Feedbacks: Financial Markets and Economic Activity[†]

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

Is credit expansion a sign of desirable financial deepening or the prelude to an inevitable bust? We study this question in modern US data using a structural VAR model of 10 monthly frequency variables, identified by heteroskedasticity. Negative reduced-form responses of output to credit growth are caused by endogenous monetary policy response to credit expansion shocks. On average, credit and output growth remain positively associated. "Financial stress" shocks to credit spreads cause declines in output and credit levels. Neither credit aggregates nor spreads provide much advance warning of the 2008–2009 crisis, but spreads improve within-crisis forecasts. (*JEL* C51, E23, E31, E43, E44, E52, G01)

Credit aggregates tend to expand faster than GDP in the long run. Studies of economic development sometimes use the ratio of credit to GDP as a measure of "financial depth" thought to boost economic growth.[1] On the other hand a number of recent studies claim to have demonstrated a predictive relation between rapid credit growth and low future GDP growth or higher likelihood of crisis.[2]

Both mechanisms, of course, could be consistent with the same reduced-form evidence. A useful analogy is the relationship between interest rates and inflation. The structural VAR literature on monetary policy effects successfully separates two opposite-sign mechanisms connecting these variables. Policy-generated increases in rates reduce inflation, while rates on average endogenously rise with inflation to compensate investors for inflation-generated losses. This pattern consistently emerges when a number of different identifying assumptions are imposed on multiple-equation models.

[1] Perspectives on this topic are found, for example, in Shaw (1973); McKinnon (1973); Goldsmith (1969); and Rajan and Zingales (1998). A summary of the related literature is available in World Bank (2012, pp. 23–25).
[2] For example, Mian, Sufi, and Verner (2017); Schularick and Taylor (2012); Jordà, Schularick, and Taylor (2016); Borio (2012); and Drehmann and Juselius (2014).

It is reasonable to suspect that the relationship between credit market outcomes (quantities and prices) and macro aggregates similarly involves multiple, opposite-signed causal mechanisms. An important such mechanism could be monetary policy itself, which strongly affects and responds to credit conditions. Unraveling this and other feedback in the data is essential for probing the policy implications of predictive findings. Is "bad credit growth" the exception or the norm in modern economies? Should central banks raise rates when credit expands rapidly? Would monitoring credit conditions significantly improve policy and/or macro forecasting?

We approach these challenges in a way that is inspired by the monetary literature: with a large-scale structural vector autoregressive (SVAR) model that jointly identifies multiple causal channels. Our focus is US data from the 1970s to the present day for credit aggregates, credit spreads, and standard macro variables. The familiarity of this context, and the large quantitative and narrative literature about it, allows a simple identification strategy (due, in economics, to Rigobon 2003): assuming that different shocks' relative variance changes across notable episodes in recent history (e.g., the Volcker disinflation versus the Great Moderation) while macro dynamics remain constant. We began by using other identifying restrictions, yet modeling this time-varying variance because it was so clearly needed to fit the data. But we discovered as we proceeded that identification via heteroskedasticity approach produced stable and interpretable results without the additional restrictions.

We find some evidence of long-run negative response to credit growth. But essentially all of this can be attributed to endogenous monetary tightening in response to associated medium-run output and inflation growth. We validate, in simulation, how omitting these full model dynamics would lead to misleading results in small, reduced-form models.

Shocks to credit spreads generate substantial contractions in output and credit, as do monetary policy shocks. The monetary policy shocks have somewhat larger effects, on average, than the spread shocks. The monetary policy shocks explain a substantial part of the variation in the credit aggregates and in one of the spread variables. These shocks are separately and stably distinguished from one another because of their variability peaking at different points in history (e.g., the Volcker disinflation versus the Great Recession).

We also explore how useful our model would be in pseudo-out-of-sample forecasting and what the contribution to prediction accuracy of the credit and spread variables might be. We find that credit aggregates are not particularly valuable for improving fit. Credit spreads are valuable for matching the behavior of the economy in the midst of crisis but, less so outside of this context. Neither spreads nor credit aggregates provide an "early warning indicator" over longer horizons.

Related Literature.—Much of the existing empirical literature in this area has used short lists of variables and has not attempted to distinguish several channels of interaction between financial variables and the macroeconomy, including the one modulated by monetary policy. Studies of the predictive power of credit growth have primarily used single-equation projection methods (e.g., Mian, Sufi, and Verner 2017; Jordà, Schularick, and Taylor 2015, 2016) or binary outcome (i.e., crisis or no

crisis) predictive models (e.g., Schularick and Taylor 2012; Drehmann and Juselius 2014).

Jordà, Schularick, and Taylor (2013) uses 11 variables and estimates impulse responses by local projection. It uses no spread variables and focuses on whether knowledge of whether a recession is "financial" changes forecasts of the length or depth of the recession. Mian, Sufi, and Verner (2017), like this paper, includes 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.

Our approach has three benefits and two costs relative to this literature. On the benefits side, we use (i) more comprehensive data on different macro and financial aggregates at (ii) a higher frequency combined with (iii) a more complex identification strategy. This allows us to more precisely hone in on the causal effect of "credit shocks," differentiated from other identified macro shocks. The costs are that we (i) focus on a single country, the United States and (ii) have a shorter time-period of study. We leave the possible extension of our more structural methods, with appropriate flexibility to capture heterogeneity in economic structure and/or monetary regimes across countries and longer time periods, to future work.

Studies focused on the information in credit spreads, similarly to the aforementioned credit growth literature, have looked extensively at single-equation models (e.g., López-Salido, Stein, and Zakrajšek 2017; Krishnamurthy and Muir 2017) and reduced form multi-equation models (e.g., Gilchrist, Yankov, and Zakrajšek 2009; Gilchrist and Zakrajšek 2012a). Gertler and Karadi (2015) and Caldara and Herbst (2019) introduce credit spread variables into structurally identified, multiple-equation frameworks with monetary policy. But these authors have a narrower focus on identifying and interpreting monetary policy shocks, relative to the rest of the system, and do not discuss the role of credit aggregates. Krishnamurthy and Muir (2017) does look at both aggregates and spreads in the same framework. But their main specifications, single-equation models which can include interactions (nonlinear transformations) of credit growth and credit spreads, do not solve the endogeneity problem. They document increased negative skewness of future GDP growth forecast errors following an increase in spreads, which our model cannot account for. Extending our model to allow such asymmetry in shock distributions is an interesting challenge for future research.

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 of monetary policy.[3] 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.

---

[3]Christiano, 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 US business cycle dynamics. The model uses data on credit spreads 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.

Stock and Watson (2012) use a six-factor dynamic factor model estimated from a collection of hundreds of time series to study some questions that overlap with those we address in this paper. They focus attention on interpreting the 2008–2012 crash and recovery, concluding that it seems to be best explained as unusually large shocks feeding through a stable dynamic structure, which matches the assumptions underlying our modeling approach. When they attempt structural interpretation of their model, they use an external instruments approach. We discuss this approach to identification and contrast it with our own in Section IB.

Details and Roadmap.—Section I describes the econometric model in detail. Our baseline specification is a Bayesian SVAR with heteroskedasticity and non-normal (Student's t) distributed errors. Section II describes the data and estimation procedure. We use 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 (2012a) corporate bond spread (GZ), and the 3-month Eurodollar over Treasury spread (ES). The sample period runs from January 1973 to June 2015.

Section III lays out our main results, the identified shocks, and impulse responses. Section IV demonstrates, via simulation, how such results are consistent with credit growth predicting future slowdowns or crises in single-equation "prediction regressions." Section V conducts the pseudo-out-of-sample forecasting experiments. In Section VI we summarize variations on our model that we have investigated to check robustness of results.

The online Appendix includes details of our estimation methods and results for the variations of our model that we examined as robustness checks. We refer to relevant sections of the online Appendix at various points in the main text of this paper.

I. Framework

A. *The Basics*

Let y_t be an $n \times 1$ vector of observed variables in time periods $t \in \mathcal{T} := \{1, \ldots, T\}$. We model y_t with the following system of equations:

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

where A_0 is an $n \times n$ matrix of simultaneous relationships, $(A_j)_{j=1}^{p}$ are $n \times n$ matrices of coefficients at each lag j, C is an $n \times 1$ vector of constants, and ϵ_t is an $n \times 1$ vector of shocks independent across equations and time.

The variance of these shocks differs by time period. Let $\mathcal{M} = \{1, \ldots, M\}$ identify regimes over which variances are constant and the function $m : \mathcal{T} \to \mathcal{M}$ map dates to their respective regime. In each regime the variance of structural shocks is a different diagonal matrix Λ_m:

$$(2) \qquad E[\epsilon_t \epsilon_t'] = \Lambda_{m(t)}.$$

The coefficients $(A_j)_{j=0}^p$ remain fixed over all time periods.

As the model is written, 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

$$\frac{1}{M}\sum_{m=1}^M \lambda_{i,m(t)} = 1, \qquad \forall i \in \{1,\ldots,n\},$$

where $\lambda_{i,m(t)}$ is the ith diagonal element of $\Lambda_{m(t)}$. This makes the cross-period average structural variance 1 in each equation. 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.[4]

B. *Interpretation*

These assumptions restrict the economy to respond to shocks in the same way at all times, but allow the relative size of those shocks and, therefore, the relative size of their effects on the economy, to change. Plots of impulse responses are implied always to have the same shape but, across different regimes, different sizes.

The term "identification through heteroskedasticity" may suggest that the place to look for possible failure of the approach is the possibility of insufficient variability in reduced-form covariance matrices. But that the volatility of macro aggregates and of financial variables varies over time is not really in doubt. Economists refer to the "Volcker disinflation" in the early 1980s, when interest rates were unusually volatile and to the "Great Moderation" during the late 1980s through the early 2000s, when macro aggregates in general were less volatile than during the 1960s and 1970s. Furthermore, if the assumption of constant $A(L)$ were correct, but there was in fact little variation over time in the reduced form covariance matrices, the identification problem would reveal itself in a flat likelihood function and, therefore, very wide error bands on estimated impulse responses.

The Achilles' heel of this approach to identification, then, is not the possibility of little heteroskedasticity, but rather the possibility that the assumption of constant $A(L)$ is incorrect. Time variation in $A(L)$ would produce time variation in estimated reduced form covariance matrices Σ for a model that was fitted under the false assumption of constant $A(L)$. We provide evidence below that our specification outperforms some models that allow $A(L)$ to vary.[5]

---

[4] A formal proof of this can be found, for instance, in Lanne, Lütkepohl, and Maciejowska (2010). Here is the basic idea. If Σ_j is the reduced form residual covariance matrix for regime j, and our assumption of constant A_0 holds, then $\Sigma_j = A_0^{-1}\Lambda_j(A_0^{-1})'$. Given covariance matrices from two regimes i and j we can calculate

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

which 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.

[5] Sims and Zha (2006) consider models with a parametric form of time variation in $A(L)$. They find that posterior odds favor models with a rich specification of time varying heteroskedasticity over models with their form of time variation in $A(L)$.

Why Identification via Heteroskedasticity?—There are other approaches to making assumptions that allow identification in structural VARs. The most common are zero restrictions on A_0, long-run response restrictions, sign restrictions on responses, narrative approaches, and "external instruments." Much of the applied literature has focused on identifying just one structural shock, usually a monetary policy shock. In this paper, we are aiming at discovering interpretable shocks that correspond to influences of credit growth and financial market disturbances on the economy, while also distinguishing them from monetary policy shocks. Identification through heteroskedasticity lets us do this in a straightforward way, with assumptions, described above, that seem to us plausible.

The external instrument approach requires finding, for each identified structural shock, an observed variable that is correlated with that shock and not correlated with any other structural shock. This approach is plausible when the shock is the monetary policy shock and the instrument is the surprise in the setting of the Federal Funds rate, measured by the behavior of the Federal Funds futures market in narrow windows of time around a policy announcements. For the financial market disturbances we are interested in, though, this approach seems to us impractical or unconvincing. There are plenty of variables that are likely to be correlated with financial market disturbances, but none that plausibly are known a priori *not* to be correlated with any other source of disturbance to the economy. The strength of such assumptions is brought out when the external instrument approach is formulated, as in Plagborg-Møller and Wolf (2021), as the equivalent set of zero restrictions on a structural VAR.

Stock and Watson (2012), like this paper, attempts to separate interpretable sources of disturbance to the economy. They apply an external instruments approach, trying to identify six named sources of variation. They find that their estimated structural disturbances, though estimated by a method that assumes structural disturbances are independent, are strongly correlated. This reflects how difficult it is to find instruments that meet the stringent requirements of this approach to identification.[6]

The sign restrictions approach to identification uses informal, qualitative beliefs about the likely shape of impulse responses to structural shocks to achieve partial identification. Our approach uses similar qualitative beliefs to attach interpretations to our estimated structural shocks. The difference is that identification through heteroskedasticity generally allows point identification and consistent estimation, while the sign restrictions approach does not.

In most of the applied literature on structural VARs using other approaches to identification, time varying heteroskedasticity has not been modeled and constancy of $A(L)$ has been assumed. The central idea of SVAR modeling is that structural disturbances should not be cross-correlated. Otherwise there is a causal mechanism at play that has been missed by the model structure. Maintaining that assumption and constancy of $A(L)$ while allowing for time variation in Σ delivers identification through heteroskedasticity, if heteroskedasticity is present. Then applying additional

---

[6]The estimated structural shocks for our model have correlations averaging in absolute value about 0.01, with none as big as 0.1. This reflects our use of monthly data, in which correlations among innovations tend to be small even in reduced form, and also the fact that our model, which fits well, exploits the assumption of no correlation in estimation.

restrictions, to implement any of the other possible approaches to identification leads to a strictly less general model.

In online Appendix Section IV, we do check the correlations of variables we could have used as monetary policy external instruments with our estimated monetary policy shocks. The correlations are substantial, and the correlation of the instruments with our nonmonetary-policy structural shocks are much smaller. This suggests that our approach is not likely to have given results for monetary policy effects much different from those we could have obtained with the external instruments approach.

Model for Regime Switching.—Our choice of variance regimes in estimated models (discussed in Section II and presented in Table 2) is motivated by observed variation in 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.[7]

Note also that our results are not likely to be sensitive to modest changes in the boundaries of our assumed regime periods. The equations in footnote 4 would apply even if Σ_i and Σ_j were average values over two regimes of covariance matrices that varied within regimes. Errors in defining the regimes would tend to weaken results, widening error bands, by reducing the amount of variation in Σ_j across regimes, but would not introduce inconsistency.[8]

Rare Events.—Finally it is worth commenting on how our specification treats episodes like the financial crisis or the Volcker disinflation. While our method does implicitly identify shocks using these big sources of variation, it also *under-weights* them for estimating shock transmission (i.e., all other parameters), just as in a regression model generalized least squares (GLS) model high-residual-variance observations are down-weighted. Our concerns about having a few influential observations are sufficiently large that we also explore several distributional assumptions for the structural shocks ϵ_t.

C. *Distribution of ϵ_t*

We now return to specifying the model. The correction for heteroskedasticity will work best if volatilities mainly change between persistent episodes or regimes.

---

[7] 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.

[8] Sims (2020) provides an argument for this claim. In the case of Gaussian likelihood, inference is based entirely on sample second moments, and reduced form residuals are consistently estimated regardless of the pattern of heteroskedasticity. Therefore, if the number observations increases within each regime, all that is required for consistency is that the expected sample covariance matrix within each regime have the $A_0^{-1} \Lambda_t (A_0^{-1})'$ form, and that the Λ_t diagonals show sufficient variability. With the likelihood of independent t-distributed structural shocks, large outliers for individual shocks are penalized less than linear combinations of structural shocks of the same length. This means that even without heteroskedasticity, in contrast to the normal likelihood, the t likelihood penalizes orthonormal rotations of the structural shocks and thus allows consistent estimation of A_0. This is true even if the structural shocks are in fact not t-distributed, so long as they are independent and fat-tailed.

But it does not allow for the possibility of a few isolated large disturbances or outliers. For instance, the bankruptcy of Lehman Brothers in September 2008 and the 650 basis point drop in the Federal Funds rate from April to May 1980 generate outliers of around six standard deviations that do not disappear when we allow variance-regime switches. To guard against such large shocks distorting inference, we consider a specification in which structural errors ϵ_t are distributed as a mixture of normal distributions.[9]

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

$$(3) \qquad \epsilon_{i,t} \sim Normal\left(0, \lambda_{m(t),i}\xi_{i,t}\right).$$

We can also think of these objects as shocks which capture, in a simple way, a high-frequency component of volatility that has no persistence across time or correlation across equations.

Our main results assume the ξ_{it} are distributed as inverse-gamma:

$$(4) \qquad \xi_{i,t} \sim Inverse\ Gamma\left(shape = \alpha/2, rate = 2/\alpha\right).$$

This implies that each $\epsilon_{i,t}$ has an independent Student-t distribution with α degrees of freedom and unit scale.[10]

We chose the degrees of freedom of the t distribution to be 5.7, by fitting the sample distribution of residuals for the Gaussian-errors model.[11] Use of the t distribution mainly changes the statistical properties very far in tails: relative to a normal distribution, the $t(5.7)$ distribution puts 4 times the probability on a three-standard-deviation event but 200,000 times the probability on the 6 standard deviation event. The t implies we expect to see about two 6-σ residuals in our 10 equations and 510 months of data, while with Gaussian errors the probability of any 6-σ residuals in our sample is essentially zero. As we show in online Appendix Section VI, the number of large shocks is greater than the model expects even with the t assumption, suggesting room for further improvement in modeling the distribution of shocks.

Instead of adopting such a "two-step" procedure to calibrate the shape and scale of the t distribution, we could have also explored extensions in which we treated α as an unknown model parameter and estimated it directly. Here, as in a number of other cases of possible extensions of the parameter list, the likely improvement in

---

[9] The mixture-of-normals assumption has been used in the time series literature to better model large movements in macro variables. Lanne and Lütkepohl (2010) introduce a maximum likelihood approach to estimating a discrete normal mixture SVAR model, and Chiu, Mumtaz, and Pinter (2017) describe a Bayesian Gibbs sampling algorithm with an application to a model with stochastic volatility for US 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.

[10] Online Appendix Section VII.2 reports results from an alternative case with $\xi_{i,t}$ as independent k-multinomial, so that the distribution is a finite mixture of normals. The results from this specification are similar, and it does not fit as well as the t specification.

[11] We scaled the distribution of shocks to have unit variance, rather than setting the scale parameter of the t distribution to 1. All that matters to the likelihood is the shape of these distributions, not their scale, since A_0 can absorb differences in scale. But since our prior is not scale invariant, someone replicating our results might need to know about this scaling.

fit and effect on conclusions did not seem to justify the increased computational burden. Moreover, specifying the error distribution exactly "correctly" (and, in the first place, adopting the t specification instead of the standard Gaussian one), affects more small-sample efficiency and correct inference than the overall consistency of our results; and we will show later that our qualitative conclusions are similar also in the Gaussian model, though the t model fits much better.

II. Estimation

A. *Data*

Our main specification uses monthly data on 10 time series (listed in Table 1) from January 1973 to June 2015. We include data from the 1970s and early 1980s because, as discussed in the previous section, they are a valuable source of identifying variation for the structural shocks. The lag length p in our model is set to 10. All quantity variables are in log levels, and interest rates are in decimal units (i.e., $0.01 = 1$ percent $= 100$ basis points).

Our measures of "household" and "business" credit are based on the Federal Reserve's weekly surveys of US commercial banks.[12] 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 nonprofits," "nonfinancial noncorporate business," and "nonfinancial corporate business"), used in some other research.[13] Online Appendix Section V.2 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 seem 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.[14]

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 inflation expectations and 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, as policy-generated rises in the short rate might be expected to have little effect on, or even to lower, long rates, if the monetary tightening does succeed in lowering expected future inflation.

---

[12] These are published in the H.8 "Assets and Liabilities of Commercial Banks in the United States" release.

[13] In particular, the cross-country database, assembled by the Bank of International Settlements, uses these quarterly data.

[14] One practical complication is dealing with breaks in the credit series introduced by changes in accounting standards or major entrances to or exits from the commercial bank industry. Our specific calculations for eliminating these breaks, which are particularly large in the real estate credit series, are detailed in online Appendix Section V.1.

Table 1—Data Series Used

	Description
IP	Industrial production
P	Personal consumption expenditures price index
HHC	Sum of commercial bank real estate and consumer loans
BC	Commercial bank commercial and 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 (2012a) bond spread
ES	TED spread of 3-month Eurodollars over 3 month Treasuries

Sources: All series except PCM and GZ based on data from FRED, the Federal Reserve Bank of St. Louis database (https://fred.stlouisfed.org/). The names of the corresponding variables on FRED are IP: INDPRO; P: PCEPI; HHC: CONSUMER + REALLN; BC: BUSLOANS; M1: M1SL; R: FEDFUNDS; TS: GS10–TB3MS; ES: MED3–TB3MS. Note that HHC and BC required adjustment for breaks in coverage; see online Appendix Section V for details. GZ was downloaded from http://people.bu.edu/sgilchri/, which updates the data available at Gilchrist and Zakrajšek (2012b). PCM was downloaded from Global Financial Data under Princeton University's license for that database, where it has the code CMCRBSPD. Exact URLs and download dates for each series are available in the readme_data file in this paper's replication package (Brunnermeier et al. 2021). Raw data are also available there for all the series except PCM.

Table 2—Dates for Variance Regimes

	Start	End	Description
1	Jan 1973	Sep 1979	Oil crisis and stagflation
2	Oct 1979	Dec 1982	Volcker disinflation
3	Jan 1983	Dec 1989	Major S&L crisis defaults
4	Jan 1990	Dec 2007	Great moderation
5	Jan 2008	Dec 2010	Financial crisis and Great Recession
6	Jan 2011	Jun 2015	Zero lower bound, recovery from Great Recession

B. *Regime Choices*

We separate the full sample into six variance regimes described in Table 2. Our goal in selecting these regimes is to identify episodes in recent history in which the relative importance of driving sources of economic fluctuations were different, producing differing covariance matrices for innovations. In our first regime, the 1970s, oil market disturbances and somewhat erratic monetary policy prevailed. During the second regime, 1979–1982, the Fed under Paul Volcker targeted unborrowed reserves, resulting in much higher variance of residuals in a policy equation with interest rates on the left. The third and fourth regimes, 1983–1989 and 1990–2007 are sometimes all thought of as part of the "Great Moderation." Monetary policy was less erratic and most macro aggregates became more predictable. We separated the third and fourth regimes based on the idea that the S&L crisis defaults might have made the earlier period different. The fifth regime is the financial crisis, in which many macro variables, and especially financial variables,

became more volatile than anywhere else in the sample. And the last, post-crisis, regime is the period of zero lower bound monetary policy during crisis recovery.

Note that to make identification through heteroskedasticity work what is essential is that structural innovation variances differ on average across regimes. Defining regimes poorly, so that innovation variances vary little across regimes, would undermine identification and lead to wide error bands and ill-determined results. But identification does not depend on the regimes exactly capturing all variation in innovation variances.

C. *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 $n^2 p$ free parameters in the A_j. The t-distributed disturbances add another nT parameters. We use Bayesian methods to update beliefs about the parameters conditional on observed data $\{y_1, \ldots, y_T\}$ and initial conditions $\{y_{-p-1}, \ldots, y_0\}$.

Priors.—For A_0 we specify independent Gaussian priors on all elements, centered around 100 times the identity matrix, with standard deviation 200. For $\lambda_{.,i} = \{\lambda_{1,i}, \ldots, \lambda_{M,i}\}$, the vector of variances in each equation i, we put a Dirichlet prior (with $\alpha = 2$) on $\lambda_{.,i}/M$. This restricts each of the relative variances to lie in $[0, M]$ (where $M = 6$ in our main model), centers the prior on equal variances, and enforces our normalization that for each structural shock the relative variances average to one across periods. The centering at $A_0 = 100I$ implies we expect residual variances in each equation to be around 0.01 in order of magnitude, though with standard deviations of 200 the prior is quite diffuse. We use a variation of the "Minnesota prior" described in Sims and Zha (1996) on the remaining regression coefficients (described in more detail in online Appendix Section I).

Posterior Sampling.—We use a Gibbs sampling method to sample from the posterior distribution of all aforementioned parameters (see online Appendix Section II for the details). This also implies prior and posterior distributions for all (potentially nonlinear) transformations of the coefficients, including the impulse response functions for variable i to each shock j. In all reported results we display horizon-by-horizon 68 percent and 90 percent highest posterior density regions as error bands.

D. *Model Fit*

Table 3 displays measures of fit for our central model (the second line of the table) and for variants of the model. The third column shows each model's log marginal data density (MDD). The MDD is the integral over the parameter space of the joint distribution of data and parameters. If we regarded these four models as the full space of possible models, with equal prior probabilities for the models, the posterior probabilities on the models would be proportional to the MDD values. The differences between MDD values, exponentiated, can be interpreted as log odds ratios between the models: that is, how much more likely is one model than

TABLE 3—MARGINAL DATA DENSITIES

Variation in…	Distribution	MDD
None	Gaussian	19,830
Structural shock variances	t	21,434
Structural shock variances	Gaussian	20,712
All of $A(L)$ and A_0	Gaussian	20,296

Notes: Differences between values are log Bayes factors, or log posterior odds with equal prior weights on each model. The first row is for a model with Gaussian shocks, but no variation in $A(L)$ or shock variances across regimes. The second row is for the main model in the text, with time variation only in the variances of the structural shocks and t-distributed errors. The third row is the same model with Gaussian-distributed errors. The fourth row fits a separate unrestricted reduced-form model to each period; though it assumes Gaussianity, this model is otherwise the least restricted.

the other to capture the data, taking into account within-model uncertainty about parameter values.[15]

The model that comes closest to our main model is the one that differs only by assuming Gaussian, rather than t-distributed, errors. The Gaussian errors model nonetheless fits much worse: its posterior probability is implied to be smaller by a factor of over 10^{300}. Posterior odds ratios among finite collections of models often, as in this case, emerge as implausibly large or small, because they condition on the list of models being exhaustive. In fact, we know that in principle we could construct models, for example, that allowed richer parametric forms for the distribution of shocks, and that probably some variant on the t or normal distributions would provide a better fit. Nonetheless the MDD values do provide a useful measure of fit, even if we do not take the odds ratios literally.

The first line of the table refers to a model with no time variation in variances or in $A(L)$. Since in this case the SVAR is not identified, we estimated this model as a simple time-invariant reduced form VAR over the whole dataset. Comparing this model to the Gaussian-errors model with time varying variances in row 3, we see again a difference in log MDD of over 700. Allowing for time varying variances is extremely important for fit.

The bottom line of the table refers to a model with both the coefficients in $A(L)$ and the residual variances allowed to change at every change in regime. Here again, since there is no time variation within regimes, the regime-specific SVARs are not identified, and we therefore fit unrestricted reduced form VARs to each regime. Each regime has its own prior, so the small number of observations in some regimes creates no numerical problems in estimation. This model fits worse than the Gaussian SVAR with fixed $A(L)$ on the third line. Since the third-line model is nested in the fourth-line model, it may seem odd that the fourth-line model can fit

---

[15] The priors on $A(L)$ for the first three lines of Table 3 are implemented as dummy observations defining a prior on the reduced form $B(L)$, together with the normal prior on A_0 used in all the SVAR models. For the fourth line, in which only $B(L)$ is identified, we used the same dummy observations as for the other three models, plus dummy observations defining an inverse-Wishart prior on the reduced form covariance matrix of residuals. Since we are comparing models, and formulating a prior with dummy observations, it is important that we normalized the prior so that it integrated to one in each case.

worse. That it does reflects the fact that Bayesian posterior odds tend to penalize models with larger numbers of free parameters, and this model has many more free parameters.[16]

Ideally, we could have calculated MDDs for models like the first and fourth line models, but with *t*-distributed errors. But calculation of MDDs for models with *t* errors is extremely time consuming, and the results would probably be similar to what we see in this table.[17]

See online Appendix Section III.4 for more details on these models and the MDD calculations.

III. The Structural Shocks

This section begins with an overview of the impulse responses. The estimation just separates them, without giving them interpretations, so as we discuss them we try to map the shocks into economic interpretations. Then we consider the estimates of time-varying variance. We finish with a closer look at financial stress build-up and credit growth. Financial stress, as indexed by rate spreads, is occasionally (but not always) an important driver of IP fluctuations while the shocks originating in credit growth are much less important.[18]

A. *The Big Picture*

Figures 1 to 4, in four 5-by-5 "blocks," show the impulse response over five years of all 10 variables to the model's orthogonal structural shocks, scaled to draws from a unit-scale *t* distribution with 5.7 degrees of freedom. Since the diagonal of Λ_i is normalized to sum to 1 across regimes, these responses are a kind of average across regimes.

In principle structural VAR shocks need not be associated with single variables in the system. However in these monthly data, the reduced form shocks, which are variable-by-variable innovations, show only modest cross-variable correlations. Since our prior mildly favors A_0 matrices with positive entries on the diagonal, the structural estimates still have shocks that are not far from being innovations in a single variable. Accordingly, with one exception, in the discussion and plots we name shocks by the variable they are most strongly associated with. The exception is the "interest rate shock," which has large contemporaneous effects not just on the short interest rate, but also on the term spread and the interbank (ES) spread.

---

[16] We also experimented with a model that allows arbitrary changes across regimes in A_0 while the rest of $A(L)$ remains fixed. This relaxes the main model's constraint that in every regime Σ_t has the same eigenvectors. This model also fits less well than our main model. The model and its MDD are discussed in online Appendix Sections VII.1 and III.4.

[17] There is also a conceptual problem in applying the *t*-errors specification to models in which structural shocks are not identified. In the SVAR, the residuals are independent univariate *t*s. In the reduced form, the residuals are linear combinations of these structural residuals, which are *not* distributed as univariate *t*. If we took the reduced-form residuals as multivariate *t*, the implied SVAR residuals would be univariate *t*, but would no longer be independent.

[18] This process also resembles at a high level, though differs greatly in the technical implementation, with recent suggestions by Antolín-Díaz and Rubio-Ramírez (2018) and Ludvigson, Ma, and Ng (2017) to identify structural shocks ex post based on matching known events in history.
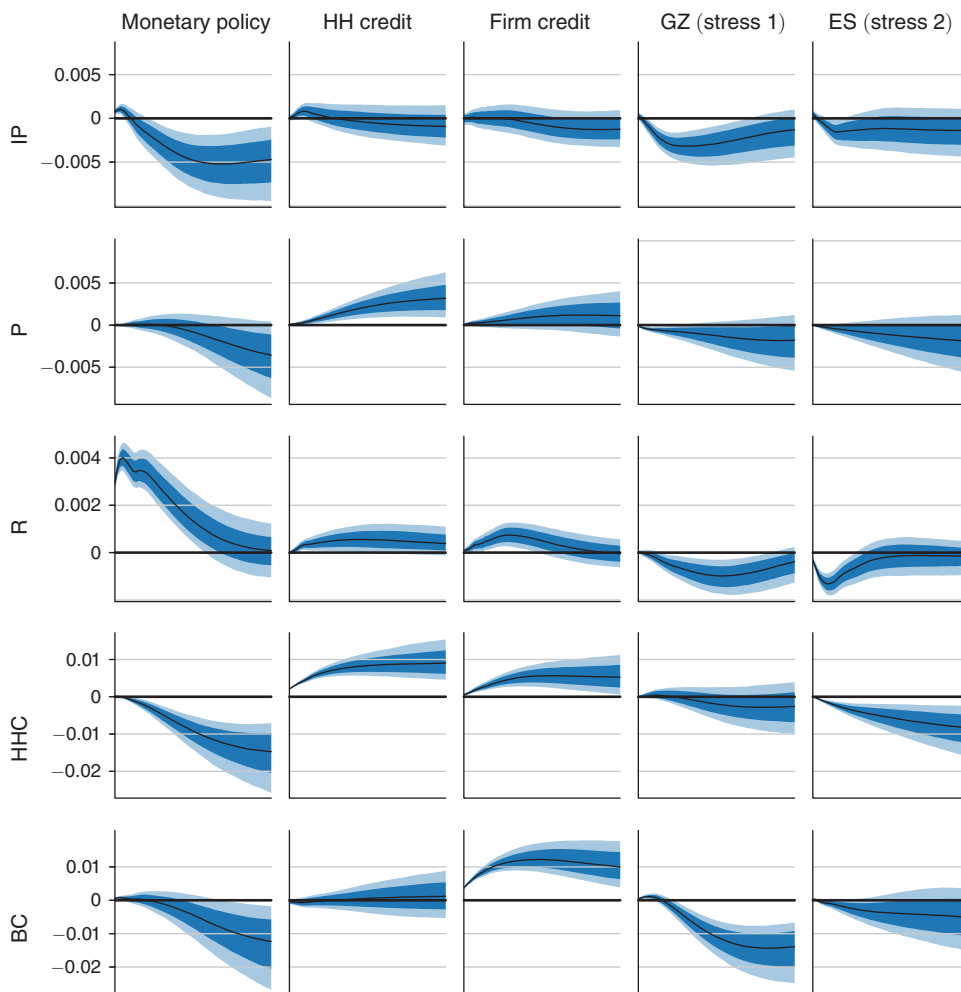
FIGURE 1. IMPULSE RESPONSE FUNCTIONS (1/4)

Notes: Impulse responses to the 10 orthogonal structural shocks in the model with *t* distributed errors over 60 months, with 68 percent (dark blue) and 90 percent (light blue) posterior uncertainty regions. Scaled to an "average" period with unit scale.

We name this shock the "monetary policy" shock. Despite the model's lack of any identifying zero restrictions on coefficients, this shock emerges as having the characteristics usually associated with monetary policy shocks. It is the only one that has an immediate positive R response, a delayed negative IP response, a negative (though ill-determined) long run P response, negative responses of 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 contemporaneous movement of credit spreads, and possibly policy amplification through these channels, corroborates the general findings of Gertler and Karadi (2015) and Caldara and Herbst (2019).

Our median estimated response of IP to the contractionary monetary policy shock is somewhat stronger and more persistent than most estimates in the literature. We
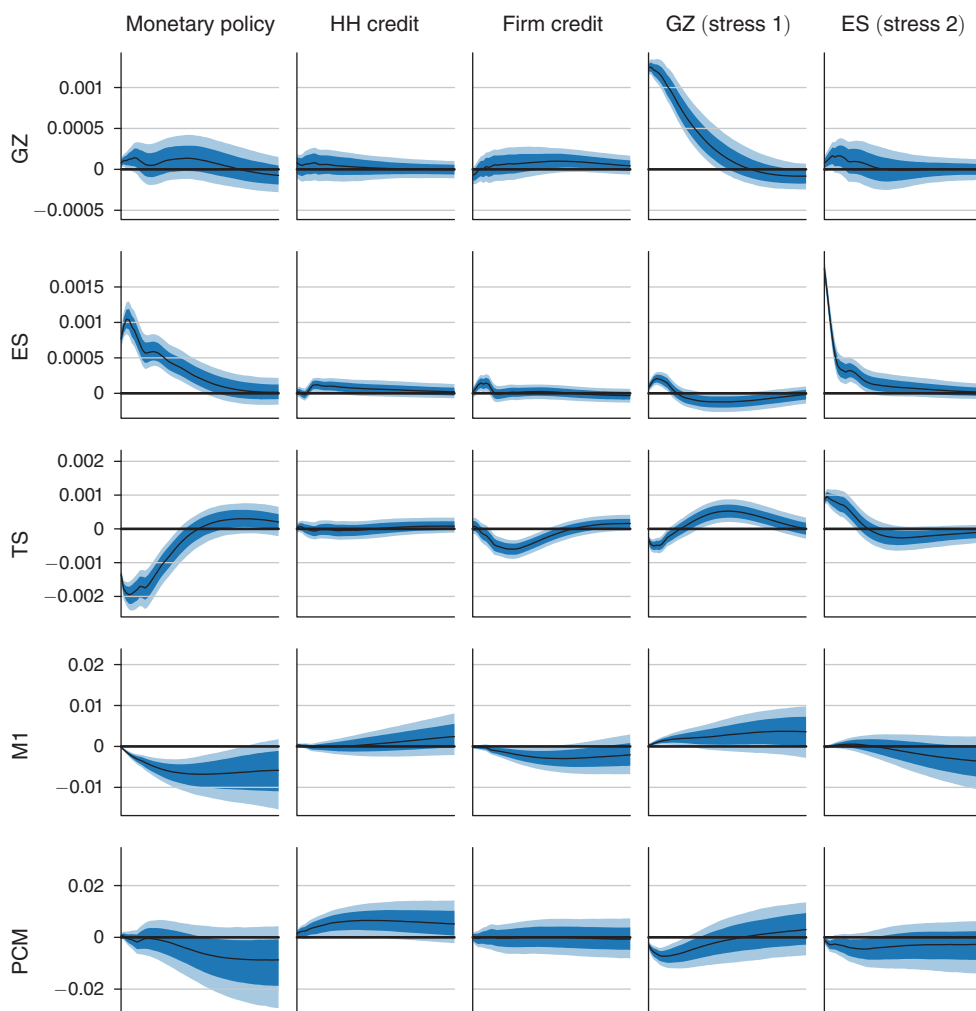
FIGURE 2. IMPULSE RESPONSE FUNCTIONS (2/4)

Notes: Impulse responses to the 10 orthogonal structural shocks in the model with t distributed errors over 60 months, with 68 percent (dark blue) and 90 percent (light blue) posterior uncertainty regions. Scaled to an "average" period with unit scale.

think this could reflect better identification of the shock in our model. By including the ES spread variable, we allow the model to capture the increased banking system financial stress that arises immediately when monetary policy is tightened. We include M1 in the model, which was definitely a component of the monetary policy reaction function in the 1970s, yet is often omitted in empirical models of monetary policy. And we include the term spread, which allows the model to use the fact that a short rate rise induced by policy contraction should be accompanied by a much smaller rise in long rates.[19]

---

[19] Contemporaneous work by Jordà, Singh, and Taylor (2020), using very different methods and a cross-country panel, also find very large and persistent effects of monetary policy which contrast to findings from conventional models.
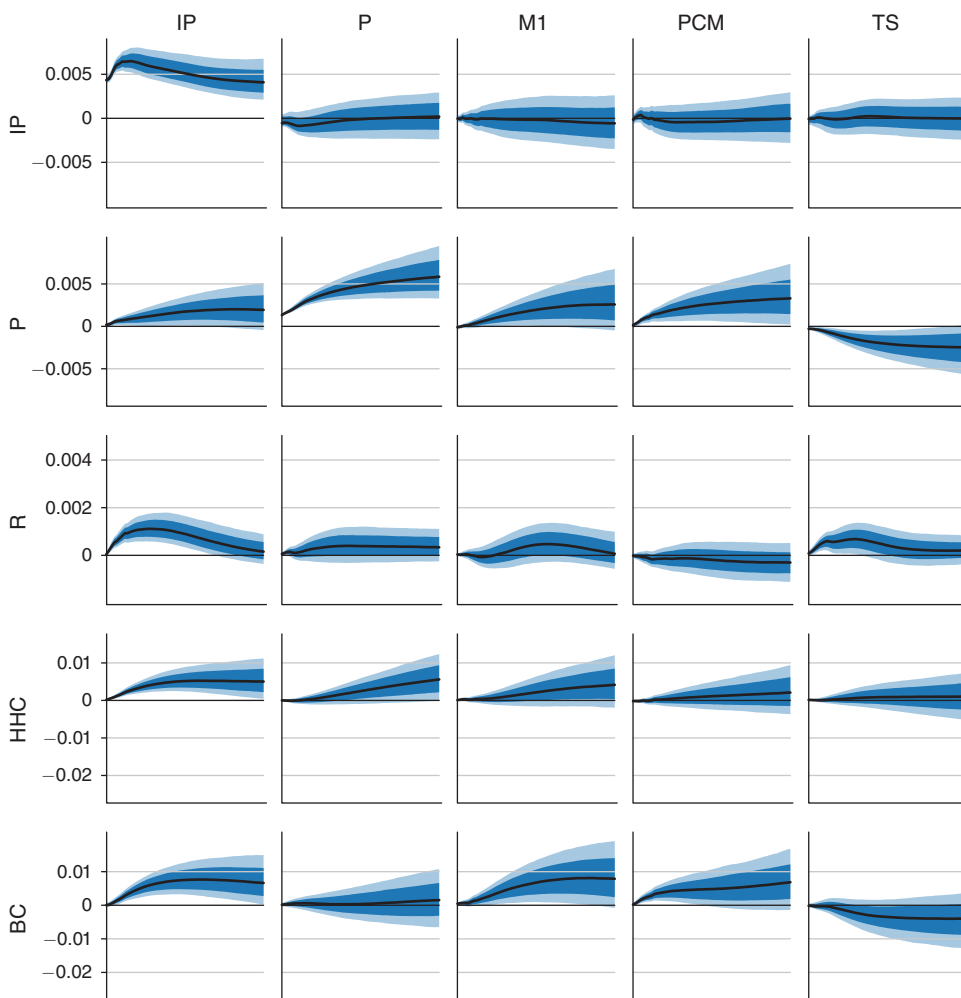
FIGURE 3. IMPULSE RESPONSE FUNCTIONS (3/4)

Notes: Impulse responses to the 10 orthogonal structural shocks in the model with *t* distributed errors over 60 months, with 68 percent (dark blue) and 90 percent (light blue) posterior uncertainty regions. Scaled to an "average" period with unit scale.

The two spread shocks are the most important sources of variation in the GZ spread and ES spreads. The two spreads do not tend to move together in response to these shocks, and the two have different patterns of effects on other variables. Both depress IP. Both depress P, though in the case of the ES shock this effect is statistically weak. The shock we label GZ has a strong delayed effect in depressing BC, but modest and indeterminate-signed effect on HHC, while the shock we label ES, which immediately impacts ES, strongly depresses HHC with ill-determined effect on BC. The ES shock produces a quick and strong expansionary movement in R, while the GZ shock is followed by a smoother, more delayed response of R. These patterns seem to fit an interpretation that distinguishes a banking credit shock (ES) from a non-bank financial disturbance (GZ). All the effects of these shocks on other variables are delayed, while their effects on the spread variables
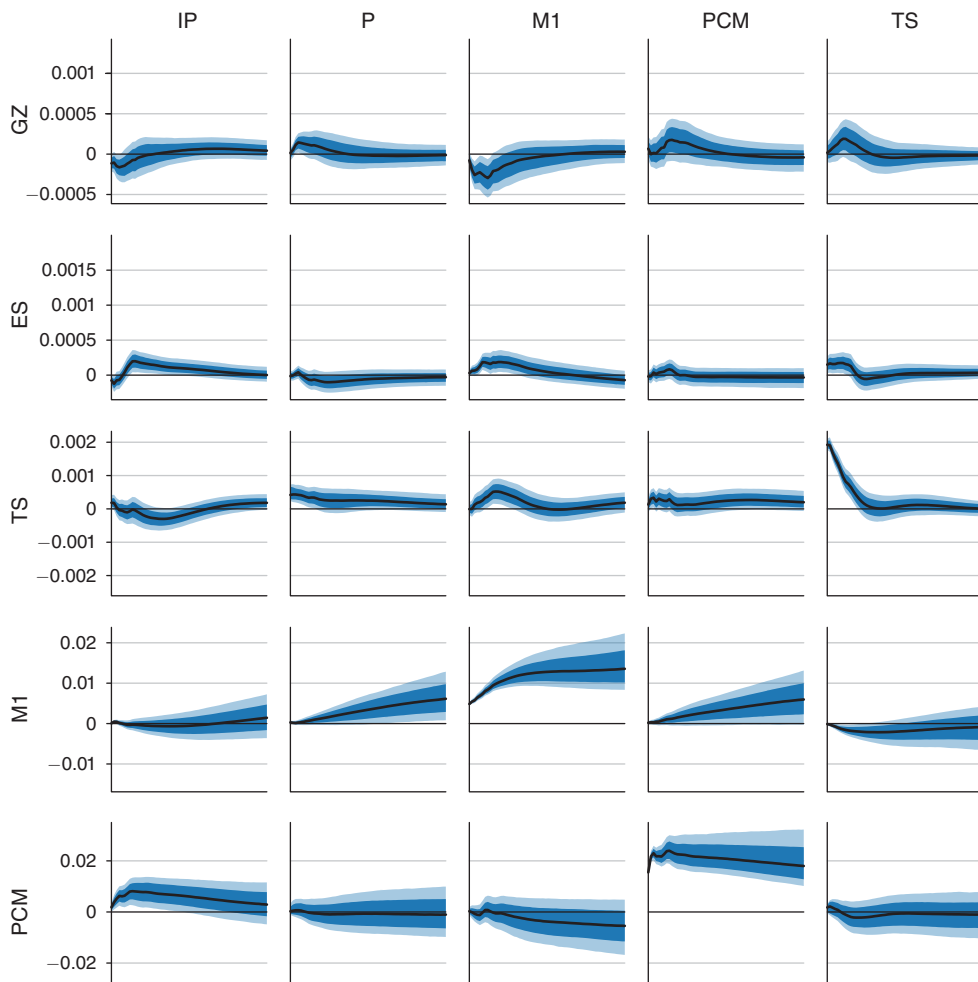
FIGURE 4. IMPULSE RESPONSE FUNCTIONS (4/4)

Notes: Impulse responses to the 10 orthogonal structural shocks in the model with *t* distributed errors over 60 months, with 68 percent (dark blue) and 90 percent (light blue) posterior uncertainty regions. Scaled to an "average" period with unit scale.

are immediate. This fits an interpretation that they reflect disturbances originating in financial markets, with monetary policy trying to partially offset their effects.

The shocks we label HH and firm credit start with an impulse to household credit and to business credit (net of inflation), respectively. They lead to persistent, but marginally statistically significant, long-term declines in output. They qualitatively match the "excessive credit growth" story demonstrated empirically by Mian, Sufi, and Verner (2017); Schularick and Taylor (2012); and others, but might seem small in the light of that earlier literature. We discuss the interpretation of these shocks and their comparison to earlier literature in more detail in Section IIIE.

The PCM shock accounts for both the main component of variation in PCM and a substantial component of the variation in P. Its immediate effect is to increase

commodity prices. The effect on commodity prices is persistent. With some delay, P (the PCE deflator) moves up. BC, but not HHC moves up slightly. This looks like a commodity supply shock.[20]

These core impulse response results seem largely robust to the alternative error specifications. We provide full impulse response plots for versions of the model with Gaussian errors in online Appendix Section VII.2.

The picture is also very similar if we estimate with data only up to December 2007. Full impulse responses from this sample period are plotted in online Appendix Section VII.3. In particular, the identification of monetary policy and spread effects is very stable. There is weak (within 68 percent, but not 90 percent bands) evidence of an output response to household and business credit expansion shocks. These effects are of comparable magnitude to the estimated credit effects in the model estimated on the full dataset.

B. *Structural Shock Volatility*

Our results suggest that the variances of these shocks change substantially among periods. Table 4 reports the variances of each of the ten structural shocks in the posterior mode *t*-errors model, over the full sample. Ninety percent 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).

Table 5 lists the four largest posterior median shocks for each equation. These are in standard deviation units and not scaled by the corresponding λ values. They thus show the biggest shocks, not the biggest "surprises" for the model.

The biggest of these, in the monetary policy shock, reflects a sudden easing of monetary policy in May 1980, during the recession of that year. The Federal Funds rate fell from 18 percent to 11 percent in that month, but soon started rising again. There were also large values for this shock in March 1980, February 1981, and May 1981. These were all during the period of Volcker unborrowed reserve targeting, which has a high value of λ for this shock. That is, though large, these shocks occur in a period the model has identified as a high-variance period.

The second and third largest shocks were in the two financial stress indicators, at the time of the Lehman collapse in October 2008. The IP shock, which accounts for much of the variance in output and looks like a "demand" shock, was sharply negative in September 2008, reflecting the large decline in industrial production as the crisis took hold. The September 2001 attack on the United States shows up in the M1 shock, which looks like an accommodated money demand shock. The shock is sharply negative in October 2001 as the Fed withdrew its temporary liquidity accommodation. It was nearly as large and positive in the previous month. The PCM shock takes on large values in 1974–1975, near the height of the oil crisis.

---

[20] The Gaussian-errors version of the model shows a modest, marginally statistically significant, negative IP response to this shock. See online Appendix Figure 7.

TABLE 4—POSTERIOR MEDIAN RELATIVE VARIANCES

Shock	Jan 1973– Sep 1979	Oct 1979– Dec 1982	Jan 1983– Dec 1989	Jan 1990– Dec 2007	Jan 2008– Dec 2010	Jan 2011– Jun 2015
IP	1.255	1.336	0.767	0.807	1.306	0.610
P	0.856	0.946	0.902	0.875	1.190	1.119
HH credit	0.676	0.533	0.413	1.700	1.675	0.934
Firm credit	0.814	0.950	1.179	1.165	1.241	0.622
M1	0.319	0.645	0.482	0.595	2.190	1.721
Monetary policy	1.031	4.147	0.585	0.109	0.087	0.004
PCM	1.275	0.832	0.760	0.722	1.729	0.575
TS	0.940	2.553	0.864	0.521	0.781	0.375
GZ (stress 1)	0.764	0.484	0.533	0.717	2.823	0.632
ES (stress 2)	1.565	1.917	0.543	0.286	1.707	0.009

Note: The estimates are posterior medians from the model with *t*-distributed innovations.

TABLE 5—LARGEST RESIDUALS FOR EACH SHOCK

Month	ϵ_{it}	dy_{it}	Month	ϵ_{it}	dy_{it}
IP			**Monetary policy**		
9/2008	−9.875	−0.043	5/1980	−22.627	−0.067
12/1974	−6.259	−0.027	3/1980	10.409	0.031
1/1983	5.160	0.022	2/1981	−10.135	−0.030
4/1980	−5.015	−0.022	5/1981	9.312	0.028
P			**PCM**		
9/2005	5.492	0.007	8/1976	−4.222	−0.066
1/1990	3.820	0.005	7/1975	4.096	0.064
11/2005	−3.748	−0.005	7/1974	3.901	0.060
11/2008	−3.453	−0.005	2/1974	3.777	0.059
HHC			**TS**		
12/1999	7.068	0.015	8/1974	−7.157	−0.014
9/2008	5.922	0.012	6/1981	6.261	0.012
10/2003	−5.861	−0.012	11/2008	5.871	0.011
10/2002	4.933	0.010	4/1980	4.411	0.008
BC			**GZ**		
10/2008	5.247	0.019	10/2008	18.275	0.023
12/1986	5.216	0.019	1/2009	−7.844	−0.010
9/2007	4.454	0.016	9/2008	7.183	0.009
11/1999	4.283	0.016	7/2002	6.771	0.009
M1			**ES**		
10/2001	−9.766	−0.047	10/2008	10.840	0.019
8/2011	8.507	0.041	7/1974	8.950	0.016
9/2001	8.431	0.041	9/2008	7.739	0.014
9/2008	6.773	0.033	12/1973	7.148	0.013

Notes: The estimates are posterior medians from the model with t-distributed innovations. The first column is the size of the shock and the second is the contemporaneous impact on the "diagonal" variable.

The largest of these shocks all correspond to events that were recognizably unusual as they occurred. But they tend to come from periods with large values of λ_{it}, so they are not necessarily the biggest surprises. Looking at the largest surprises, the $\epsilon_{it}/\sqrt{\lambda_{it}}$ values, is useful because it is these residuals that have the biggest impact on model fit. Also the size and distribution across variables of these large surprises casts some doubt on our assumption of i.i.d. *t*-distributed scaled shocks. We discuss the surprises in online Appendix Section VI.

C. *External Identifying Evidence*

Online Appendix Table 5 shows the correlation with our estimated shocks of several variables that have been used by others as external instruments for monetary policy. In general, when the sample periods do overlap, the high-frequency instruments agree with our identification. They show the instruments highly correlated with our estimated monetary policy shocks, and not highly correlated with our other estimated shocks. This suggests that if we followed the methodology of Gertler and Karadi (2015) and Caldara and Herbst (2016) to identify a monetary policy shock, we would get back something quite similar to what we have found with identification through heteroskedasticity.

D. *Spread Spikes and Early Warning*

Three independent shocks can be identified by sharp increases in spreads at $t = 0$, but only the ES and GZ shocks (the responses to which are plotted in Figure 5) have significant output effects. The output effect is larger and more significant for the bond spread (GZ) shock, associated with an initial surge in the corporate bond spread and a long-term contraction in business credit. The ES shock, in contrast, begins with a shock to the interbank lending rate of comparable magnitude to the impulse following a monetary policy shock, a significant long-term contraction in household credit, and a modestly significant short-term output contraction. It also shows a quicker and stronger R decline following the shock, implying a stronger monetary loosening, in response to the generated output decline. The fitting of two independent "stress shocks," with quantitatively different macro effects, suggests the importance of a multidimensional approach to measuring financial stress.

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 4). 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. Interest rates react to the corporate and inter-bank spread shocks by falling, as would be expected from a monetary policy easing. Because the response of IP to monetary policy is slower and more persistent than its response to these spread shocks, it is not clear that monetary policy could do more to mitigate the output decline without creating more instability.[21]

We will return to discussion of the contribution of financial variables to forecasting in Section V, where we show results of pseudo-out-of-sample forecasting.

---

[21] Of course this presumes the rest of the system remains stable as the monetary policy rule shifts. In principle, we expect that eventually a systematic change in the policy rule would result in changes in the rest of the system.
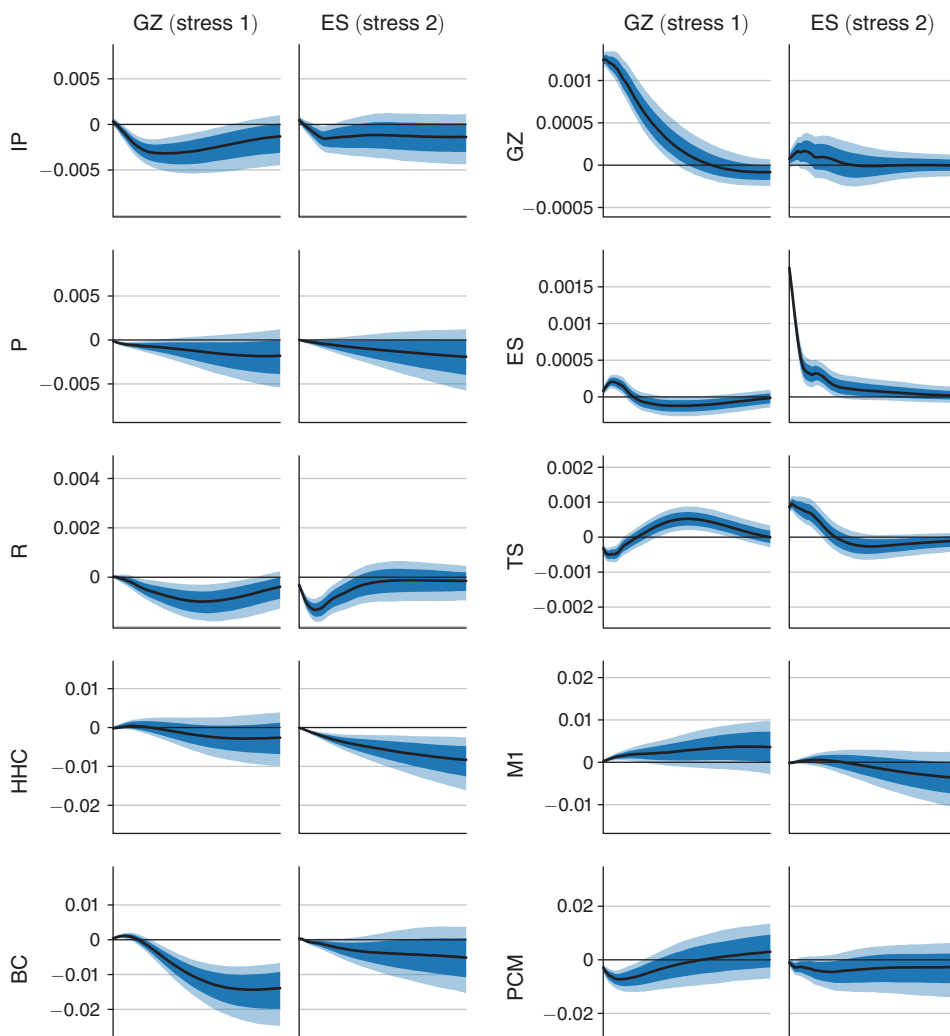
FIGURE 5. FINANCIAL STRESS SHOCKS

Note: Impulse responses to the bond spread (GZ) and inter-bank spread (ES) shocks, with 68 percent (dark blue) and 90 percent (light blue) posterior uncertainty regions.

E. *Credit Growth and Recessions*

Our main model offers some support, within 68 percent error bands, of the hypothesis that excessive growth in household credit can forecast negative long-term real output growth (Figure 6). The shape of our estimated output responses to the two credit shocks HHC and BC are similar to those found in a small-system (household credit to GDP, business credit to GDP, and real GDP) VAR by Mian, Sufi, and Verner (2017). In our model the response to the HHC shock, which raises HHC by about 1 percent over five years, is a small increase in IP, on the order of 0.1 percent, that lasts less than a year, followed by a decline that reaches −0.1 percent after
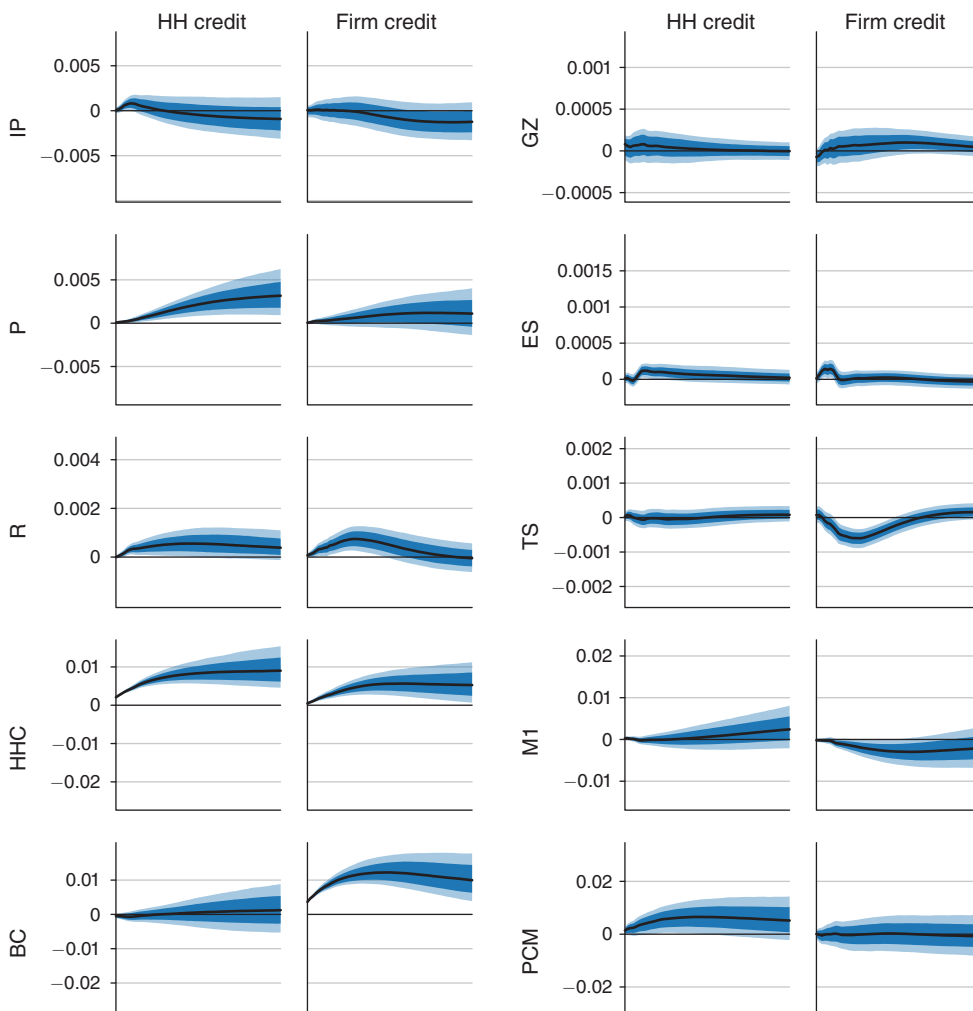
FIGURE 6. CREDIT AGGREGATE SHOCKS

Notes: Impulse responses to the household credit (HHC) and firm credit (BC) shocks, with 68 percent (dark blue) and 90 percent (light blue) posterior uncertainty regions. Scaled to an "average" period with unit scale.

five years. Mian, Sufi and Verner (forthcoming) find, as a response to a shock that raises the ratio of household credit to GDP by 1.6 percent over two years, an initial rise in GDP of about 0.2 percent, over two years, but also a subsequent decline that reaches about −0.1 percent after five years. Because they ordered log GDP first in their triangularly orthogonalized VAR, the innovation in their credit-to-GDP ratio is in fact an innovation in credit itself, holding GDP constant. Because the ratio of household credit to GDP in the United States around 2014 was roughly 0.8, their 1.6 percent rise after two years in that ratio corresponds to a rise of 2 percent over that span in household credit itself. Thus their impulse response is to a larger shock than ours, and their initial expansion in GDP is larger by the same factor. When their shock is rescaled to make the shock sizes match, the effects on IP (in our case) or GDP (in their case) match quite closely.

Our estimates show a five-year negative response to the BC shock that is similar in size and statistical significance to the estimated response to HHC, though without the initial positive response. This also is consistent with the Mian, Sufi, and Verner (2017) business credit estimated responses. All these results have fairly wide error bands, so there is no statistically sharp difference between the models. This might not have been expected, since these other authors use an international panel of annual data and a smaller model with no structural identification, while we use monthly US data, a big model, and separate out a structural shock.

However, our model implies that the decline in output growth following this shock can be entirely accounted for by the rise in interest rates it elicits. The response of the system to the credit shocks, combined with a sequence of monetary policy shock values that keep the interest rate constant, eliminates the decline in output (Figure 7). Of course this does not imply that the monetary contraction following a credit aggregate shock is a mistake. The credit shocks are followed by an increase in inflation; without the monetary policy reaction, the inflation would be larger and more persistent.

Though the credit shocks are followed by future declines in output growth, they are not followed by substantial movements in the spread variables GZ and ES. The spread shocks, on the other hand, are followed by substantial declines in credit aggregates. Our interpretation is that the credit expansions generated by the credit shocks are followed with a delay by slow growth due to monetary tightening, not financial market distress.

A different way to assess the economic significance of credit shocks in our model is to calculate the share of forecast error variances that are explained by each shock (Figure 8).[22] The importance of what we label as the HHC and BC shocks for explaining credit variation starts very high (as they have by far the largest contemporaneous impact on credit) but decays over time. Over five years, these shocks explain 20 percent and 24 percent, respectively, of forecast error variance in household and business credit. The remainder of credit variation over these horizons is explained by the other shocks in the model, which are all associated with credit and output moving in the same direction, or with output scarcely moving at all. Over five years, the same "bad credit shocks" explain 1.4 percent and 1.6 percent, respectively, of forecast error variance in industrial production.

In the period 1990–2007, when the variance of the credit shocks is relatively high and the variance of others relatively low, the two credit shocks explain a higher fraction of five-year-ahead variance in the two aggregates (49.5 percent for household credit and shock 3, and 37.6 percent for business credit and shock 4) and a higher, though still small, percentage of output variation at the same horizon (4.1 percent and 3.2 percent, respectively).

---

[22] These are the squared impulse responses scaled to sum to 1 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*.
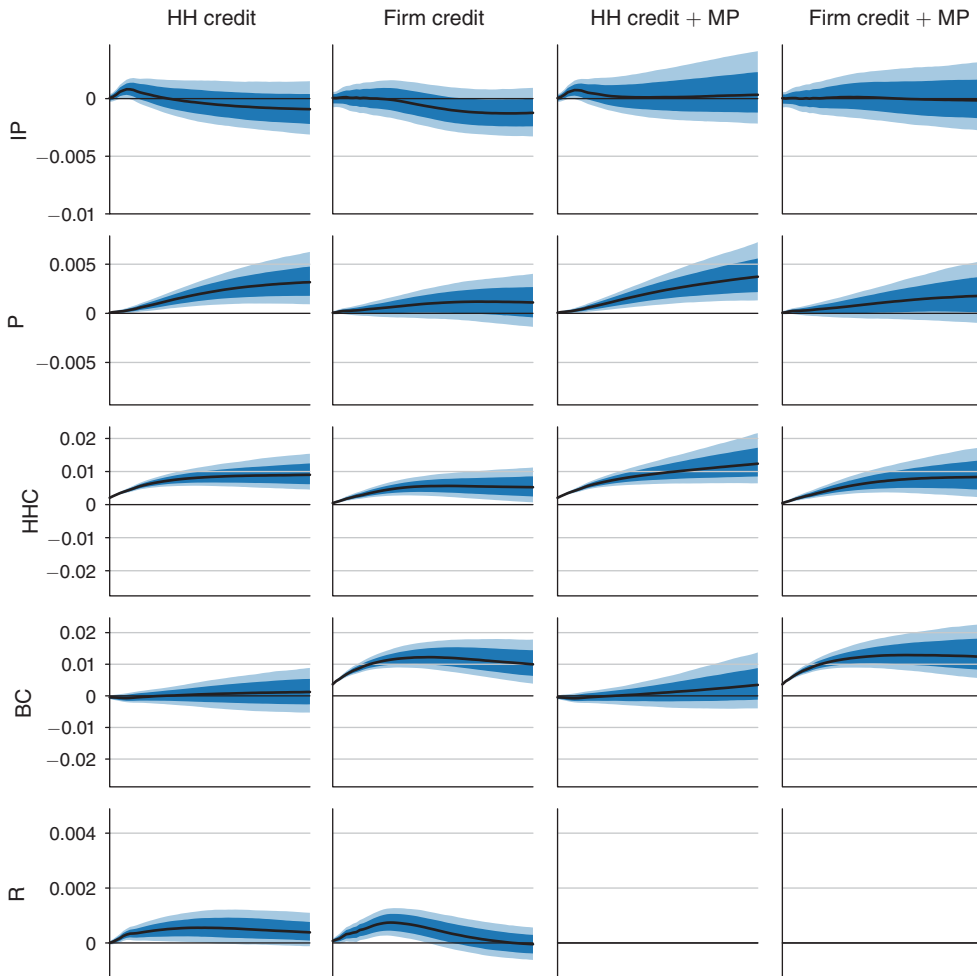
FIGURE 7. CREDIT SHOCKS, WITH AND WITHOUT MONETARY RESPONSE

Notes: The first two columns reproduce the responses to the HHC and BC shocks in Figures 1 and 2. The third and fourth columns combine each shock with a series of shocks to the monetary policy shock, that keep R constant.

IV. Credit Aggregates in Smaller Models

There is a recent literature that claims to show that credit expansion predicts negative growth and/or financial crises. Our model does not suggest that credit expansion predicts financial crises and shows the negative association of credit growth with future output growth as a small component of the overall relation between credit and output. But our model does not contradict the earlier results with smaller models. It is compatible with them precisely because it endogenizes important omitted variables including interest rates.

As we have already observed, the responses to shocks the HHC and BC shocks in our model lie within error bands of the responses to credit-to-GDP-ratio shocks in Mian, Sufi, and Verner (2017)—henceforth, MSV. Those authors also estimate
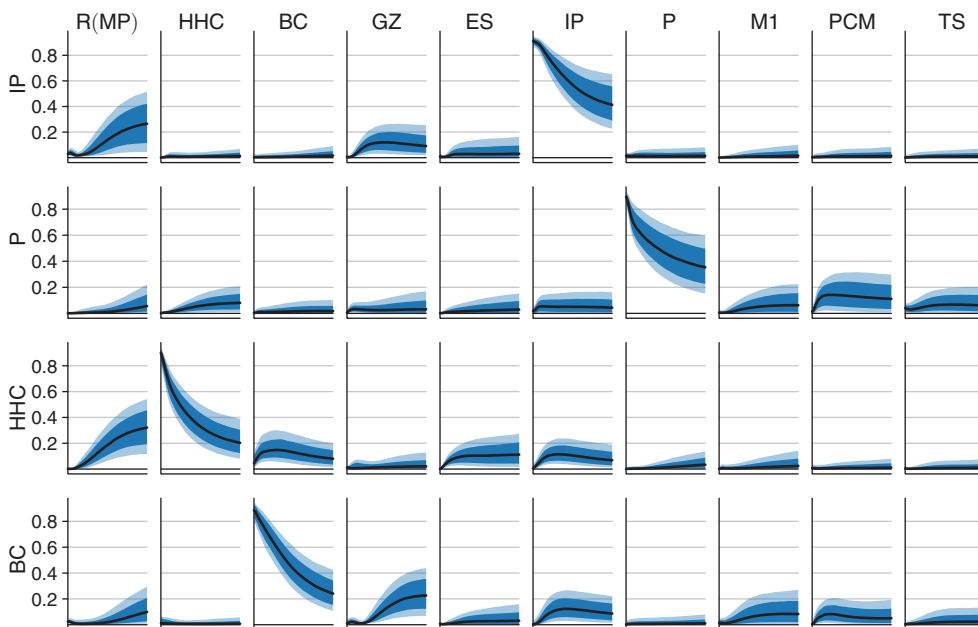
FIGURE 8. FORECAST ERROR VARIANCE DECOMPOSITION

Notes: Forecast variance decompositions in the *t*-distributed errors model for IP, prices, household credit, and business credit over 60 months, with 68 percent (dark blue) and 90 percent (light blue) posterior uncertainty regions. Scaled to an "average" period with unit scale. Variables are in the order listed in Table 1.

single-equation regressions of growth in real output over the next three years on growth in credit in the past three years, finding that a one-standard-deviation increase in 3-year household credit growth predicts a 2.1 percent lower growth rate of output in the subsequent three years. They find no such effect with business credit. When we replicate those regressions using our data, we find that a one standard deviation increase in the three-year growth rate of BC predicts a 1.7 percentage point decline in the subsequent three year growth rate of output. We find no such predicted effect from increased household credit growth. When we substitute the flow of funds data on business credit and household credit that MSV used for our HHC and BC, the effects of HHC and BC increases by 1 standard deviation are both negative, with HHC producing a 1.6 percentage point decline and business credit a 0.6 percentage point decline. The main difference between our results and theirs in this single-equation exercise seems therefore to be in the differing definitions of the credit variables.

Though the two types of data give different answers as to which type of credit growth predicts low future output growth, they agree that there is at least one type of credit growth that predicts lower future output growth. Note, though, that the VAR results, both MSV's 3-variable VAR with quarterly data and our full 10-variable model, imply that household credit growth shocks predict a rise in output, followed by a return over five years to trend line or slightly below. In other words, they predict growth rates lower three years in the future, but higher in the immediate

TABLE 6 – SIMULATION TEST FOR PREDICTIVE RELATIONSHIPS

	Household credit	Business credit
Proportion of negative effects		
Without lagged IP	0.55	0.56
With lagged IP	0.53	0.57
Proportion of economically significant negative effects		
Without lagged IP	0.29	0.31
With lagged IP	0.30	0.33

Notes: Probability, in simulated draws of coefficients and data from the posterior distribution under the *t*-distributed errors model, of a negative or "economically significant" effect of previous 3-year credit growth on subsequent three-year output growth. An economically significant effect is one that makes a one-standard-deviation change in credit growth produce a 2 percentage point or larger change in output growth rate.

future, leaving it unclear whether there is any net, persistent negative effect on the level of output.

The contrast between our data, in which business credit but not household credit predicts future output, and the MSV data, in which the reverse is true, is not statistically sharp, according to our model. We replicate the single-equation growth rate regressions on data simulated from our model, and we find that with substantial probability our results from simulated data are of similar size and sign to those found in the actual data, and that the model is about as likely to generate predictive value for HHC as for BC.[23] The results are in Table 6.

To form the table we used three-year forward differences of the log of real output (IP) as the dependent variable, three-year backward differences in real credit (credit over price level) over output as the main independent variable, and lags of first differences of log output as an extra independent variable. In equation form, with y_t denoting log IP, hc_t the ratio of household credit deflated by PCEPI to IP, and bc_t the ratio of business credit deflated by PCEPI to IP,

$$y_{t+3} - y_t = \alpha + \beta_h(hc_t - hc_{t-3}) + \beta_b(bc_t - bc_{t-3}) + \sum_{i=1}^{k}\gamma_i(y_{t-i+1} - y_{t-i}).$$

Table 6 reports probabilities for β_h and β_b. The results "with lagged IP" set $k = 3$; otherwise we set all $\gamma_i \equiv 0$.

In these simulated draws from our model's data distribution we find the probability of credit growth coefficients smaller than 0 to be over 0.5 and negative enough that a 1 standard deviation increase in credit growth reduces output growth by 2 percentage points or more with probability about 0.3. Note that in the simulated data, it is about equally likely that the BC or the HHC predictive effect will be large and negative.

---

[23] The simulated data take into account uncertainty from the model parameters and the identity of the model innovations. We sample from our saved posterior draws for the coefficients, and for each draw, we simulate monthly frequency data starting from the observed initial conditions and using structural innovations from the appropriate *t* distribution. We next transform the data to the annual frequency. And we finally apply the log or ratio transformations needed to match the earlier estimates, as described below.

V. 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.

A. *Forecasting in the Recent Financial Crisis*

We first focus on the 2007–2008 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 9, 10, and 11 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 11) 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 interbank 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 spread variables alone, with them and also 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.

B. *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
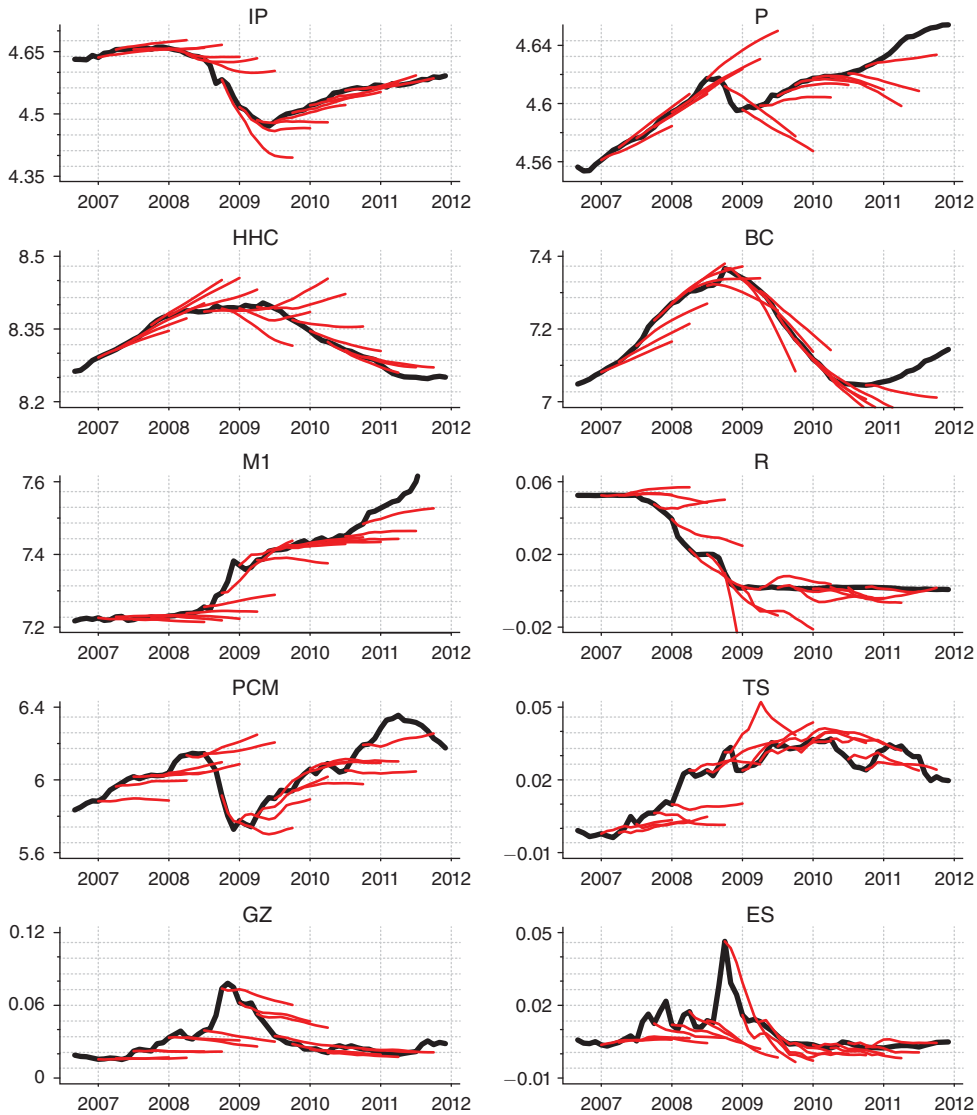
FIGURE 9. POSTERIOR MODE OUT-OF-SAMPLE FORECASTS, FROM A MODEL WITH ALL TEN VARIABLE

up to each month from October 1979 to June 2015.[24] We focus on root mean squared error (RMSE) for forecasts of all variables common to the models.

Figures 12 and 13 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

[24] The 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.
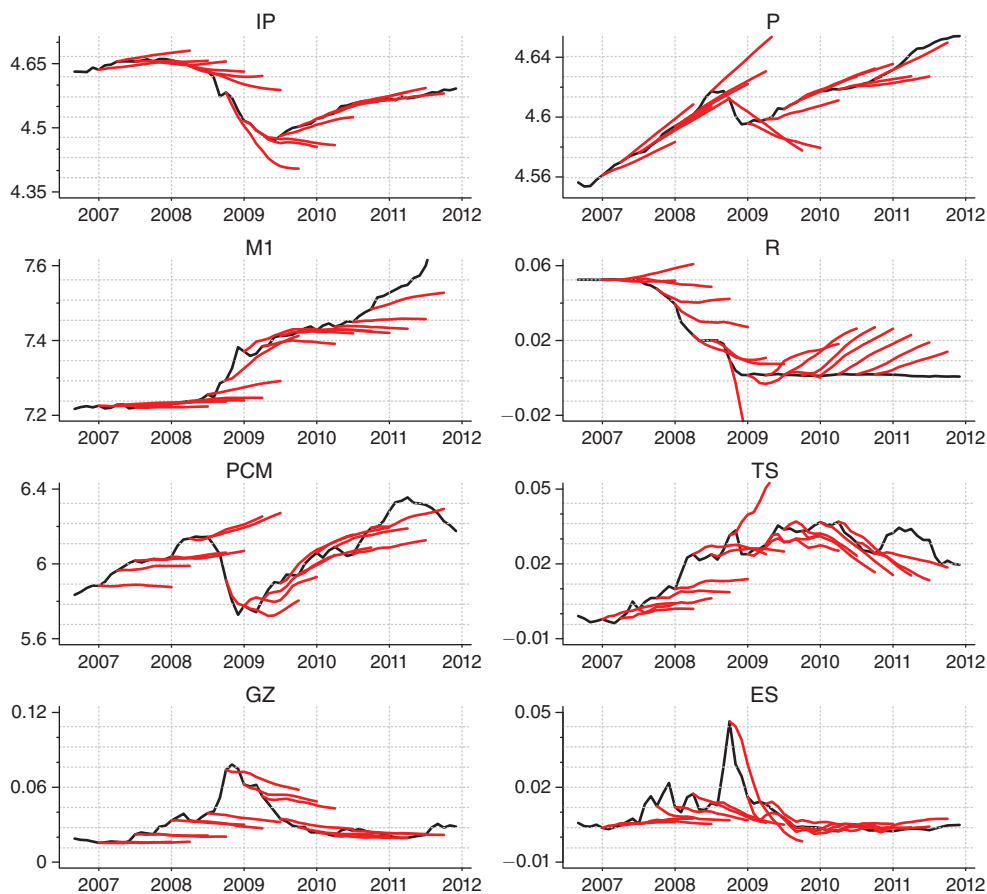
FIGURE 10. POSTERIOR MODE OUT-OF-SAMPLE FORECASTS, FROM A MODEL WITHOUT CREDIT AGGREGATES

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 better *at* the zero lower bound, though this comes at the cost of one set of very poor forecasts right around the major rate reduction 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.[25] 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. It could also be that the small fluctuations

---

[25] Del Negro and Schorfheide (2013) find a similar result, using a DSGE model with or without financial variables. And Stock and Watson (1989) much earlier found a spread variable to aid forecasting around a recession period, but not at other times.
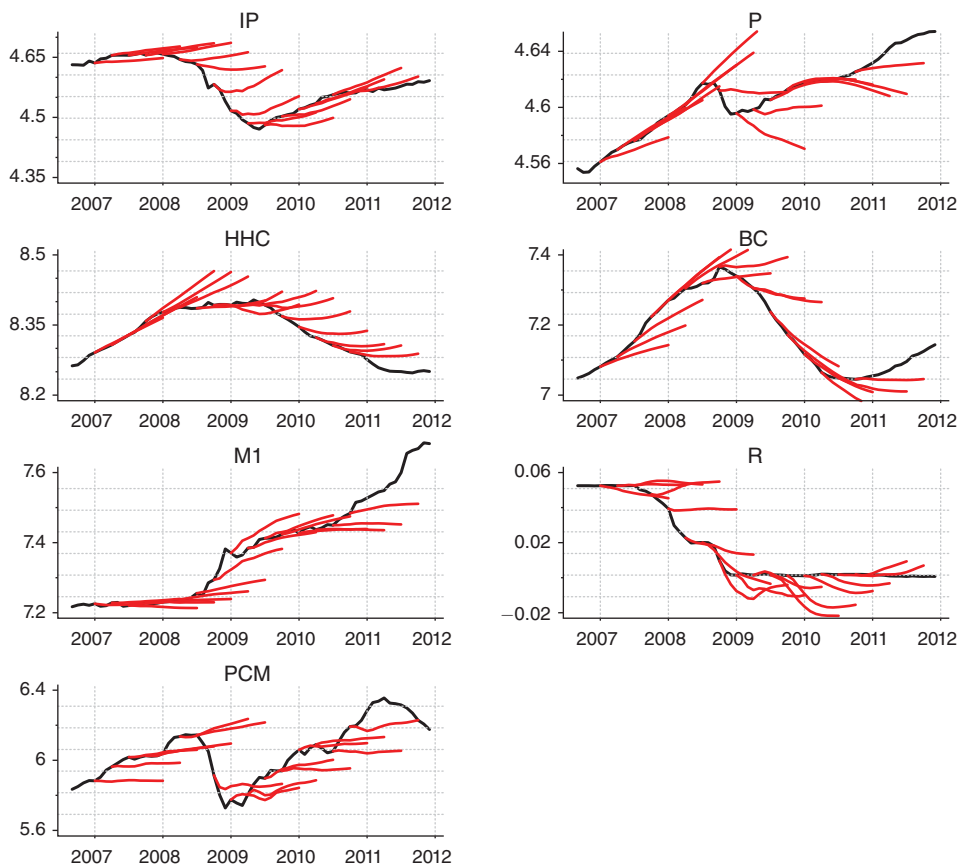
FIGURE 11. POSTERIOR MODE OUT-OF-SAMPLE FORECASTS, FROM A MODEL WITHOUT CREDIT SPREADS

in spreads in "normal times" are not useful for forecasting; only the large move-ments around recessions or crises matter.[26]

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 1990s 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 US recessions.

---

[26] These 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.
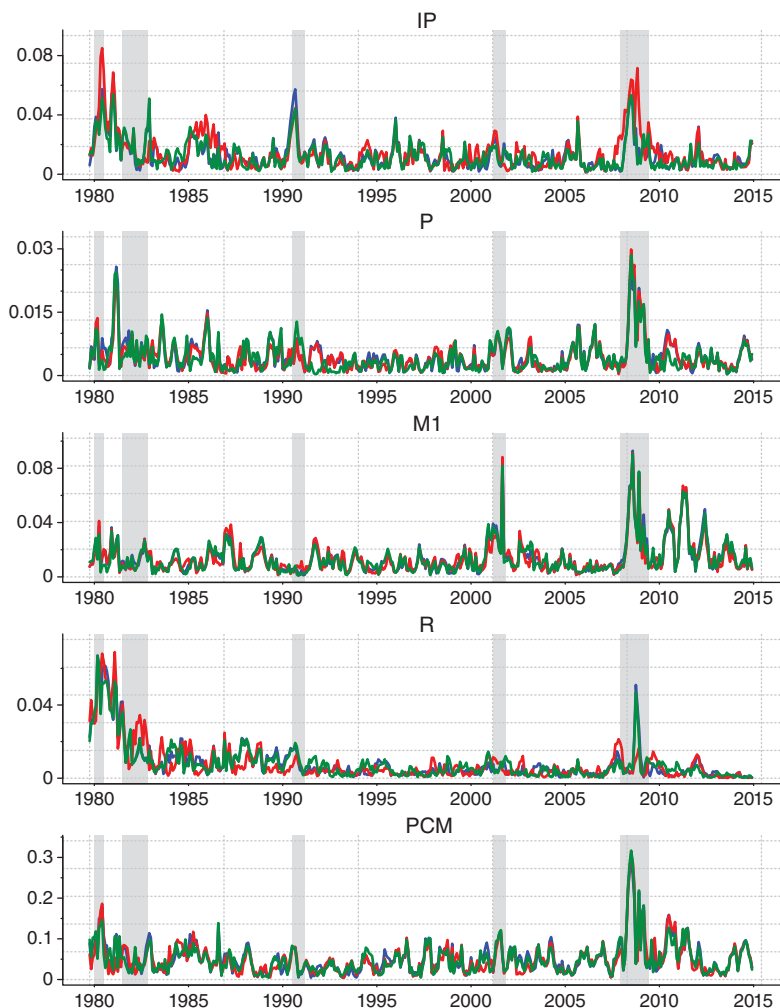
FIGURE 12. OUT-OF-SAMPLE RMSE, 6-MONTH FORECASTS

Notes: 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.

VI. Robustness

We have tried a number of variants on our model to check robustness of our results. These are described in detail in sections of the online Appendix. Our checks include varying the assumption on error distributions (Section VII.2), estimating models with triangular Cholesky identification (Section VII.4), looking for nonlinearity via various nonlinear transformations of the data (Section VII.6), and estimating quarterly models (Section IX). Of course the model we present as our main model is itself the result of experiments like this, where we have adopted specifications when we found them fitting better. As a result, the robustness checks in the online Appendix do not cast doubt on our main specification.
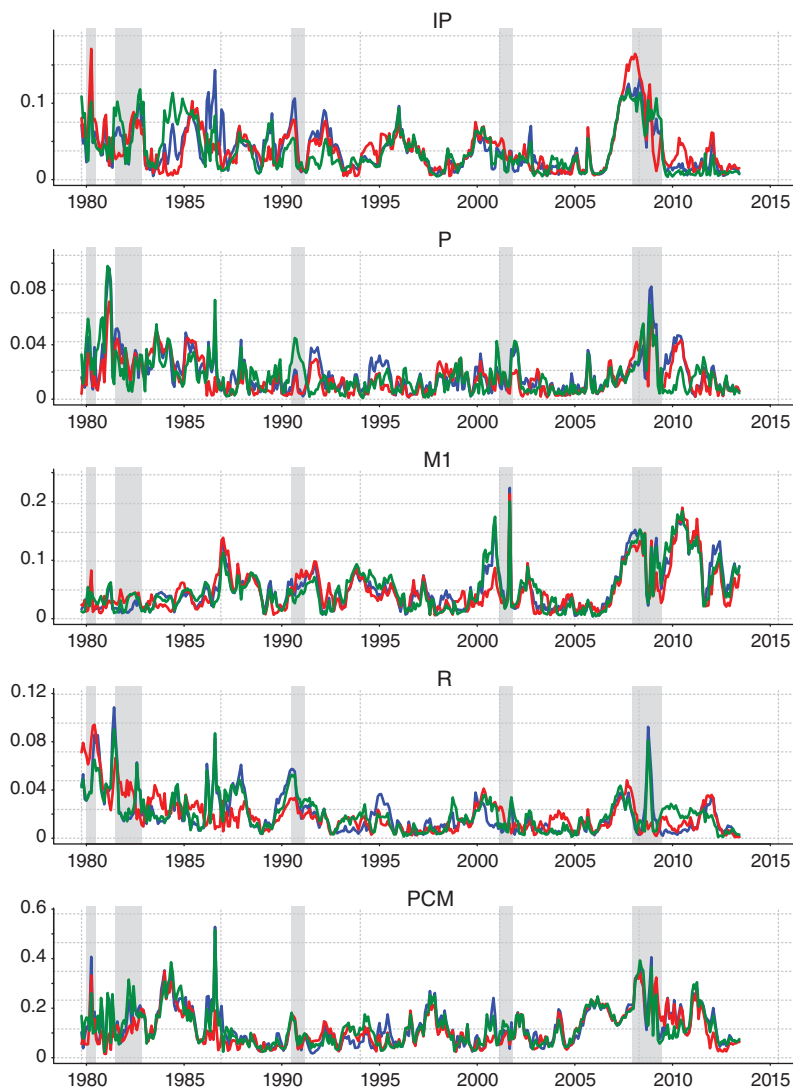
FIGURE 13. OUT-OF-SAMPLE RMSE, 24-MONTH FORECASTS

Notes: 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.

VII. 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 interbank 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 model mainly move "passively" in the same direction as output. Two shocks generate opposite movements in household (real estate plus consumer) or business credit and output, but in all periods the magnitude of these effects is relatively small. They are accompanied by rising interest rates, and if monetary policy offsets that rise, the effect of the credit shocks on output would disappear. To the extent that this effect is quantitatively important, a multivariate model is necessary to properly separate it from other effects.

The weakness in our model of the predictive value of credit aggregates for output growth contrasts with some results in the literature. While we argue that some of the apparent contrast stems from our use of a richer variable list and more careful identification, some of the contrast could arise from our use of data from a single country rather than several, or from our use of data over a shorter time span.

REFERENCES

Antolín-Díaz, Juan, and Juan F. Rubio-Ramírez. 2018. "Narrative Sign Restrictions for SVARs." *American Economic Review* 108 (10): 2802–29.

Bernanke, Ben S., Mark Gertler, and Simon Gilchrist. 1999. "The Financial Accelerator in a Quantitative Business Cycle Framework." In *Handbook of Macroeconomics,* Vol. 1C, edited by John B. Taylor and Michael Woodford, 1341–93. Amsterdam: Elsevier.

Borio, Claudio. 2012. "The Financial Cycle and Macroeconomics: What Have We Learnt?" Bank for International Settlements Working Paper 395.

Brunnermeier, Markus, Darius Palia, Karthik A. Sastry, and Christopher A. Sims. 2021. "Replication Data for: Feedbacks: Financial Markets and Economic Activity." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E121521V1.

Caldara, Dario, and Edward Herbst. 2019. "Monetary Policy, Real Activity, and Credit Spreads: Evidence from Bayesian Proxy SVARs." *American Economic Journal: Macroeconomics* 11 (1): 157–92.

Chib, Siddhartha, and Srikanth Ramamurthy. 2014. "DSGE Models with Student-*t* Errors." *Econometric Reviews* 33 (1–4): 152–71.

Chiu, Ching-Wai, Haroon Mumtaz, and Gabor Pinter. 2017. "Forecasting with VAR Models: Fat Tails and Stochastic Volatility." *International Journal of Forecasting* 33 (4): 1124–43.

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): 1031–52.

Del Negro, Marco, and Frank Schorfheide. 2013. "DSGE Model-Based Forecasting." In *Handbook of Economic Forecasting,* Vol. 2A, edited by Graham Elliott and Allan Timmermann, 57–140. Amsterdam: Elsevier.

Drehmann, Mathias, and Mikael Juselius. 2014. "Evaluating Early Warning Indicators of Banking Crises: Satisfying Policy Requirements." *International Journal of Forecasting* 30 (3): 759–80.

Gertler, Mark, and Peter Karadi. 2015. "Monetary Policy Surprises, Credit Costs, and Economic Activity." *American Economic Journal: Macroeconomics* 7 (1): 44–76.

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): 471–93.

Gilchrist, Simon, and Egon Zakrajšek. 2012a. "Credit Spreads and Business Cycle Fluctuations." *American Economic Review* 102 (4): 1692–720.

Gilchrist, Simon, and Egon Zakrajšek. 2012b. "Credit Spreads and Business Cycle Fluctuations: Dataset." *American Economic Review*. Inter-university Consortium for Political and Social Research, distributor.

Goldsmith, Raymond. 1969. *Financial Structure and Development*. New Haven, CT: Yale University Press.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor. 2013. "When Credit Bites Back." *Journal of Money, Credit, and Banking* 45: 3–28.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor. 2015. "Betting the House." *Journal of International Economics* 96: S2–18.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor. 2016. "The Great Mortgaging: Housing Finance, Crises, and Business Cycles." *Economic Policy* 31 (85): 107–52.

Jordà, Òscar, Sanjay R. Singh, and Alan M. Taylor. 2020. "The Long-Run Effects of Monetary Policy." NBER Working Paper 26666.

Krishnamurthy, Arvind, and Tyler Muir. 2017. "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 and Economic Statistics* 28 (1): 159–68.

Lanne, Markku, Helmut Lütkepohl, and Katarzyna Maciejowska. 2010. "Structural Vector Autoregressions with Markov Switching." *Journal of Economic Dynamics and Control* 34 (2): 121–31.

López-Salido, David, Jeremy C. Stein, and Egon Zakrajšek. 2017. "Credit-Market Sentiment and the Business Cycle." *Quarterly Journal of Economics* 132 (3): 1373–1426.

Ludvigson, Sydney C., Sai Ma, and Serena Ng. 2017. "Shock Restricted Structural Vector-Autoregressions." NBER Working Paper 23225.

McKinnon, Ronald. 1973. *Money and Capital in Economic Development*. Washington, DC: Brookings Institution.

Mian, Atif R., Amir Sufi, and Emil Verner. 2017. "Household Debt and Business Cycles Worldwide." *Quarterly Journal of Economics* 132 (4):1755–1817.

Plagborg-Møller, Mikkel, and Christian K. Wolf. 2021. "Local Projections and VARs Estimate the Same Impulse Responses." *Econometrica* 89: 955–980.

Rajan, Raghuram G., and Luigi Zingales. 1998. "Financial Dependence and Growth." *American Economic Review* 88 (3): 559–86.

Rigobon, Roberto. 2003. "Identification through Heteroskedasticity." *Review of Economics and Statistics* 85 (4): 777–92.

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 Deepening in Economic Development*. New York: Oxford University Press.

Sims, Christopher A. 2020. "SVAR Identification through Heteroskedasticity with Misspecified Regimes." Unpublished.

Sims, Christopher A., and Tao Zha. 1996. "Bayesian Methods for Dynamic Multivariate Models." Federal Reserve Bank of Atlanta Working Paper 96-13.

Sims, Christopher A., and Tao Zha. 2006. "Were There Regime Switches in US 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.

Stock, James H., and Mark W. Watson. 1989. "New Indexes of Coincident and Leading Economic Indicators." In *NBER Macroeconomics Annual 1989,* Vol. 4, edited by Olivier Jean Blanchard and Stanley Fischer, 351–94. Cambridge, MA: MIT Press.

Stock, James, and Mark Watson. 2012. "Disentangling the Channels of the 2007–2009 Recession." *Brookings Papers on Economic Activity*: 81–135.

World Bank. 2012. *Global Financial Development Report 2013: Rethinking the Role of the State in Finance.* Washington, DC: World Bank.