accounts for a substantial component of variation in IP and P. Its immediate effect is to increase commodity prices, and to some degree to increase the GZ spread. The effect on commodity prices is persistent. With some delay, P (the PCE

deflator) moves up and IP moves down. Neither BC nor HHC moves much. This looks like a “supply shock” originating in commodity markets.

These core impulse response results seem robust to the alternative error specifications. Median impulse responses calculated from the Gaussian model are nearly the same as those calculated from the t or mixed-normal model when put on a common scale. The main difference is that most of the error bands are somewhat narrower for the models based on normal mixture residuals. This is what would be expected if the normal mixture models were correct. The impulse responses from the misspecified Gaussian model in that case would be consistent, but somewhat inefficient estimates. Figure 3 compares the median impulse responses for IP, R, GZ and ES to shocks 6 (monetary policy), 9 (GZ) and 10 (ES), with each shock scaled so that its largest initial component (R, GZ and ES, respectively) is the same size for both the t and Gaussian models. It is clear that the differences between them are well within the error bands. On the other hand, the Gaussian model does just barely include zero within the entire 90% error band around the response of IP to shock 3, which we have interpreted as possibly corresponding to an “unhealthy” expansion of household credit. Full impulse responses from the Gaussian errors model, and a model with multinomial normal mixture errors, can be found in Appendix E.

The picture is also very similar if we estimate with data only up to December 2007 (Figure 4). In particular, the identification of monetary policy and spread effects is very stable. There is additional (weak) evidence that a shock which induces business credit expansion (ordered fourth) has a negative long-term output effect, of a similar magnitude as the effect of the third (household credit) shock.

Our results suggest that the variances of these shocks change substantially among periods. Table 3 reports the variances of each of the ten structural shocks in the posterior mode Gaussian errors model, over the full sample. 90% probability bands for these relative variances are quite tight, mostly within 0.8 to 1.2 times the posterior median estimate. In general, there is strong evidence of time-varying variance. Several of the shocks spike in variance during the financial crisis (period 5). The sixth shock, which we identify as

a monetary policy shock, has a considerably inflated variance in the Volcker disinflation period and almost zero variance in the most recent period (near the Zero Lower Bound).

While the variance differences among regimes are large, the model still needs the normal mixture assumption to justify large residuals. Table 5 lists the 10 largest posterior median $\zeta_{i,t}$, or “variance adjustments” for individual equations and time periods. These are in standard deviation units and are still associated with large residuals — so the equivalent residuals in the Gaussian error models are all larger than 4, with tail probabilities lower than 0.0006!¹⁷ The biggest of these shocks were those in the fifth disturbance at and just after September 11, 2001. This shock primarily affects M1, with small effects on other variables, and these two shocks reflect an accommodated brief increase in demand for liquidity. The next largest is a surprise easing of monetary policy in June 1980, during the recession of that year. Interest rates had already been brought down sharply by the Fed, and the model is surprised at a further decrease. October 1987 was the time of a stock market crash, that shows up in the model as a large upward shock to the interbank spread.

The remainder of these “top 10” shocks occurred during 2007-2008. One, in September 2008, was to the first disturbance, which weights heavily on industrial production. One, at the time of the Bear-Stearns crisis (March 2009), was a surprise monetary policy easing. The remaining four were all upward shocks to disturbance 10, which we have interpreted as an interbank spread shock or, in one case, disturbance 9, the corporate spread shock. These large shocks have reasonable economic interpretations, and the model avoids having several supposedly independent shocks coincidentally extremely large at the same date. The surprising sequence of upward shocks in spreads, even when higher volatility in this period has been allowed for, fits the descriptions by policy-makers of their uncertainties during 2007-2008. Yet the shocks are not enough for the model to predict a sharp downturn very far in advance.

¹⁷The Gaussian errors model, at the posterior mode, has 42 residuals larger than 3 in absolute value, 14 greater than 4, and 4 greater than 5. For independent standard normal draws, the expected values should be 13.5, 0.3, and 0.002 respectively.

Date	Shock	ζ Value	Diagonal impact
11/2001	5	4.82	-0.0128
10/2001	5	4.27	0.0133
6/1980	6	3.71	-0.0224
11/2008	9	3.39	0.0089
9/2007	10	3.35	0.0025
11/1987	10	3.18	0.0033
9/2008	11	2.98	-0.0147
1/2008	10	2.98	0.0023
12/2007	10	2.93	0.0018
3/2008	6	2.90	-0.0028

TABLE 5. Largest $\zeta_{i,t}$ values in t -errors model, with dates and “diagonal” impacts. The latter are the effects of the shocks on the variable with the shock’s number. 5 is M1, 6, is Fed Funds, 9 is corporate spread, 10 is interbank spread, and 11 is industrial production. Units are logs for M1 and industrial production, raw annual rates (not per cent) for interest rate and spread variables.

These largest shocks do concentrate on just a few of the structural disturbances and, despite our attempt to allow for extended periods of high and low variance, the extreme shocks concentrate in 2007-2008. This suggests we might in future work modify the dates of our regimes or try estimating the regime switches using a Markov-switching extension of the model.

The remainder of this section reviews the model dynamics of key shocks in greater detail.

III.1. The Credit Channel of Monetary Policy. The sixth ordered shock of Figure 2 satisfies the description of a monetary policy shock in the initial impulse to the Federal Funds rate, initial decrease in the 10 year over 3 month Treasury term spread, and persistent negative impact on output. At the 68% level, there is still substantial uncertainty about the responses of both consumer and commodity prices, but point estimates from the posterior mode model show persistent declines. This identification comes despite the lack of any identifying restrictions on contemporaneous responses (or the monetary policy reaction function).¹⁸

¹⁸It is interesting that the uncertainty about the sign of the responses of P and PCM to this shock arises only from the data after 2007, as can be seen by comparing Figures 2 and 4. Apparently the positive interest

The model provides substantial evidence that the effect of short rate movements is amplified by corresponding movements in interest rate spreads. In particular, the 3-month Eurodollar spread over Treasuries increases about 1 basis point per 3 in the Federal Funds rate and decays over a similar horizon. There is very scant evidence, in contrast, that the bond premium moves in the short run (i.e., within the first 6 months) and no evidence that the effect persists longer. The general finding of financial “amplification” of monetary policy shocks is consistent with the empirical results of Gertler and Karadi (2015) and a variety of theoretical models which suggest that *risk premia* should move in response to monetary policy (e.g., Drechsler, Savov and Schnabl (2016) and Brunnermeier and Sanikov (2016)). Our empirical result is focused, however, on *inter-bank* credit conditions separate from *firm-level* credit conditions. This is concurrent with our broader empirical point that the information content of credit spreads in the multi-variate model is *multi-dimensional*—movements in different spreads forecast different macro-financial dynamics and, potentially, relate to different economic mechanisms.

The model’s opposite implications for household and business lending, the former of which sharply declines over all horizons and the latter of which increases slightly over the 18 months, potentially speaks to the differential access of large and small borrowers to credit at the onset of a recession. The theory and empirical result in monetary VARs goes back to Gertler and Gilchrist (1993), and a version of it is corroborated in bank balance sheet data in the recent financial crisis by Ivashina and Scharfstein (2010).

III.2. Spread Spikes and Early Warning. Three independent shocks, ordered eight to ten, can be identified by sharp increases in spreads at $t = 0$. The eighth shock hits the term spread only and has no long-term consequences for output growth. In contrast, the tenth shock, which hits both the term spread and the interbank spread, results in a sharp recession and contraction in household credit. The ninth shock, which begins with a change in the corporate bond premium and small associated movements in ES or TS,

rates in 2008-9, when the model expected to see them go negative, appeared to the model as a contractionary monetary policy shock with surprisingly little negative effect on prices

is followed by a slightly longer recession and a major contraction in business rather than household credit.

The fitting of two independent stress shocks suggests the importance of a multidimensional approach to measuring financial stress. In their long-run macro implications, the two shocks can be distinguished by sharply different implications for credit aggregates and prices. The “bond spread shock” is associated with a persistent reduction in the consumer and commodity price level and a significant decline in loans to businesses. The “inter-bank shock,” in contrast, has no long-term price effect and predicts a significant long-term contraction in household, not business, credit. It also seems to pick up some delayed monetary loosening, potentially in response to the generated recession.

In the historical record, inter-bank shocks have almost as high a variance in the early sample (1973 to 1982) as they do in the financial crisis (Table 3). With post-2009 interbank rates very close to short rates at zero, this channel almost completely shuts down in the final variance period. The corporate bond spread shock, in contrast, is by some margin highest variance during the 2008 financial crisis.

Taken together, the impulse response and the estimated variances suggest that the macro importance of spread shocks—closely related to the forecasting value of the spread variables—is concentrated during certain high variance episodes and largest at *short* horizons. Both spread spikes only precede the trough of the output effect by six months to a year. This is enough “early warning” to react more quickly at the *onset* of a recession, but likely not enough to steer an economy around the risk completely through policy intervention.

III.3. Credit Growth and Recessions. Our model offers some support, within 90% error bands, of the hypothesis that excessive growth in household credit can forecast negative long-term real output growth. The shape of our estimated output response to shock three, with a short-term output boost and long-term contraction, matches that of the small-system (household credit to GDP, business credit to GDP, and real GDP) VAR of Mian, Sufi and Verner (2015). Because our system has many more variables, we can be more

confident that the result does not pick up the effects of financial stress, monetary tightening, or inflation. Furthermore, like the *single-equation* results of Mian, Sufi and Verner (2015), our results suggest that movements in household credit are substantially more predictive than those in business credit. More specifically, the jump of the former independent of the latter seems to be the “signal” for this particular shock.

That said, our results still raise questions about the importance of this phenomenon and the wisdom of encouraging policy to actively “lean against it.” First, the estimated magnitude of this credit to real output channel is still relatively low, with a one per cent expansion of household credit over about 2.5 years followed by an output decline reaching 25 basis points below trend after five years. Since the structural shocks are t distributed with 5.7 degrees of freedom, this would need to be scaled up by 1.24 to be a one-standard-deviation shock, or by 1.1 to be at the same quantile (.84) of its distribution as a one-standard-deviation normal shock. In the period before the 2008 crisis, our model has this shock scaled up by a factor 1.18. But even rescaled by $1.18 \cdot 1.24 = 1.46$, this shock’s predicted output decline is considerably less than the 2.1 percentage point GDP response to a one-standard deviation event (a credit-to-GDP shock) reported by Mian, Sufi and Verner (2015). It is comparable to, but still smaller than, the response observed in the same authors’ 3-variable VAR system (close to 0.7 percentage points over 10 quarters).

One way of quantifying the importance of this credit shock relative to others in the model is to calculate forecast error variance decompositions (Figure 5).¹⁹ The importance of third shock for explaining credit variation starts very high (as it has by far the largest contemporaneous impact on credit) and decays over time. Over the “long-run” of five years, this shock explains only about 20 percent of household credit variation, while other shocks with “passive” credit reactions moving in the same direction as output explain the

¹⁹These are the squared impulse responses scaled to sum to one for each response variable in each period. Precisely, the variance decomposition of variable i is, for each j and each time horizon s , the proportion of s -step ahead forecast error variance in variable i attributable to shock j .

remainder. At this same horizon, the explanatory percentage for output is, at the posterior mode estimate, is only about 10%. This story mostly holds in all variance periods, including the two in which the credit shock is the largest (1/1990-12/2007 and 1/2008-12/2010). In these periods the shock explains a greater proportion of household credit variation (44% and 56% respectively), but close to the same amount of output variation (10% and 8%).

IV. CREDIT CONDITIONS AND FORECASTING

So far we have demonstrated that credit variables have an interesting interpretation within the model. But are they practically helpful to include, and could this have been realized before the 2008 financial crisis? We find that information in spreads can be useful for short-term forecasting at the onset of a crisis. The model with spreads does not, however, provide much advanced warning of a crisis or any clear advantage in “normal” times outside of recessions.

IV.1. Forecasting in the Recent Financial Crisis. We first focus on the 2007-08 financial crisis and its immediate aftermath. At each month between January 2007 and December 2010, we estimate (posterior modes of) models with and without credit variables using data only up to that point and then calculate 12-month forecasts. This “pseudo-out-of-sample forecasting” exercise offers a dimension in which to compare models with different data lists and gives a sense of how much changing the emphasized data in macro models would have helped in real time. We focus on the Gaussian errors specification, despite its fitting more poorly than the t model, because it seems to capture the main model dynamics and is much easier to do recursive computations with.

Figures 6, 7, and 8 plot posterior mode forecasts from our (Gaussian error) model with 10 variables, a version without the credit aggregates, and a version without the spreads, respectively, at 3-month intervals from January 2007 to October 2010. The model without spreads (Figure 8) never fully “accepts” the crisis, predicting a return to near pre-crisis growth rates at each point during the deepest contraction. The models with spreads (with

or without credit aggregates) give slightly less optimistic forecasts in early 2008, at which point the bond and inter-bank spreads have elevated slightly over mid-2000s levels. But the most obvious improvement is the models' ability to grasp the severity of the crisis during the deepest fall from mid 2008 to mid 2009. This observation is consistent with the previous section's analysis of impulse responses, which suggested that the model could identify spread shocks which have macro effects within the first few months. The spreads provide little advance warning of severe recession but do enhance recognition of the severe recession, and its likely persistence, once it is underway.

The addition of credit aggregates seems considerably less important. With or without credit aggregates, the model is quicker to recognize a persistent downturn. While forecasts of IP are little affected by excluding credit aggregates the model without them consistently predicts that interest rates will start reverting to positive values from the zero lower bound, though this seems to have limited effects on forecasted output or consumer prices.

IV.2. Forecasting Power in the Entire Sample. We generalize the exercise of the previous section by calculating forecasts with versions of the main model, the no spreads model, and the no credit model estimated up to each month from October 1979 to June 2015.²⁰ We focus on root mean squared error (RMSE) for forecasts of all variables common to the models.

Figures 9 and 10 display the evolution of these RMSE for the base model with all variables (blue), a model without spreads (red), and a model without credit aggregates (green). As suspected from the previous section, the models with spreads does a significantly better job predicting output just before and during the 2007-2009 financial crisis and recession. The model's internal projections for the Federal Funds rate are quite a bit

²⁰The truncation at the beginning of the sample comes from the requirement of having two variance regimes to identify the parameters. Unfortunately, this cuts out some interesting macroeconomic turbulence in the 1970s.

Additionally, for the period October 1979 to December 1982, we use models with six lags because of the smaller availability of data.

better *at* the zero lower bound, though this comes at the cost of one set of very poor forecasts right around the final major reduction in the rate in late 2008. Any advantages in forecasting the price level and credit aggregates in the crisis are less obvious.

Outside the recent financial crisis, and potentially the early 1980s and early 2000s recessions, the no spread model seems to perform just as well if not better. We might suspect that a formal or informal comparison of models before 2008 would not clearly support the inclusion of the financial variables, even if the estimated dynamics from such a model look like they have “economically interesting” transmissions from spreads to macro variables. Alternatively, the model with spreads might only be better for forecasting when economic conditions worsen.²¹

The model without credit aggregates, but with credit spreads, seems to match the full model quite closely throughout the sample. One exception seems to be the early part of the 1981-1982 recession and the subsequent uptick in growth around 1984. In several periods, including post-recession growth in the early 90s and 2010s, the no credit model is significantly *better* at predicting output. In general there is no clear pattern of the model with credit aggregates, after including spreads, doing a better job of forecasting the timing or severity of U.S. recessions.

V. WHY DO CREDIT AGGREGATES HAVE LOW PREDICTIVE VALUE?

Our results de-emphasizing the negative growth consequences of credit growth are at odds with a recent literature that empirically claims to show that credit expansion predicts negative growth and/or financial crises. The goal of this section is to demonstrate how our findings are compatible with others from smaller models. We first show that our main result is not affected by the frequency or transformation of our credit variables. We then discuss how both specification of a larger model and the assumption of heteroskedasticity

²¹These nuances could be captured formally by taking posterior forecasts averaged across an “ensemble” of models, the weights on which change over time (for instance, with some approximation of posterior odds). To capture them within the model might require some more complex (and possibly endogenous) modeling of regime switching.

combine to reduce the prominence of apparently negative consequences of credit growth in simpler models.

V.1. Data Transformation. We estimate our main specifications at the monthly frequency in order to exploit the higher frequency variation in credit spreads and more precisely identify the horizons at which “financial stress shocks” propagate to the macroeconomy. The majority of empirical studies showing high credit growth to predict long-term output declines have used lower frequency data, often with other transformations applied. We find that several alterations of our analysis along these lines do not change our main results.

V.1.1. Frequency. To check whether the level of time aggregation explains why our results show at most small negative effects on output of credit expansion, we estimate the model with quarterly averages of the same data. We also try replacing the averages of our monthly credit data with the quarterly credit data made available by the U.S. Federal Reserve and Bank of International Settlements.²²

Models estimated with quarterly data, using credit series from either source and an SVAR with time varying variances like our monthly model, give results quite similar to our monthly model. In particular, they show only one shock in which a credit variable moves promptly and IP moves with a delay in the opposite direction. As in our monthly model, this shock shows up as the third shock. With our monthly credit data simply aggregated to quarterly, the negative response of output to this shock is very similar in magnitude to what we found with the monthly data (left panel of Figure 11).

When we use the BIS quarterly credit series, the situation is slightly more complicated. Of the ten shocks, two are candidates for “dangerous credit growth.” The first involves substantial nominal, but not real, credit growth. This suggests the credit growth is a response to inflation, and we therefore think it more reasonable to interpret it as a supply shock than as reflecting independent excessive credit growth. The second does have a larger output response than the equivalent in the monthly data model, though the implied

²²We use the “credit to households” and “credit to non-financial corporations.”

timing (with the majority of the output and credit movement within the first quarter) is very different than what the explosive growth hypothesis predicts. This shock still explains only 6.6 percent of output variation and 5.0 percent of household credit variation at the five-year horizon with “average” scaling.²³ In the “Great Moderation” period of 1990:I to 2007:IV, these increase to 16.0 and 11.0% respectively, but not because of increases in the size of these shocks. The other structural shocks had low variance in this period, both in the quarterly model with BIS data and in our monthly model, shrinking the denominator of “percentage of variance explained”. While with the BIS data give stronger evidence for a “dangerous credit growth” hypothesis than our monthly frequency results, we still do not think it constitutes the type of result reported in the single-equation literature.²⁴

V.1.2. *Long-term Growth Rates.* Another possible issue is that we use only 10 months of information. A p lag system can fit dynamics with any linear combination of the variables including an $m \leq p$ month difference, but would not capture even lower frequency growth rate effects. 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 12 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 variation in the growth rate (49% at 5 years), but still only a modest amount (9%) 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.²⁵ The differences model fits considerably worse than the levels model, with an estimated marginal likelihood of 51244.7 (compared to 51530.5).

²³At the posterior median of draws.

²⁴These models use a slightly different Minnesota prior than the monthly ones, with tightness 5 and decay 1.

²⁵The Jacobian of any linear transformation like this is one. But now, if the model is with three lags, we are implicitly conditioning on the values of $\{c_0 - c_{-36}, c_1 - c_{-37}, \dots\}$, where c_t is the credit variable. This is different but not higher dimensional information.

V.1.3. *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 ones should have the opposite effect.²⁶

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.²⁷ In all optimization exercises to find the posterior mode of such a model, the data favored models without any nonlinear transformation.²⁸

V.2. Model Size and Underspecification. In smaller models estimated with the same methods, it is possible to recover results suggesting that impulses to the credit aggregates lead to negative output growth. Figure 13 presents the impulse response of a model (estimated with heteroskedasticity, like the base model) with four variables: real output, the

²⁶Jordà, Schularick and Taylor (2013), for instance, focus on upper n percentile credit growth events.

²⁷In our trials so far, we have 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

²⁸In this specification, $\beta_1 = 1, \alpha_1 = \alpha_3 = 0, \alpha_2 = 1$.

price level, nominal household credit, and nominal business credit. This is the smallest model that accounts for “standard” linear combinations of these variables (like credit to output ratios or real credit levels). 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 73% of business loan variation and 10% of output variation in one-step ahead forecasts, and 19% and 25% respectively in four-year-ahead forecasts. One might conclude from estimating this small system that business loans are a strong driving force in the macroeconomy and predictor of future economic activity over long horizons, though in the large system this quite clearly is not the case.

We can conclude that the same data will imply, when analyzed with our 10-variable model identified through heteroskedasticity, that credit expansion is rarely and weakly associated with low future IP growth, and, when analyzed with small models omitting spread and interest rate variables, that credit expansion reliably predicts low future IP growth. This could have two interpretations. One is that the big model has become so complicated that it is missing an important and powerful regularity in the data. The other is that the small models are too simple to give a clear picture. The first interpretation implies that the large model is incorrectly specified.

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. If the large model were misspecified, the apparent predictive power of credit expansion for future low IP growth should not be reproduced with the simulated data. We find that in fact it is reproduced.

In tests with 5000 simulated samples from the 10-variable estimated model with Gaussian innovations, we find that in a four-variable VAR (IP, PC, HHC, BC) we get a negative response of IP to a one-standard-deviation credit shock of one percent or larger at the three-year horizon with the probabilities shown in Table V.2. These are done with data at different frequencies, aggregated by taking averages of the monthly data our model

	BC Shock	HHC Shock
Monthly	.271	.001
Quarterly	.680	.018
Yearly	.670	.358

TABLE 6. Probability, in simulated data, of a -0.5 per cent 3-year response of IP to a positive credit shock in a 4-variable VAR

	Lagged IP?	
	No	Yes
HHC	.033	.128
BC	.553	.603

TABLE 7. Probability of finding a coefficient of -.2 or larger in a regression of 3-year IP growth on lagged 3-year credit growth

generates. By this metric, our own observations of approximately a 0.5 percentage point negative response in output to a business loans shock, in monthly data, would not be so uncommon. As the frequency is reduced, seeing similar results with a household credit shock is also quite possible.

When we imitate the “projection method” regressions of Mian, Sufi and Verner (2015) (with and without lagged output growth),²⁹ we find the probability of a coefficient of -.2 on lagged three-year credit growth in predicting three-year IP growth as in Table V.2. Again, these probabilities are reasonably high, suggesting that this one-equation model could suggest strong predictive power when the fully specified structural model attributes a limited causal role to credit innovations.

V.3. The Importance of Identification through Heteroskedasticity. So far we have tried to argue that model size does matter and that data selection and treatment do not matter in obtaining our core conclusions about the role of credit spreads and aggregates in

²⁹We imitate these exactly by using three-year forward differences on real output (IP) as the dependent variable, three-year forward differences in real credit (credit over price level) over output as the main independent variable, and lags of first differences of output as an extra independent variable. Everything is aggregated to the annual frequency.

linear models. We have not yet fully explored the importance of our assumption of heteroskedasticity and corresponding identification of contemporaneous relations.

As a first exercise, Figure 14 shows the impulse response of a “full-sized” ten variable constant-variance model identified through triangular (Cholesky) restrictions and with no allowance for changing covariances of shocks. This model, like a small four-variable model with or without heteroskedasticity, includes a rather large output reduction in response to a positive “business loans” shock. Unlike the corresponding shock in the full model estimated *with* heteroskedasticity and *without* triangular restrictions on A_0 , the fourth shock of this model includes immediate and substantial increases in the corporate bond spread (GZ), interbank lending spread (ES), and Federal Funds Rate (R). We might suspect this is a shock that represents business (but not household) borrowing in the midst of financial stress and monetary policy stringency.

Which of the two changes—adding heteroskedasticity or removing restrictions—seems more important in this context? To investigate, we can estimate a model with regime-switching heteroskedasticity and over-identifying triangular restrictions. The impulse response of such a model (with Gaussian errors) is plotted in Figure 15. The fourth shock of this model has a considerably smaller (and less confidently measured) output effect. We consider this to be evidence that downweighting high variance observations, perhaps to a greater extent than removing contemporaneous restrictions, has a large effect on results.

VI. CONCLUSION

Credit conditions, monetary policy, and real activity interact dynamically through multiple channels. To study these interactions, we construct and estimate structural multiple-equation models that are identified without strong *a priori* assumptions. Our analysis distinguishes impulses and feedbacks that focused study of individual channels might miss.

Our main model includes ten independent shocks that are identified by substantially changing volatilities across exogenously specified regimes. The data strongly favor additional corrections for “fat tails” in the distributions of the structural innovations, though the main qualitative conclusions are the same without them. Further refining (and possibly endogenizing) a model specification for volatility remains a task for future research, but addressing the issue in some way greatly improves model fit and affects implied dynamics.

Monetary policy is identified without any timing restrictions and seems to be amplified through inter-bank credit spreads. Two other model shocks look like “stress shocks” which originate in the financial sector and propagate to the real economy after several months of delay. The distinction between these shocks, which start with impulses to corporate bond spreads and interbank rate spreads respectively, is potentially very important for emerging research on the role of lending frictions and risk premia in the macroeconomy. A related takeaway for forecasters is that one-dimensional metrics of financial conditions may be insufficient for capturing risks for the real economy.

While these credit spread shocks do have strong real effects, they do not provide more than a few months of “advance warning” of an output contraction. In recursive-out-of-sample forecasts around the 2008 financial crisis, including additional credit spread variables only improves forecasts in a narrow window at the beginning of the downturn. Across the entire data sample, there is no clear evidence that including credit variables improves forecasting performance.

Credit aggregates, in this framework, almost always moving “passively” in the same direction as output. A single shock generates opposite movements in household (real estate plus consumer) credit and output, but in all periods the magnitude of this effect is relatively small. To the extent that this effect is quantitatively important, a multivariate model is necessary to properly separate it from other effects. We run multiple checks that suggest that models with fewer variables and no correction for heteroskedasticity can imply drastically more statistically and economically significant effects of credit growth

on output even if credit is mostly passive. A more comprehensive check of all possible non-linearities and verification against an international panel of data is left to further research. But our results demonstrate how multiple time-series analysis, without strong *a priori* restrictions, can shed light on complex interactions among policy, financial markets, and the real economy.

REFERENCES

- Bernanke, Ben S., Mark Gertler, and Simon Gilchrist.** 1999. "The financial accelerator in a quantitative business cycle framework." In *Handbook of Macroeconomics*. Vol. 1 of *Handbook of Macroeconomics*, , ed. J. B. Taylor and M. Woodford, Chapter 21, 1341–1393. Elsevier.
- Brunnermeier, Markus K., and Yuliy Sannikov.** 2016. "The I Theory of Money." Princeton University Working Paper.
- Caldara, Dario, and Edward Herbst.** 2016. "Monetary Policy, Real Activity, and Credit Spreads: Evidence from Bayesian Proxy SVARs." Board of Governors of the Federal Reserve System Finance and Economics Discussion Series Paper 2016-049.
- Chib, Siddhartha, and Srikanth Ramamurthy.** 2014. "DSGE Models with Student-t Errors." *Econometric Reviews*, 33(1-4): 152–171.
- Chiu, Ching-Wai, Haroon Mumtaz, and Gabor Pinter.** 2015. "Forecasting with VAR models: fat tails and stochastic volatility." Bank of England Bank of England working papers 528.
- Christiano, Lawrence J., Roberto Motto, and Massimo Rostagno.** 2014. "Risk Shocks." *American Economic Review*, 104(1): 27–65.
- Cúrdia, Vasco, Marco Del Negro, and Daniel L. Greenwald.** 2014. "Rare Shocks, Great Recessions." *Journal of Applied Econometrics*, 29(7): 1031–1052.
- Del Negro, Marco, and Frank Schorfheide.** 2013. "DSGE Model-Based Forecasting." *Handbook of Economic Forecasting*, , ed. Graham Elliott and Allan Timmermann Vol. 2, Part A of *Handbook of Economic Forecasting*, Chapter 2, 57 – 140. Elsevier.
- Dembiermont, Christian, Mathias Drehmann, and Siriporn Muksakunratana.** 2013. "How much does the private sector really borrow - a new database for total credit to the private non-financial sector." *BIS Quarterly Review*, , (March): 65–81.
- Drechsler, Itamar, Alexi Savov, and Philipp Schnabl.** 2016. "A Model of Monetary Policy and Risk Premia." *Journal of Finance*, NA((forthcoming)): NA.
- Drehmann, Mathias, and Mikael Juselius.** 2014. "Evaluating early warning indicators

- of banking crises: Satisfying policy requirements." *International Journal of Forecasting*, 30(3): 759–780.
- Gertler, Mark, and Peter Karadi.** 2015. "Monetary Policy Surprises, Credit Costs, and Economic Activity." *American Economic Journal: Macroeconomics*, 7(1): 44–76.
- Gertler, Mark, and Simon Gilchrist.** 1993. "The Role of Credit Market Imperfections in the Monetary Transmission Mechanism: Arguments and Evidence." *The Scandinavian Journal of Economics*, 95(1): 43–64.
- Gilchrist, Simon, and Egon Zakrajšek.** 2012. "Credit Spreads and Business Cycle Fluctuations." *American Economic Review*, 102(4): 1692–1720.
- Gilchrist, Simon, Vladimir Yankov, and Egon Zakrajšek.** 2009. "Credit market shocks and economic fluctuations: Evidence from corporate bond and stock markets." *Journal of Monetary Economics*, 56(4): 471 – 493.
- Goldsmith, Raymond.** 1969. *Financial Structure and Development*. Yale University Press.
- Hamilton, James D., Daniel F. Waggoner, and Tao Zha.** 2007. "Normalization in Econometrics." *Econometric Reviews*, 26(2-4): 221–252.
- Ivashina, Victoria, and David Scharfstein.** 2010. "Bank lending during the financial crisis of 2008." *Journal of Financial Economics*, 97(3): 319 – 338. The 2007-8 financial crisis: Lessons from corporate finance.
- Jordà, Òscar, Moritz Schularick, and Alan M. Taylor.** 2013. "When Credit Bites Back." *Journal of Money, Credit and Banking*, 45(s2): 3–28.
- Jordà, Òscar, Moritz Schularick, and Alan M. Taylor.** 2014. "The Great Mortgaging: Housing Finance, Crises, and Business Cycles." National Bureau of Economic Research Working Paper 20501.
- Jordà, Òscar, Moritz Schularick, and Alan M. Taylor.** 2015. "Betting the house." *Journal of International Economics*, 96, Supplement 1: S2 – S18. 37th Annual NBER International Seminar on Macroeconomics.
- Krishnamurthy, Arvind, and Tyler Muir.** 2016. "How Credit Cycles across a Financial Crisis." Stanford GSB Working Paper.

- Lanne, Markku, and Helmut Lütkepohl.** 2010. "Structural Vector Autoregressions With Nonnormal Residuals." *Journal of Business & Economic Statistics*, 28(1): 159–168.
- Lanne, Markku, Helmut Lütkepohl, and Katarzyna Maciejowska.** 2010. "Structural vector autoregressions with Markov switching." *Journal of Economic Dynamics and Control*, 34(2): 121 – 131.
- Lopez-Salido, David, Jeremy C. Stein, and Egon Zakrajsek.** 2015. "Credit-Market Sentiment and the Business Cycle." Board of Governors of the Federal Reserve System Working Paper.
- McKinnon, Ronald.** 1973. *Money and Capital in Economic Development*. Washington, D.C.:Brookings Institution.
- Mian, Atif R., Amir Sufi, and Emil Verner.** 2015. "Household Debt and Business Cycles Worldwide." National Bureau of Economic Research Working Paper 21581.
- Rigobon, Roberto.** 2003. "Identification through Heteroskedasticity." *The Review of Economics and Statistics*, 85(4): 777–792.
- Schularick, Moritz, and Alan M. Taylor.** 2012. "Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870-2008." *American Economic Review*, 102(2): 1029–61.
- Shaw, Edward Stone.** 1973. *Financial Deepning in Economic Development*. New York:Oxford Univ. Press.
- Sims, Christopher A., and Tao Zha.** 1996. "Bayesian methods for dynamic multivariate models." Federal Reserve Bank of Atlanta FRB Atlanta Working Paper 96-13.
- Sims, Christopher A., and Tao Zha.** 1999. "Error Bands for Impulse Responses." *Econometrica*, 67(5): 1113–1155.
- Sims, Christopher A., and Tao Zha.** 2006. "Were There Regime Switches in U.S. Monetary Policy?" *American Economic Review*, 96(1): 54–81.
- Smets, Frank, and Rafael Wouters.** 2007. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review*, 97(3): 586–606.

DEPARTMENT OF ECONOMICS, PRINCETON UNIVERSITY

E-mail address: markus@princeton.edu

RUTGERS BUSINESS SCHOOL

E-mail address: dpalia@rci.rutgers.edu

DEPARTMENT OF ECONOMICS, MIT

E-mail address: ksastry@mit.edu

DEPARTMENT OF ECONOMICS, PRINCETON UNIVERSITY

E-mail address: sims@princeton.edu

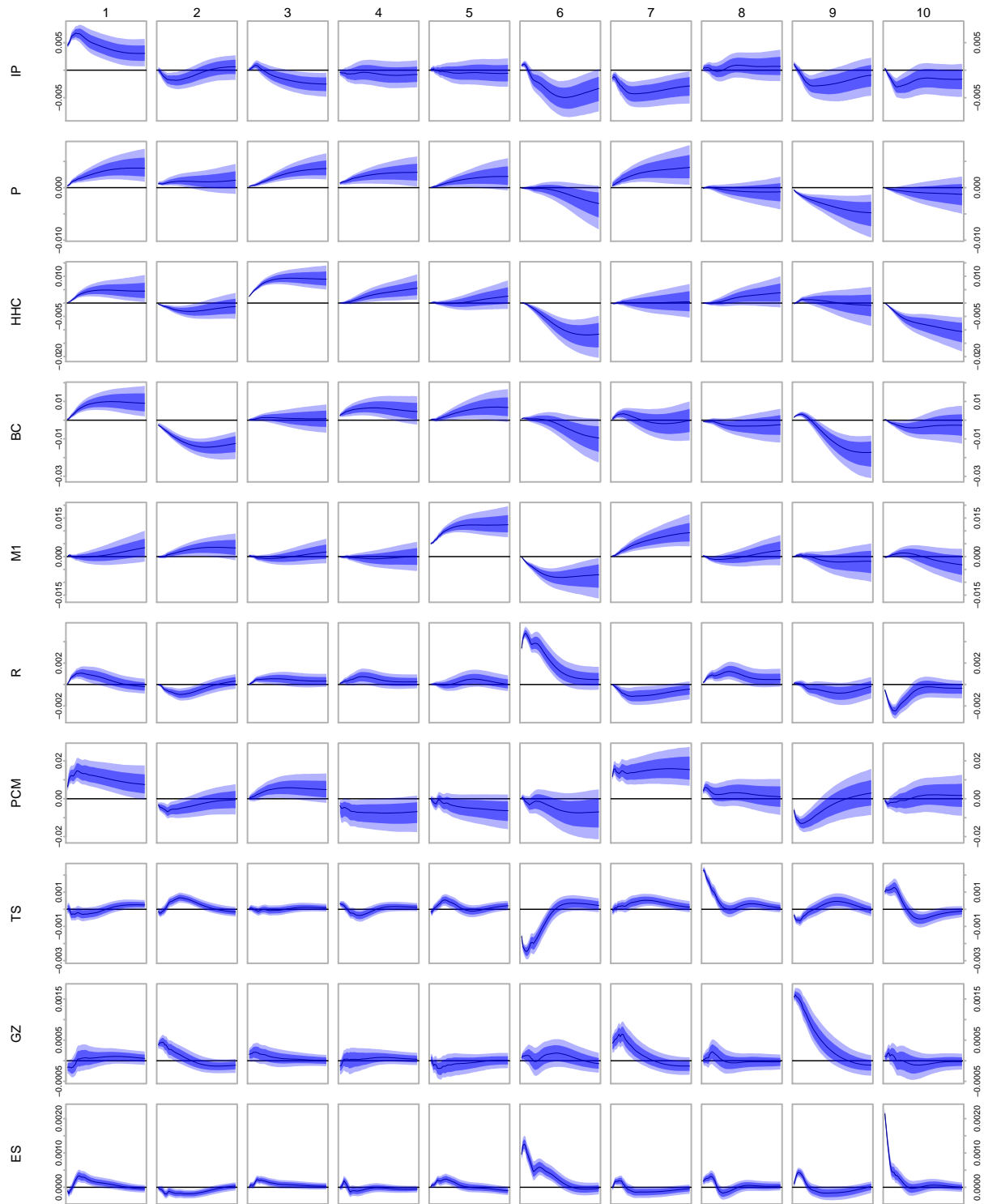


FIGURE 2. Impulse responses to the ten orthogonal structural shocks in the model with t distributed errors over 60 months, with 68% (dark blue) and 90% (light blue) posterior uncertainty regions. Scaled to an “average” period with unit variances. Variables are in the order listed in Table 1.

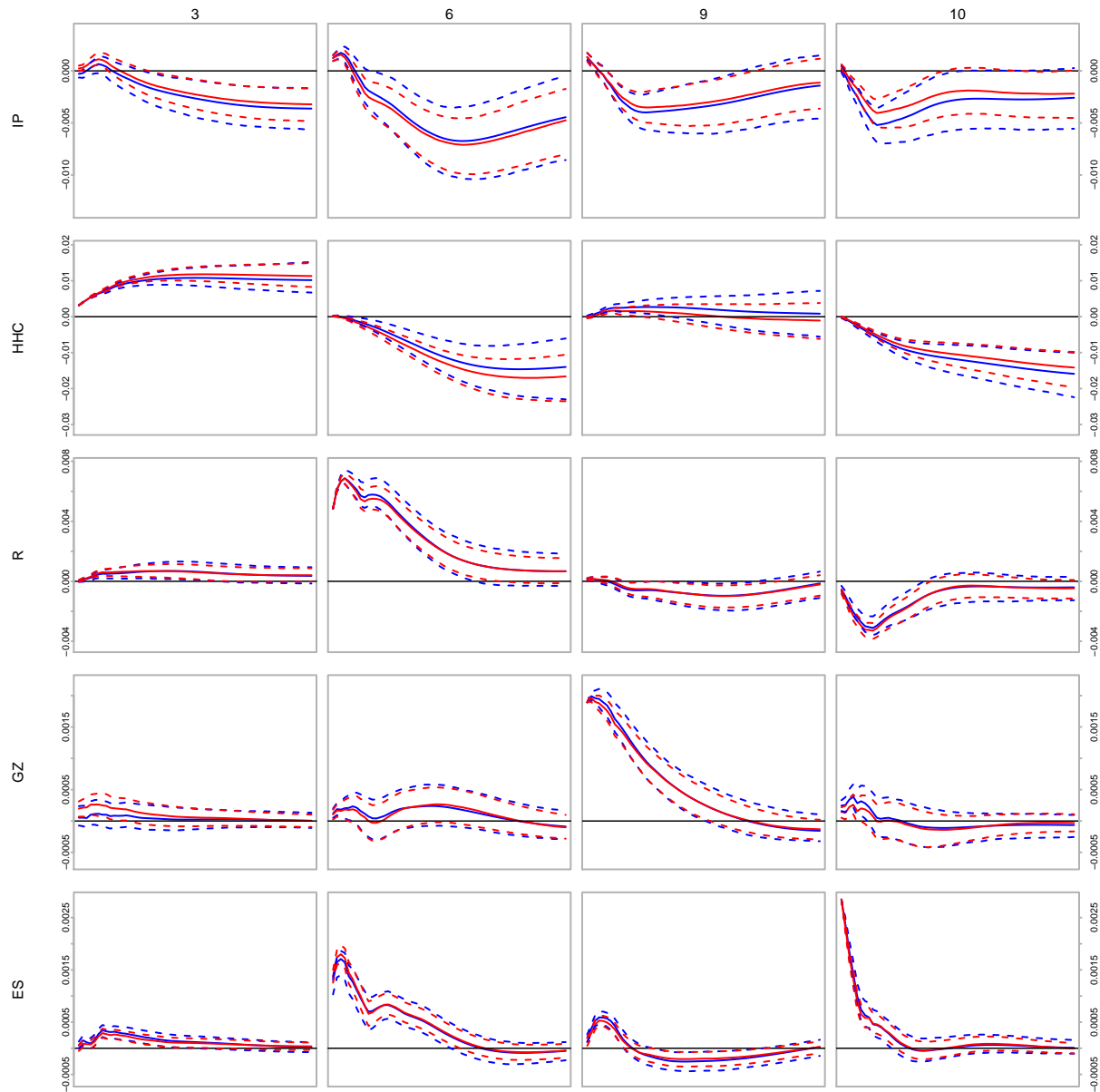


FIGURE 3. Impulse responses for shocks 3, 6, 9 and 10, Gaussian model in blue, *t* model in red, scaled to have identical sized initial shocks in HHC, R, GZ and ES, respectively. The solid lines are from the posterior mode and median respectively, and dotted lines are 68% bands.

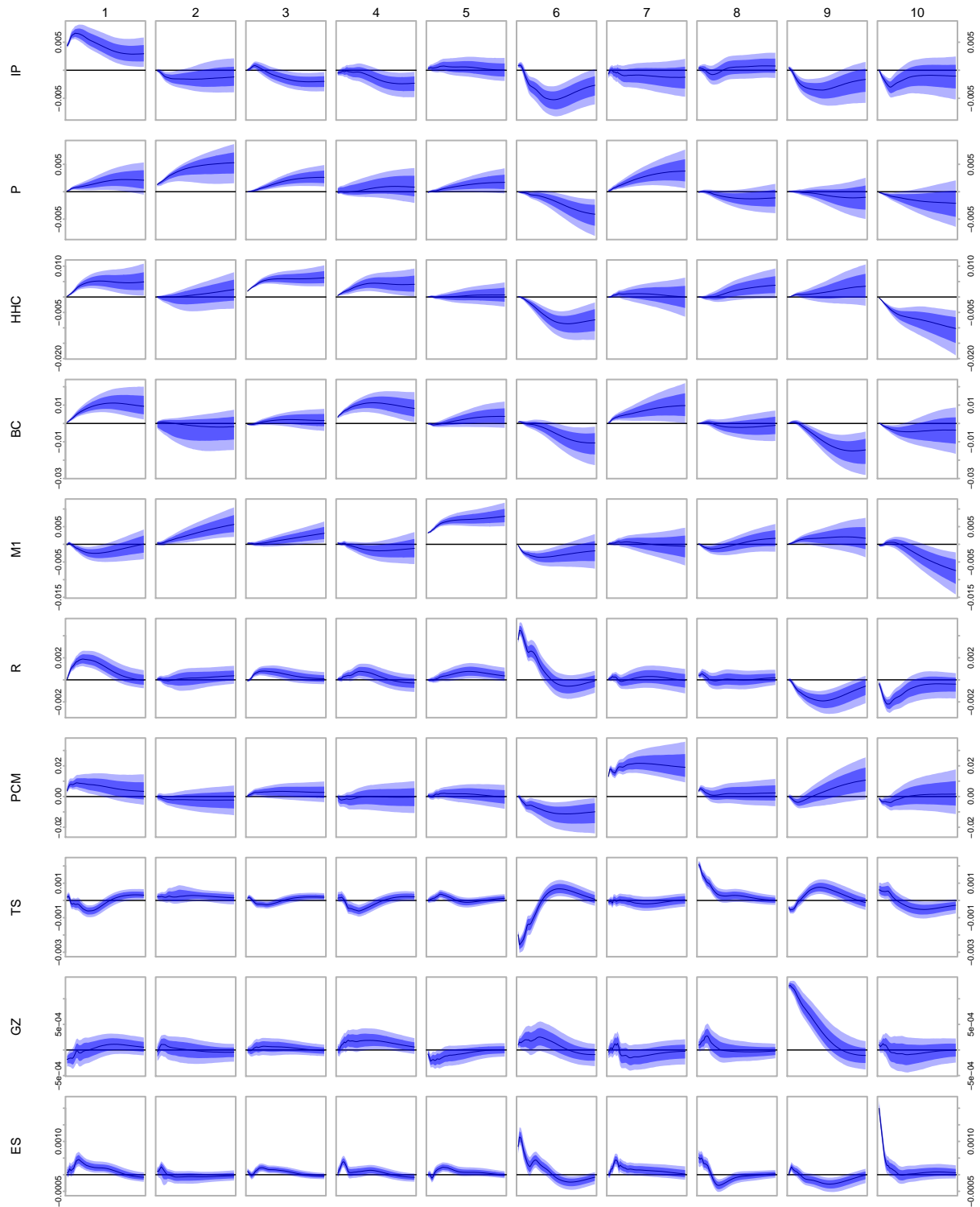


FIGURE 4. Impulse responses to the ten orthogonal structural shocks in a model with *t*-distributed errors estimated up to December 2007.

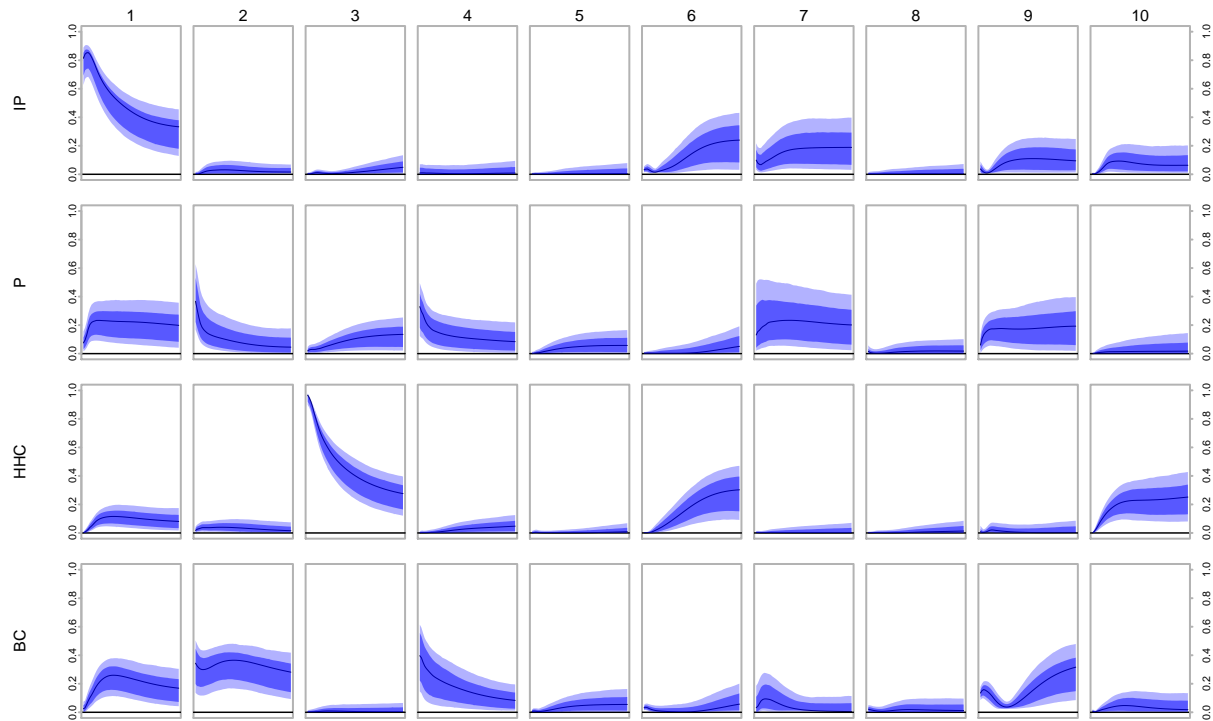


FIGURE 5. Forecast variance decompositions in the t -distributed errors model for IP, prices, household credit, and business credit over 60 months, with 68% (dark blue) and 90% (light blue) posterior uncertainty regions. Scaled to an “average” period with unit variances. Variables are in the order listed in Table 1.

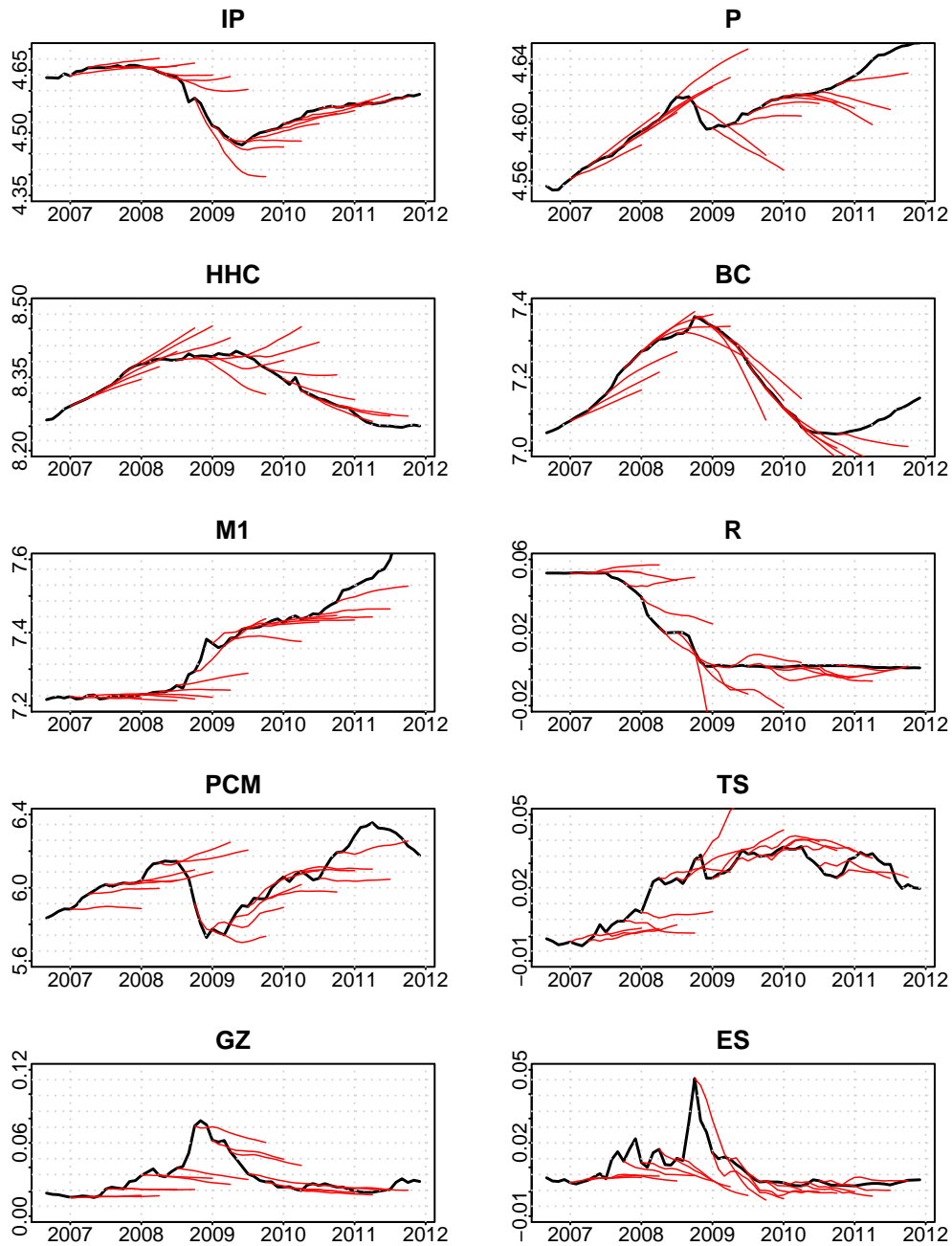


FIGURE 6. Posterior mode forecasts, from a model with all ten variables, estimated up to points in and around the Great Recession. (NOTE: need to fill out the missing forecasts)

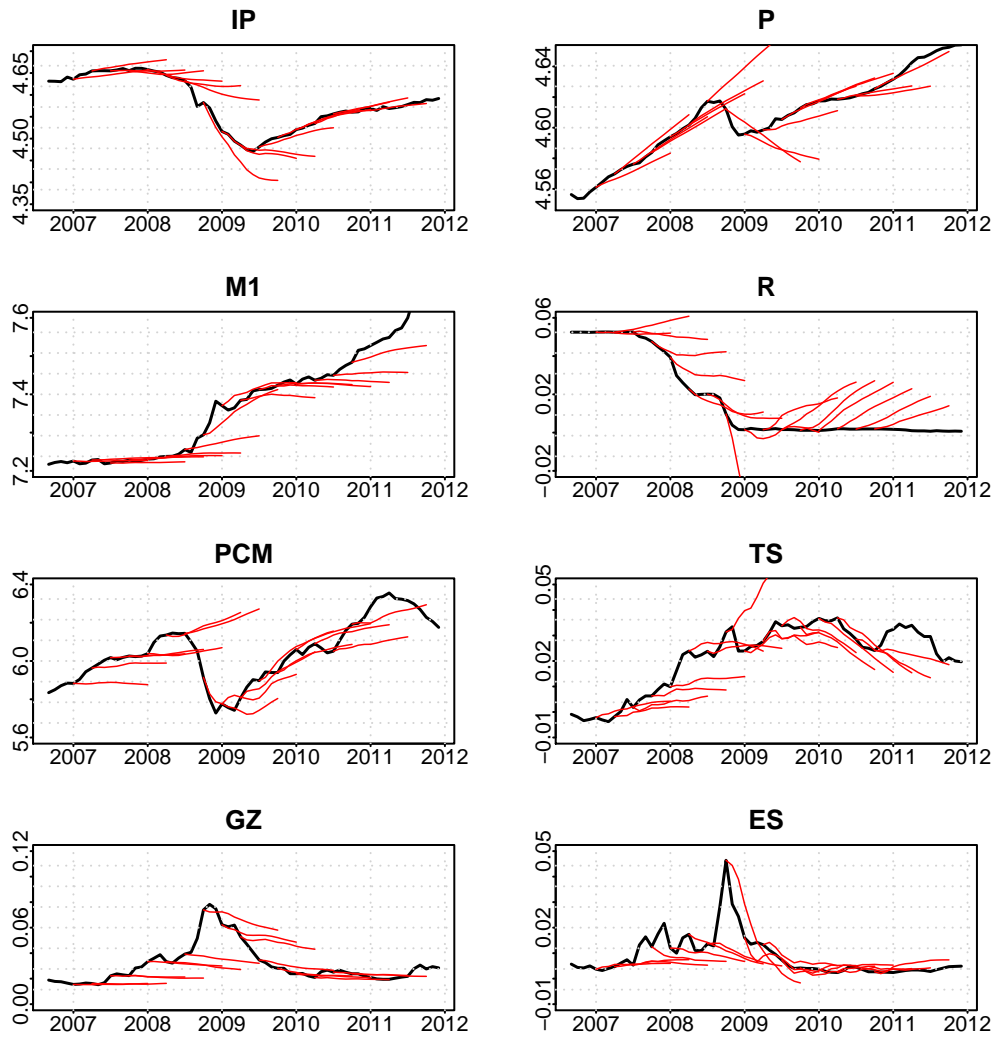


FIGURE 7. Posterior mode forecasts, from a model without the two credit aggregates, estimated up to points in and around the Great Recession.

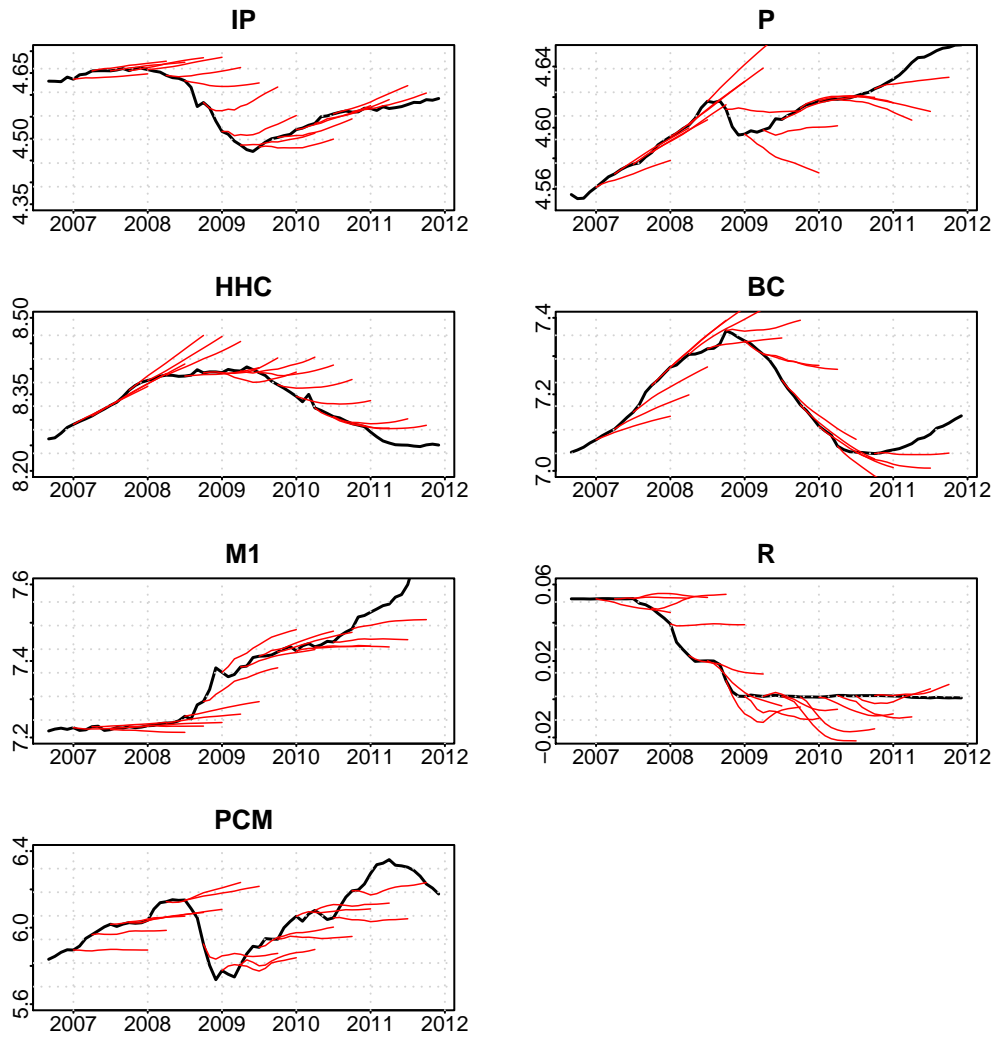


FIGURE 8. Posterior mode forecasts, from a model without the three credit spreads, estimated up to points in and around the Great Recession.

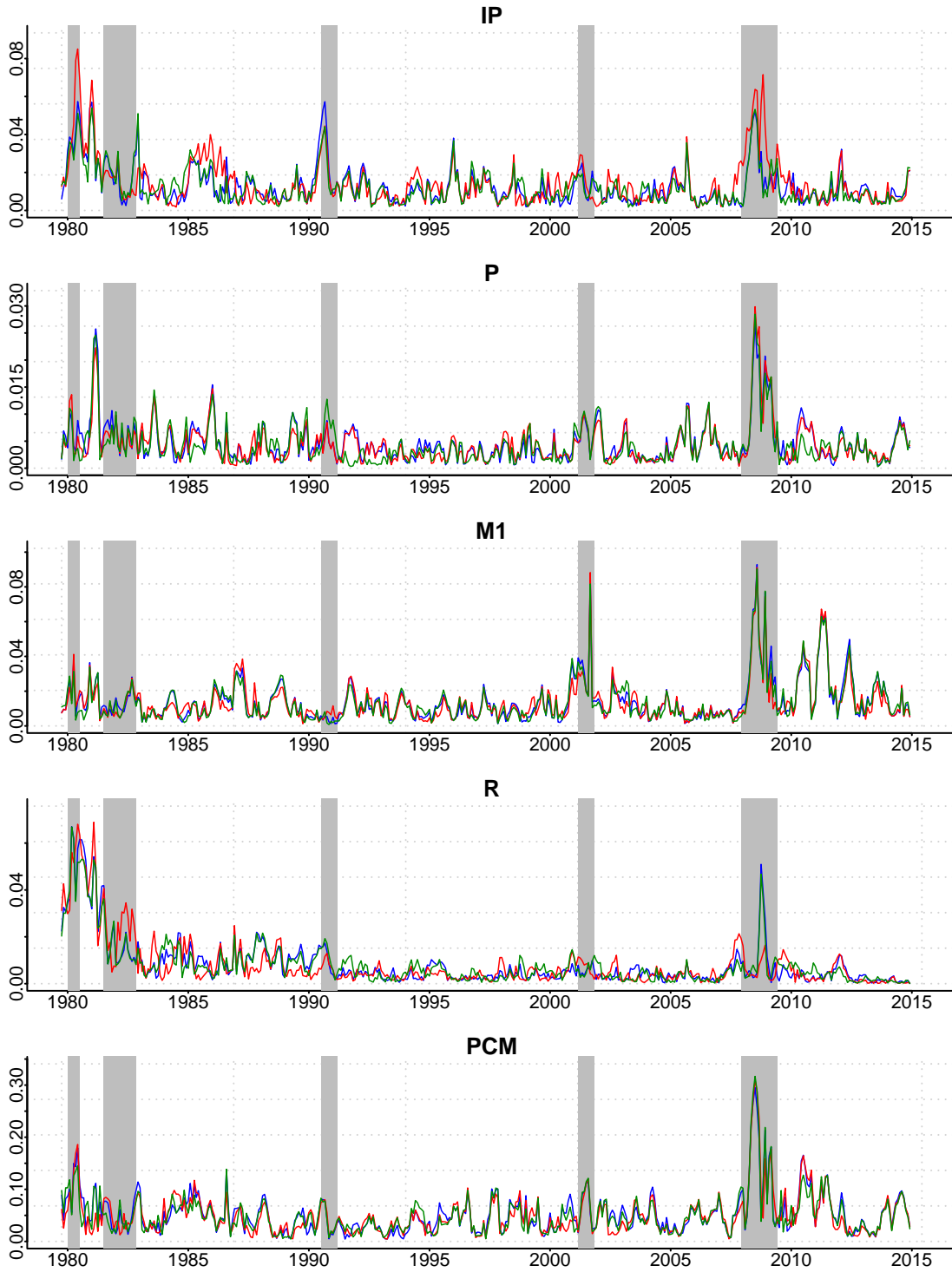


FIGURE 9. Root mean squared error (RMSE) for 6-month forecasts from rolling estimations of the Gaussian errors model with all variables (blue), no credit spreads (red), and no credit aggregates (green). NBER recessions are shaded.

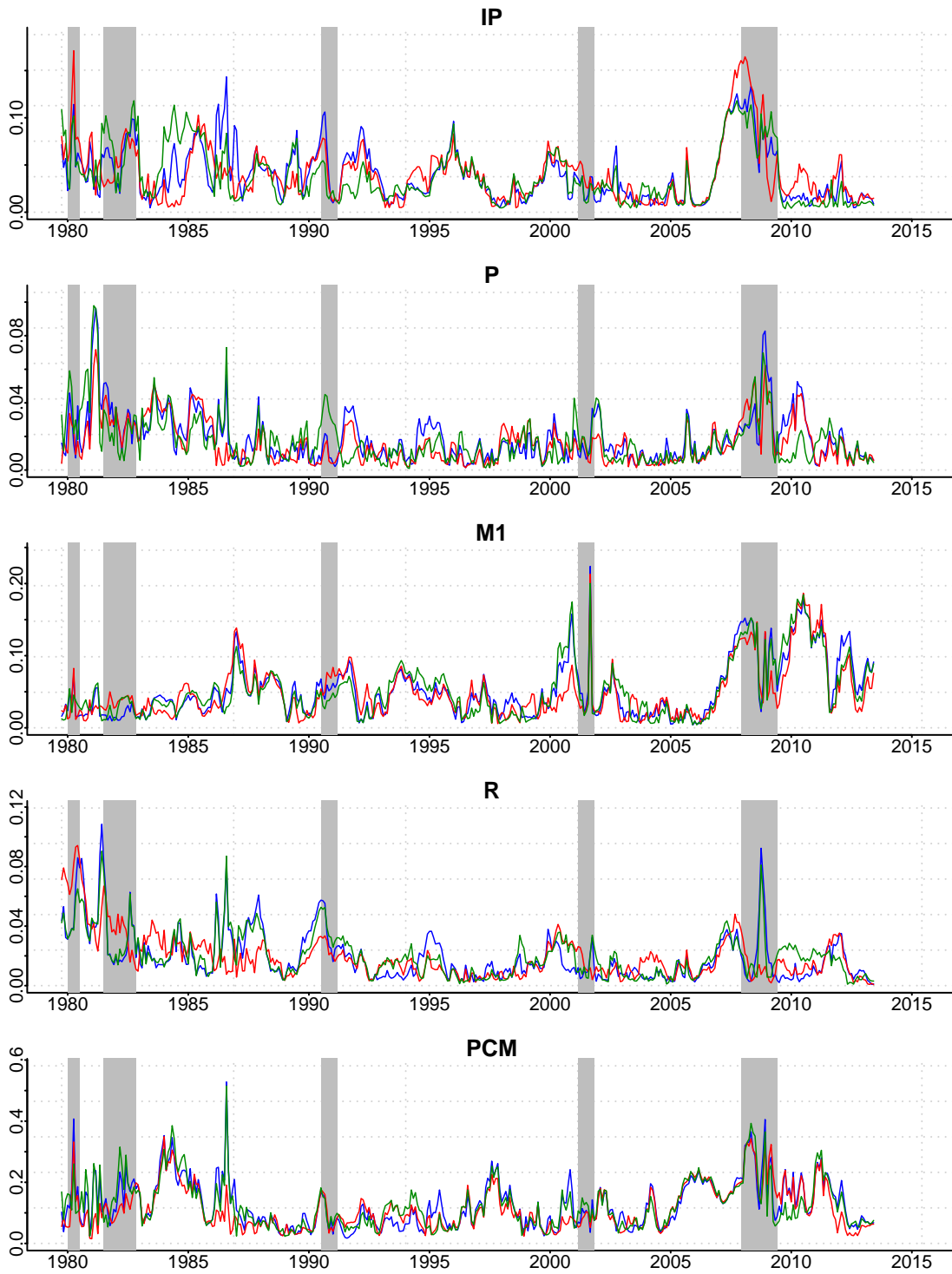


FIGURE 10. Root mean squared error (RMSE) for 24-month forecasts from rolling estimations of the main model with all variables (blue), no credit spreads (red), and no credit aggregates (green). NBER recessions are shaded.

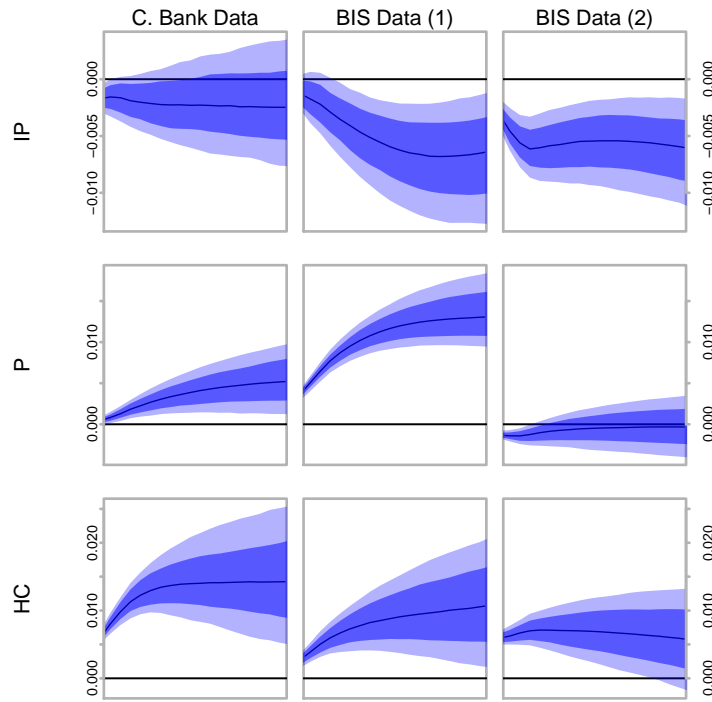


FIGURE 11. Impulse response to candidate bad credit growth shocks in the quarterly commercial bank data model (leftmost) and BIS credit data (middle and right), over 5 years. Shaded areas are 68 and 90% posterior bands.

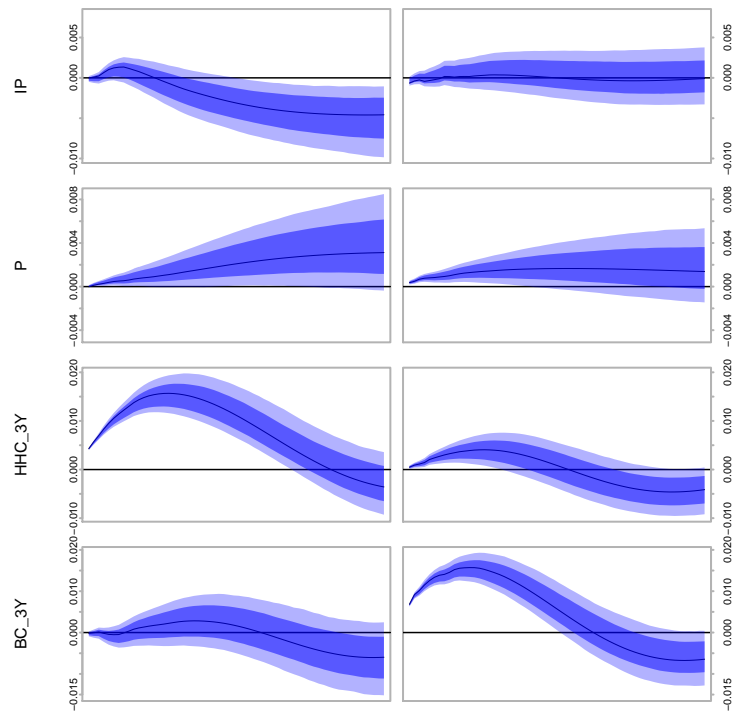


FIGURE 12. Impulse responses to “candidate” credit growth shocks in a (Gaussian errors) model with 3 year growth rates of credit aggregates.

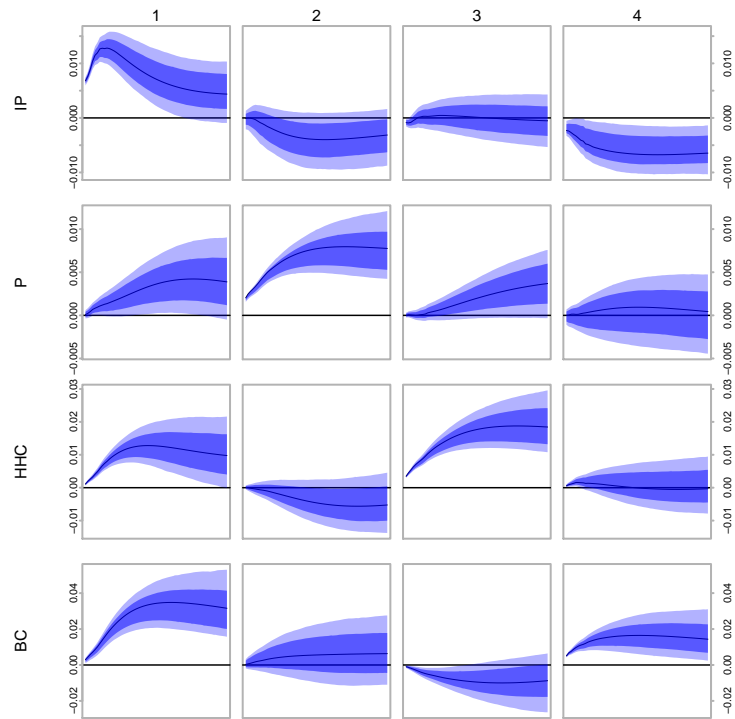


FIGURE 13. Impulse responses to four structural shocks in a “small model” over 60 months, with 68% (dark blue) and 90% (light blue) posterior uncertainty regions. Scaled to an “average” period with unit variances.

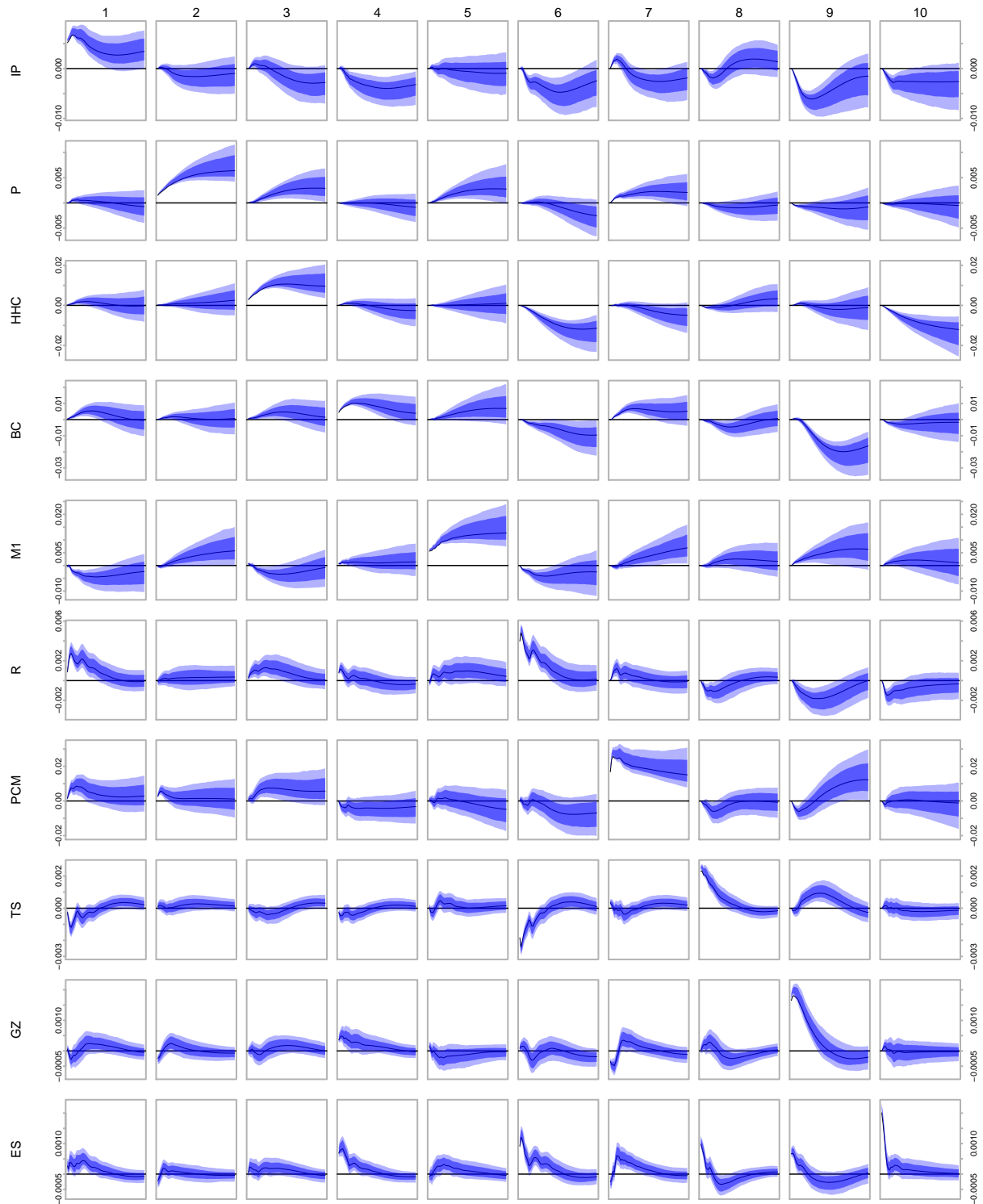


FIGURE 14. Impulse responses to 10 structural shocks in a Cholesky-identified VAR with constant structural variances over 60 months, with 68% (dark blue) and 90% (light blue) posterior uncertainty regions.

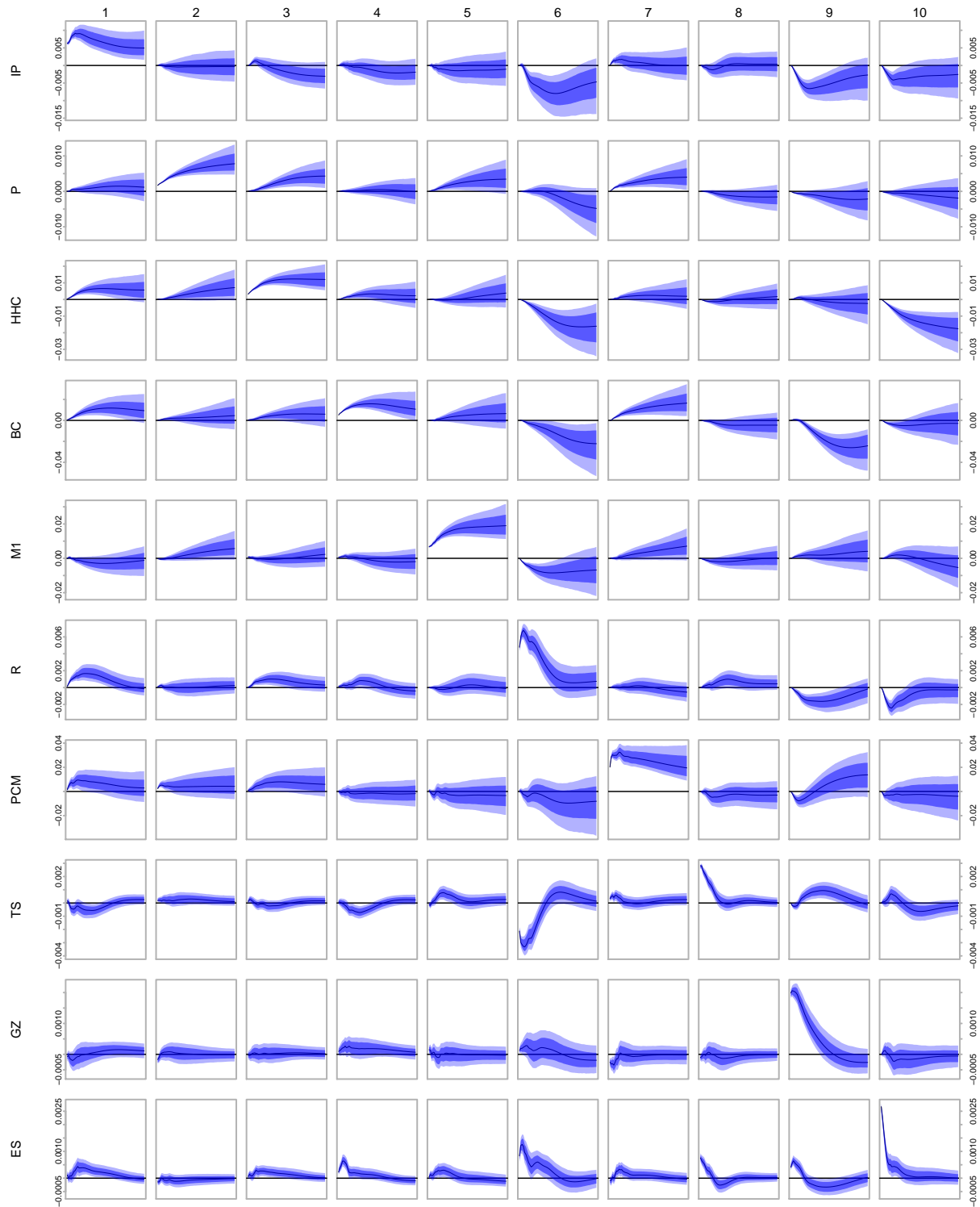


FIGURE 15. Impulse responses to 10 structural shocks in a model with triangular restrictions and heteroskedasticity over 60 months, with 68% (dark blue) and 90% (light blue) posterior uncertainty regions.