ABSTRACT. We examine the relation among measures of credit expansion, measures of financial market stress, and standard macroeconomic aggregates. We use a form of structural VAR with monthly data on 10 variables. The model explains observed variation as driven by 10 mutually independent structural disturbances. We identify the shocks from variation across time in their relative variability. One of them emerges as representing monetary policy. We find two distinct financial stress shocks, suggesting that attempts to create a one-dimensional index of financial stress may be misguided. While our results are consistent with the finding by others of a negative reduced form relation between credit expansion and future output growth at certain frequencies, we find the output decline to be explained by the monetary policy response to the inflation that accompanies the credit expansion. 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.

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. 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. 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.

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

Date: June 9, 2017.

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 early perspectives on this topic are found, for example, in Shaw (1973), McKinnon (1973), and Goldsmith (1969). A summary of the related literature is available in World Bank (2012), p.23-25.
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.

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 (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)). Studies focused on the information in credit spreads have looked extensively at single-equation models (e.g., Lopez-Salido, Stein and Zakrajšek (2015), and Krishnamurthy and Muir (2016)) and reduced form multi-equation models (Gilchrist, Yankov and Zakrajšek (2009) and Gilchrist and Zakrajšek (2012)). Gertler and Karadi (2015) and Caldara and Herbst (2016) introduce credit spread variables into structurally identified, multiple-equation

---

Mian, 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.
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.  

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. 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 corporate bond spread (GZ) and the 3-month Eurodollar over Treasury spread (ES).  

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 usually assumed about the responses to a monetary policy shock in SVAR models. 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 corporate bond and Eurodollar 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 household and business credit, and all but one of them move IP, if at all, 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. This disturbance increases inflation as well as credit, and initially it increases IP, so it is natural that monetary policy responds to it by increasing interest rates. Our model implies that if the shock were accompanied by monetary policy holding interest rates fixed, the decline in IP following the shock would disappear, though of course at the cost of increased inflation.

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

---

6 Though 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.
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 included. 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 specification account for the apparent difference in results. Even a small, four-variable multivariate model, which treats all predictive variables as jointly endogenous, implies a relatively small negative effect on credit growth on output. Increasing the number of variables, which further reduces endogeneity problems, and accounting for heteroskedasticity, which allows the model to down-weight high variance periods, further reduces estimated effects. In a Monte Carlo exercise with simulated data, we find that the results others have obtained with smaller models would not be unlikely if our full estimated model represented the true structure of the economy.

II. MODELING FRAMEWORK

This section describes our empirical approach. The first two parts introduce our multivariate time series models, 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

\[ A_0 y_t = \sum_{j=1}^{p} A_j y_{t-j} + C + \epsilon_t \]  

\(^7\)This 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).
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 $\epsilon_t$ is a vector of independent shocks. In the base model, these are Gaussian (normally distributed).

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

$$
E[\epsilon_t \epsilon_t'] = \Lambda_m \quad \text{if } t \text{ is in period } m \in \{1 \ldots M\}
$$

(2)

where $\Lambda_m$ 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 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. 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,

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

(3)

with

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

(4)
Some normalization is required, as we could multiply the rows of $A_0$ and $\Lambda$ by scale factors without changing the implied behavior of the data. We impose the restriction

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

where $\lambda_{m,i}$ is the $i$th diagonal element of $\Lambda_m$. 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). Thus the variance switching eliminates the need for arguable linear restrictions on the $(A_i)_{i=0}^P$, such as “short-run” restrictions on contemporaneous responses in $A_0$ or “long-run” restrictions on the sums of coefficients in the $(A_i)_{i=1}^P$. Avoiding such restrictions on $A_0$, while still maintaining full identification, is a particularly appealing feature for a model including financial time series which would plausibly react to all shocks immediately at the monthly or lower frequency.

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. In our later analysis, 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.

---

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

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

has the form of an eigenvalue decomposition, with the columns of $A_0'$ the eigenvectors. As long as the eigenvalues, the diagonal elements of $\Lambda_j^{-1} \Lambda_j$, 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 $\Sigma_j$ and $\Sigma_i$. A more formal proof of this can be found, for instance, in Lanne, Lütkepohl and Maciejowska (2010).
Aside from flexible identification of impulse responses, another benefit of our method is efficiency. For the same reasons as standard GLS (Generalized Least Squares), our method prevents periods of large shocks from inefficiently dominating the likelihood.

II.2. Non-normality in the error distribution. The previous section’s correction for heteroskedasticity will work best if volatilities mainly change between persistent episodes or regimes. 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 an alternative specification in which structural errors \( \epsilon_t \) have Student \( t \) distributions.\(^9\)

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

\[
\epsilon_{i,t} \sim \text{Normal} \left( 0, \lambda_{i,t} \xi_{i,t} \right)
\]  

(5)

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.

Giving the \( \xi_{i,t} \) an inverse-gamma distribution, i.e.

\[
\xi_{i,t} \sim \text{Inverse Gamma} \left( \text{shape} = \alpha / 2, \text{rate} = 2 / \alpha \right).
\]  

(6)

\(^9\)We also estimated a version of the model assuming a finite scale-mixture of normals as the distribution of the residuals. That model produced very similar results to those we display and discuss from the \( t \)-distributed errors model, and had lower likelihood. (Both the \( t \) and mixture-of-normals models had much higher likelihood than models that assumed Gaussian errors.) 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 (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 Curdia, 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.
implies that each $\epsilon_{i,t}$ has an independent Student-$t$ distribution with $\alpha$ degrees of freedom and unit scale.\footnote{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.\footnote{All 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), instead of the unit scale.}}

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 $\Lambda_m$, and $n^2p$ free parameters in the $A_j$. The model with $t$-distributed disturbances has another $nT$ parameters.\footnote{As mentioned previously, the $\zeta_{it}$ 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.}

We use Bayesian methods to update beliefs about the parameters conditional on observed data \{y₁. . . yₜ\} and initial conditions \{y₋ₚ₋₁ . . . y₀\}. For $A_0$ we specify independent Gaussian priors on all elements, centered around 100 times the identity matrix, with standard deviation 200. For $\lambda_{·,j} = \{\lambda_{1,i} . . . \lambda_{M,i}\}$, the vector of variances in each equation $i$, we put a Dirichlet prior (with $\alpha = 2$) on $\lambda_{·,j}/M$. This restricts each of the relative variances to lie in $[0, M]$ (where $M = 10$ 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. 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 A, 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 case with normal structural shocks, we sample from the posterior distribution for these parameters in a two-step process that exploits the fact that, conditional on knowing...
and \( \{\Lambda_1 \ldots \Lambda_M\} \), 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 \( \{\Lambda_1 \ldots \Lambda_M\} \) using the likelihood integrated over the \( A_j \) (which is available analytically). Then, after drawing a large MCMC sample from the marginal posterior distribution of \( A_0 \) and \( \{\Lambda_1 \ldots \Lambda_M\} \) we can, for each of the \( A_0, \Lambda \) draws, sample from the coefficients in the \( A_j \) which have a known conditional normal distribution.

The model with \( t \)-distributed shocks requires a more involved method, a (non-standard) Metropolis-in-Gibbs algorithm. The first part of the algorithm, a Monte Carlo update of \( A_0 \) and \( \{\Lambda_1 \ldots \Lambda_M\} \) conditional on the \( \xi_{i,t} \) and integrated over the \( (A_j)_{j=1}^p \), is just as in the normal shocks model. But now we need to draw \( (A_j)_{j=1}^p \) after each draw of \( (A_0, \Lambda) \) to form implied normalized residuals \( \epsilon_{i,t} \). The \( \xi_{i,t} \) conditional on the \( \epsilon_{i,t} \) are distributed independently across \( i \) and \( t \), allowing us to draw directly from their exact conditional posterior distribution. The process is repeated recursively. Appendix B describes the mechanics of the both algorithms in more detail, and Appendix C reports diagnostic evidence of its convergence to randomly sampling the model posterior.

For a given set of model parameters, we could change the sign of the coefficients in an equation (a row of \( (A_0, A_1, \ldots) \)) or change the order of the equations (permute the rows of \( (A_0, A_1, \ldots) \)), 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,
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.

if applied on data for which identification did not emerge as strongly, might need to test for and eliminate permuted or sign-switched models.\(^\text{13}\)

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 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 from 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)). The lag length \(p\) in our model is set to 10.

Our measures of “household” and “business” credit are based on the Federal Reserve’s weekly surveys of U.S. commercial banks.\(^\text{14}\) 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

\(^{13}\)This is a special case of the kind of normalization issue discussed by Hamilton, Waggoner and Zha (2007).

\(^{14}\)These are published in the H.8 “Assets and Liabilities of Commercial Banks in the United States” release.
Table 2. Dates for variance regimes in full model specification.

<table>
<thead>
<tr>
<th>Start</th>
<th>End</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Jan 1973</td>
<td>Sep 1979</td>
<td>Oil crisis, stagflation, and Burns Federal Reserve</td>
</tr>
<tr>
<td>2 Oct 1979</td>
<td>Dec 1982</td>
<td>Volcker disinflation</td>
</tr>
<tr>
<td>3 Jan 1983</td>
<td>Dec 1989</td>
<td>Recovery from early 1980s recession</td>
</tr>
<tr>
<td>4 Jan 1990</td>
<td>Dec 2007</td>
<td>Great Moderation and Greenspan Federal Reserve</td>
</tr>
<tr>
<td>5 Jan 2008</td>
<td>Dec 2010</td>
<td>Great Recession</td>
</tr>
<tr>
<td>6 Jan 2011</td>
<td>Jun 2015</td>
<td>Zero Lower Bound, Recovery from Great Recession</td>
</tr>
</tbody>
</table>

noncorporate business,” and “nonfinancial corporate business”), used in some other research.

Appendix E.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 seems 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.

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 lower, long rates, if the monetary tightening does succeed in lowering inflation.

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

\[\text{Appendix E.1 discusses the differences. Although our “household credit” series includes commercial real estate loans (which cannot be separately identified for the entire sample in the data) and our “business credit” data seems to have more high-frequency variation than the corresponding quarterly series, we believe these data capture the majority of the low-frequency behaviors that are critical for existing empirical evidence of their forecasting power.}\]

\[\text{The inclusion of three credit spreads (of interest rates over short-term Treasuries) is meant to capture several possible dimensions of credit market stress: the term spread captures inflation expectations and uncertainty about future movements in fundamentals, the bond spread captures tightness in business financing, and the TED spread captures tightness in bank financing. The first was also expected to, along with the Federal Funds rate, M1, and commodity prices, provide a sharper identification of a monetary policy shock, as policy-generated rises in the short rate might be expected to have little effect on, or even lower, long rates, if the monetary tightening does succeed in lowering inflation.}\]
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 model with Gaussian distributed 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. The $t$ model clearly fits much better than the model with Gaussian errors.\(^\text{17}\)

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

---

\(^{17}\)See Appendix section C.4 for more details.
distribution with 5.7 degrees of freedom. Since the diagonal of $\Lambda_i$ 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 recognizable monetary policy (number 6) and credit spread (9-10) shocks with significant long-term real consequences. Shock 6 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 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 the 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 68% confidence bands is barely significantly less than zero at the five-year or smaller horizon. Second, several other shocks to which the credit variables respond substantially (shocks 1, 6, 9 and 10) move output and credit in the same direction. Over
medium and long horizons, observed credit growth is as likely to result from these shocks as from the model’s third shock. Distinguishing “good” from “bad” credit growths, then, requires observing all ten variables in this model.

Note that this shock also elicits a positive movement in R, which is not surprising given that it is associated with rising inflation and, initially, rising output. If we accompany this shock with a sequence of shock-6 values that suppress this R response, we find that the inflation and credit expansion are somewhat greater and there is no output decline. In other words, the model implies that the decline in output that follows this type of credit expansion can be accounted for by the systematic monetary policy contraction that the credit shock elicits. Figure 5 shows the responses to the third shock, and to the third shock with policy keeping the interest rate constant. (The figure omits responses that were nearly the same for the two patterns of shock.)

The seventh shock accounts for a substantial component of variation in 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 by a non-trivial, but statistically marginal, amount. 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 when put on a common scale. The main difference is that most of the error bands are somewhat narrower for the t-residuals model. This is what would be expected if the t model were correct. The impulse responses from the misspecified Gaussian model in that case would be consistent, but somewhat inefficient estimates. Figure 6 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. The differences between them are mainly within the error bands, with two major exceptions. The normal errors model puts more posterior probability on a nonzero output effect for
the household credit shock and the interbank spread shock. The output response of the former, with normal errors, is comfortably significant (less than zero) with 68% bands and barely significant with 90% bands (which are not pictured). The output response of the latter is comfortably significant, from the first 18 months of response, at both levels. Full impulse responses from the Gaussian errors model, and a model with multinomial normal mixture errors, can be found in Appendix D.

The picture is also very similar if we estimate with data only up to December 2007, as shown in Figures 7 to 10. In particular, the identification of monetary policy and spread effects is very stable. There is weak (within 68%, but not 90% bands) evidence of an output response to household and business credit expansion shocks. These effects are of comparable magnitude to the estimated household credit effect in the model estimated on the full dataset.

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 $t$-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 $t$-errors assumption to justify large residuals. Table 5 lists the 10 largest posterior median shocks for individual equations and time periods. These are in standard deviation units and not scaled by the corresponding $\lambda$ values. They thus show the biggest shocks, not the biggest “surprises” for the model.

The biggest of these shocks reflects a surprise easing of monetary policy in May 1980, during the recession of that year. The Federal Funds rate fell from 18 per cent to 11 percent
in that month, but soon started rising again. The two next largest were in the two financial
stress indicators, shocks 9 and 10, at the time of the Lehman collapse in October 2008.

The fourth and tenth biggest shocks, in shock 3, March and April 2010, reflect a mis-
judgment in our correction for a break in the credit series between March and April of
that year. We discuss in appendix section E.1 the adjustments for breaks that we made to
the series and the nature of this misjudgment. Though these residuals are large in stan-
dard deviation units, since they occurred during the high-variance crisis period, they are
not among the top 10 “surprises”, and estimation with data spliced correctly at this date
does not visibly change the plotted responses.

The fifth biggest shock is shock 1 in September 2008, reflecting the large decline in
industrial production as the crisis took hold.

The sixth biggest is in the monetary policy shock 6, March 1980, as the Federal Funds
rate jumped from 14 per cent to 17 per cent.

The next is October 2001, just after the 9/11 attack, as the Fed withdrew the temporary
liquidity it had provided in the immediate wake of the attack.

Finally there are two shocks from the period in early 1981 when the Federal Funds
rate, over 19 per cent in January and June of that year, briefly dropped to 14.7 per cent in
March.

Except for the two that reflect mis-splicing of one of the credit series, these large shocks
all correspond to events that were recognizably unusual as they occurred. But they tend
to come from periods with large values of $\lambda_{it}$, 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, as we will see, the size
and distribution across variables of these large surprises conflicts with our assumption of
i.i.d. $t$-distributed scaled shocks.

Half of the 10 biggest unscaled residuals appear also among the 10 biggest surprises.
The high value of $\lambda_6$ during the 1979-82 period leaves all but one of the 1980-81 shock
6 values that appear in Table 5 out of the “big surprise” category, and the two shocks
arising from mis-splicing are also not big surprises. Shock 5 in September 2001 joins the October 2001 shock 5 as a big surprise, so both the initial liquidity injection and the later withdrawal of it were important to model fit. An increase in interbank financial stress (shock 10) in August 2007 appears as a big surprise, as does the subsequent easing of monetary policy (shock 6) in February 1980. July 2002, during the bursting of the internet bubble, appears as a surprise increase in the corporate bond spread financial shock (9). August 1974, the month of President Nixon’s resignation, appears as a shock to the term spread (shock 8).

The last column of Table 6 shows that shocks of these sizes are not, under our assumption of i.i.d. $t(5.7)$ distributed residuals, completely unexpected in a sample of the size of our model’s residuals. However there are nonetheless too many of these large shocks to be consistent with our assumed distribution. We would expect about two shocks of the size of the 10th largest in Table 6 or larger in a sample of size 5000, whereas we see 10. Under our assumed distribution, seeing 10 instead of 2 has probability on the order of $10^{-7}$.

Half of the 10 biggest surprises occurred during 2007-8. These surprises do not occur in the same months, but their bunching in time suggests a possible problem with our break dates. Furthermore the large financial stress shocks (9 and 10) are all positive. These observations suggest directions for improvement in the model that we take up in more detail below.

The surprising upward movement in shock 10 in August 2007 by itself moved the model’s forecast path for IP downward, but it was offset by the surprise easing of the monetary policy shock 6 in February 2008. Thus despite the occurrence of one large financial shock a year in advance, as we will see below the model does not give much advance notice of the crisis.

The remainder of this section reviews the model dynamics of key shocks in greater detail.
### Table 5. Ten largest structural residuals $\epsilon_{i,t}$ in the model and the corresponding “diagonal impact” ($dy_{i,t}$) in the units of variable $i$ (logs for industrial production, price indices, credit aggregates, and the money supply; raw annual rates (not percent) for interest rate and spread variables). Shocks are identified in Table 2. Point estimates are posterior median values.

<table>
<thead>
<tr>
<th>Month</th>
<th>Shock</th>
<th>$\epsilon_{i,t}$</th>
<th>$dy_{i,t}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5/1980</td>
<td>6</td>
<td>-21.314</td>
<td>-0.067</td>
</tr>
<tr>
<td>10/2008</td>
<td>9</td>
<td>16.064</td>
<td>0.023</td>
</tr>
<tr>
<td>10/2008</td>
<td>10</td>
<td>10.499</td>
<td>0.019</td>
</tr>
<tr>
<td>4/2010</td>
<td>3</td>
<td>-10.174</td>
<td>-0.025</td>
</tr>
<tr>
<td>9/2008</td>
<td>1</td>
<td>-9.837</td>
<td>-0.043</td>
</tr>
<tr>
<td>3/1980</td>
<td>6</td>
<td>9.814</td>
<td>0.031</td>
</tr>
<tr>
<td>10/2001</td>
<td>5</td>
<td>-9.609</td>
<td>-0.047</td>
</tr>
<tr>
<td>2/1981</td>
<td>6</td>
<td>-9.509</td>
<td>-0.030</td>
</tr>
<tr>
<td>5/1981</td>
<td>6</td>
<td>8.904</td>
<td>0.028</td>
</tr>
<tr>
<td>3/2010</td>
<td>3</td>
<td>8.609</td>
<td>0.021</td>
</tr>
</tbody>
</table>

### Table 6. Ten Largest surprises. Residuals have been scaled by the regime scale factors and should thus be approximately $t$-distributed with 5.7 degrees of freedom. The last column shows how many residuals of this size or larger would be expected in a sample of about 5000 (roughly matching the number of residuals in our data) if they were in fact i.i.d. draws from $t(5.7)$.

<table>
<thead>
<tr>
<th>Month</th>
<th>Shock</th>
<th>$\epsilon_{i,t}/\sqrt{\Lambda_{ii}}$</th>
<th>Expected count</th>
</tr>
</thead>
<tbody>
<tr>
<td>10/2001</td>
<td>5</td>
<td>-12.218</td>
<td>0.132</td>
</tr>
<tr>
<td>9/2001</td>
<td>5</td>
<td>10.454</td>
<td>0.312</td>
</tr>
<tr>
<td>5/1980</td>
<td>6</td>
<td>-9.987</td>
<td>0.400</td>
</tr>
<tr>
<td>2/2008</td>
<td>6</td>
<td>-9.070</td>
<td>0.672</td>
</tr>
<tr>
<td>10/2008</td>
<td>9</td>
<td>8.399</td>
<td>1.015</td>
</tr>
<tr>
<td>9/2008</td>
<td>1</td>
<td>-8.097</td>
<td>1.233</td>
</tr>
<tr>
<td>7/2002</td>
<td>9</td>
<td>7.984</td>
<td>1.328</td>
</tr>
<tr>
<td>8/2007</td>
<td>10</td>
<td>7.678</td>
<td>1.633</td>
</tr>
<tr>
<td>8/1974</td>
<td>8</td>
<td>-7.580</td>
<td>1.747</td>
</tr>
<tr>
<td>10/2008</td>
<td>10</td>
<td>7.497</td>
<td>1.851</td>
</tr>
</tbody>
</table>

### III.1. The Credit Channel of Monetary Policy. The sixth ordered shock of the model, the impulse responses of which are collected in Figure [II], 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).\footnote{It 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 Figure 11 with Figures 8 and 10. Apparently the positive interest 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.}

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 \cite{Gertler2015} and a variety of theoretical models which suggest that risk premia should move in response to monetary policy (e.g., \cite{Drechsler2016} and \cite{Brunnermeier2016}). Our empirical result is focused, however, on \textit{inter-bank} credit conditions separate from \textit{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 \cite{Gertler1993}, and a version of it is corroborated in bank balance sheet data in the recent financial crisis by \cite{Ivashina2010}. It also anticipates the kind of under-specification problems discussed in greater detail in Section \ref{sec:under-specification}.
spread, it might suggest that contemporaneous innovations to business credit “cause” recessions, while in the full model, temporary credit expansions and long-term output and price contractions are both results of monetary policy shocks.

III.2. Spread Spikes and Early Warning. Three independent shocks, ordered eight to ten, can be identified by sharp increases in spreads at $t = 0$, but only the the latter two (the responses to which are plotted in Figure 12) have significant output effects. The output effect is larger and more significant for the ninth shock, associated with an initial surge in the corporate bond spread and a long-term contraction in business credit. The tenth shock, in contrast, begins with a shock to the inter-bank 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.

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 main model offers some support, within 68% error bands, of the hypothesis that excessive growth in household credit can forecast negative long-term real output growth (Figure 13). 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.

However, as mentioned earlier in the qualitative discussion of results, 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 shock 3, combined with a sequence of shock 6 values that keep the interest rate constant, eliminates the decline in output.

The estimated magnitude of this credit to real output channel is low relative to previous estimates in the literature. A one per cent expansion of household credit over about 2.5 years is followed by an output decline reaching 25 basis points below trend after five years. A one-standard-deviation shock during the period before the 2008 crisis would imply an eventual output decline after five years of 37 basis points. This is half the size of the effect of a one-standard-deviation shock in the household-credit-to-gdp ratio found by Mian, Sufi and Verner (2015) in their 3-variable VAR model, and many times smaller than what they found in their single-equation estimates.
One way of quantifying the importance of this credit shock relative to others in the model is to calculate forecast error variance decompositions (Figure 14). 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 28 percent of household credit variation, while other shocks with “passive” credit reactions moving in the same direction as output explain the remainder. At this same horizon, the explanatory percentage for output is, at the posterior median estimate, is only about 2%. This decomposition, based on the average impulse responses, is somewhat different in the two periods in which the variance of shock is the largest (1/1990-12/2007 and 1/2008-12/2010). In these periods, the shock explains a greater proportion of 60-month forecast household credit variation (57.7% and 51.7% respectively, at posterior median values) but still relatively little output variation (5.1% and 4.1% respectively).

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

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 $j$. 
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 15, 16, and 17 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 17) 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,
We focus on root mean squared error (RMSE) for forecasts of all variables common to the models.

Figures 18 and 19 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 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.

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

---

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.

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

V.1. Data Selection and 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. It is possible that using higher frequency data (monthly instead of quarterly) explains the small negative effects of credit expansions on output. As a first step, we try estimating the same model with quarterly averages of all the monthly time series.\textsuperscript{22} Like the baseline model with monthly data, this model (estimated with Gaussian errors) fits one shock in which a credit variable moves promptly and IP moves with a delay in

\textsuperscript{22}One technical deviation is that these models use a slightly different “Minnesota” prior on the reduced form coefficients, with a tightness of 5 and decay of 1. Qualitatively, the prior is tighter because the same “number” of lag in a quarterly model corresponds to older information. This is close to, but not exactly, what we would get from averaging the monthly dummy observations into quarterly dummy observations.
the opposite direction. The magnitude of the response is quite similar for the monthly and quarterly estimates (left panel of Figure 20).

V.1.2. Alternative Quarterly Data. A closely related question is whether the results would change if we used quarterly data reported by the Federal Reserve in the Z.1 Financial Accounts table and by the Bank of International Settlements in its cross-country panel of credit to the private sector. It is possible that either the lower frequency or arguably more comprehensive construction of these data would make a difference.

The results of a model estimated from these data are slightly different than those from our baseline, but still do not constitute, in our view, strong evidence for the hypothesis that a channel of “explosive credit growth” and opposite output movement explains a significant amount of output dynamics. 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 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. 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.” This still falls short of the type of estimates reported in the single-equation literature.

V.1.3. Long-term Growth Rates. It is possible that the most relevant frequency credit growth for these results is much larger than the amount of lags that we include. A $p$ lag system can fit dynamics with any linear combination of the variables including an $m \leq p$ month lag system. At the posterior median of draws.
(or quarter) difference, but obviously does not capture longer spans. The models estimated so far include 10 months or 4 quarters and thus cannot capture growth effects averaged over multiple years. 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 21 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\textsuperscript{24}. The differences model fits considerably worse than the levels model, with an estimated marginal likelihood of 51244.7 (compared to 51530.5). So even if the model did imply a very strong connection between credit growth and recessions, we would have strong evidence to reject that model’s predictions.

V.1.4. Non-linear Transformation. Potentially our method is failing to capture a positive long-term relationship because it considers only linear effects. This implies that small credit movements have effects in proportion to large ones, and that negative shocks have an equal and opposite effect as positive shocks\textsuperscript{25}.

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

\textsuperscript{24}The 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}, \ldots \}, where \(c_t\) is the credit variable. This is different but not higher dimensional information.

\textsuperscript{25}Jordà, Schularick and Taylor\textsuperscript{[2013]}, for instance, focus on upper \(n\) percentile credit growth events.
\[ f(c_t) = \begin{cases} 
  c_t & \text{if } c_t < a \\
  \alpha_1 + \alpha_2 c_t + \alpha_3 c_t^2 & \text{if } a \leq c_t < b \\
  \beta c_t & \text{if } c_t \geq b 
\end{cases} \]

for \( c_t \) as credit growth rates, \( b > a > 0 \), and with the coefficients \( \alpha_i \) and \( \beta \) 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 \( \beta \), as well as the other parameters of the model. In all optimization exercises to find the posterior mode of such a model, the data favored models without any nonlinear transformation.

V.1.5. Interaction Terms. Many other studies have focused exclusively on various interaction terms of credit aggregates, output, and interest rates. We omit discussion of the ratio of credit aggregates to real GDP (\( \log \text{CGDP} = \log C - \log Y - \log P \)) because the specification separately in terms of \( \log C, \log Y, \) and \( \log P \) nests the former and more transparently separates the role of each component.

A different suggestion comes from studies which emphasize interactions of credit growth and interest rates, either with the specific interpretation of a debt service burden or the more “reduced-form” interpretation of intersecting signals of loan losses with signals of credit market fragility (see, for example, Drehmann and Juselius (2014) and Krishnamurthy and Muir (2016)). In our own trials replacing credit aggregates with these interaction variables (the full results of which we omit for brevity), the story is much the same. The “credit times interest rate” variables spike along with interest rates, and necessarily lead output downturns, but there are no separately identified shocks which hit the “credit times interest rate” variables and significantly affect output.

\(^{26}\)In 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 \( \beta \), with an exponential prior with parameter 2

\(^{27}\)In this specification, \( \beta_1 = 1, \alpha_1 = \alpha_3 = 0, \alpha_2 = 1 \).

\(^{28}\)In the spirit of the last subsection, we do not suspect that exponentiating to get credit-to-GDP units would make a significant difference.

\(^{29}\)It is possible that the more appropriate interpretation of this argument is that impulse responses should be state-dependent, and the state of the economy endogenously depends on credit growth. We leave this and related investigations of endogenous state-dependence of dynamics to further research.
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 22 presents the impulse response of a model (estimated with heteroskedasticity, like the base model) with four variables: real output, the 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. This is considerable relative to the fraction of business loans variation produced by the corresponding shock in the 10-variable model, as visualized in Figure 14 and discussed in Section III.3. 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.

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

As a first test, we simulated 10,000 samples from the 10-variable estimated model with Gaussian or \(t\) innovations. and, for each artificial time series, estimated four-variable VAR (including output (IP), the price level (P), household credit (HHC), and business credit (BC)). We obtained negative and “large” (greater than a 0.5% output contraction) response to a one-standard deviation credit shock with the probabilities shown in Table 7. These calculations are done with data at different frequencies, aggregated by taking averages of the monthly data our model generates. Observing negative point estimates (suggesting a negative impulse response of output to credit shocks) is not so uncommon, particularly as the frequency of data is reduced to annual.

When we imitate the “projection method” regressions of Mian, Sufi and Verner (2015), which predict three-year growth rates of output with lagged three-year averages of credit growth and controls for lagged annual output growth, the results are similar. We find

---

31 We set the “posterior mode point estimate” \(\hat{\theta} = \{\hat{A}_0, \hat{\Lambda}, \hat{A}_{1:p}\}\) to maximize the conditional posterior density \(P[A_0, \Lambda, A_{1:p} | \xi_{it}]\). For the Gaussian errors model, for which \(\xi_{it} \equiv 1\), we have a reliable estimate of this mode from a numerical optimization method. For the \(t\)-errors model, we choose the highest density draw from our MCMC sample.

32 We simulate samples of the same length as the data (500 months) with the same initial conditions and sequence of regime variances. In the \(t\)-errors model case, we simulate from the unconditional \(t\) distribution rather than the normal distribution conditional on the \(\xi_{it}\); results with the other method are comparable.

The entire exercise, described differently, involves drawing artificial data \(\tilde{Y}\) from \(P[Y \mid \hat{\theta}, \mathcal{M}]\) (where \(\mathcal{M}\) denotes our model), forming some transformation \(g(\tilde{Y})\), and estimating the relative likelihood of observing \(g(Y) = \hat{g}(Y)\), where the latter is the empirical estimate from observed data. We could also have done this exercise by forming \(\tilde{Y}\) from \(P[Y \mid \theta, \mathcal{M}]\). Practically, this would involve first sampling from the posterior of \(\theta\) (for instance, by picking uniformly over stored MCMC draws) and then, conditional on each draw \(\theta^{(i)}\), forming \(\tilde{Y}^{(i)}\) from the conditional distribution \(P[Y \mid \theta^{(i)}, \mathcal{M}]\). This exercise produces very similar results.

33 These are the point estimate impulse response from lower triangular (Cholesky) identified VARs, with “standard” coefficient shrinkage, unit root, and covariance priors.

34 We 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. In equation form, with \(y_t\) denoting log IP deflated by PCEPI, \(hc_t\) the household credit to GDP ratio, and \(bc_t\) the business credit
the probability of credit growth coefficients smaller than 0 and -0.2 (a benchmark for an economically significant result) as in Table 8. 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.

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

and the table reports probabilities for \( \beta_h \) and \( \beta_b \). The results “with lagged IP” set \( k = 3 \); otherwise we set all \( \gamma_i \equiv 0 \).
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 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 23 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 sizable long-term output contraction (about 0.4% over 48 months, explaining 13.0% of variance at that horizon) in response to a business loans 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 Federal Funds Rate (R), corporate bond spread (GZ), and interbank lending spread (ES), explaining 4.0%, 4.2%, and 6.8% of the first-period variance of those variables respectively. We might suspect that this shock at least partially represents business (but not household) borrowing in the midst of financial stress and monetary policy stringency.

Which of the two changes from the baseline—adding heteroskedasticity or removing restrictions—seems more important for generating this result? 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 24. The fourth shock of this model has a considerably smaller (and less precise) output effect.

Table 9 sheds some light on the discrepancy by reporting the correlations between shocks 4 (business loans), 6 (monetary policy), 9 (corporate bond spread), and 10 (interbank spread) in the two models. The business credit shock in the plain, Cholesky-identified VAR (rows) “absorbs” small amounts of each of the three interest rate shocks in the model with regime switching heteroskedasticity. Breaking this down by regime illustrates the
pattern more closely. In the period of aggressive Federal Reserve action (October 1979 to December 1982), the Cholesky VAR’s business loans shock is very closely correlated with the other model’s monetary policy shock (0.468). In the financial crisis and Great Recession (January 2008 to December 2010), the Cholesky VAR’s business loans shock is positively correlated with the corporate bond and inter-bank shocks (0.637 and 0.601 respectively) and negatively correlated with the monetary policy shock (-0.107) to partially offset the effect. In this context, identification through heteroskedasticity delivers a different, and perhaps a priori more reasonable, separation of policy, financial stress, and credit expansion effects than a more standard approach.

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 model mainly move “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. It is accompanied by rising interest rates, and if monetary policy offset that rise, the effect of the credit shock 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. 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 \textit{a priori} restrictions, can shed light on complex interactions among policy, financial markets, and the real economy.
REFERENCES


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 1. (1/4) 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 scale. Variables are in the order listed in Table 1.
FIGURE 2. (2/4) 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 scale. Variables are in the order listed in Table 1.
FIGURE 3. (3/4) 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 scale. Variables are in the order listed in Table 1.
FIGURE 4. (4/4) 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 scale. Variables are in the order listed in Table [1].
Figure 5. The first column reproduces the responses to shock 3 shown in Figures 1 and 3. The second column combines shock 3 with a series of shocks to shock 6, the monetary policy shock, that keep R constant.
Fig. 6. 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 7. (1/4) Impulse responses to the ten orthogonal structural shocks in a model with $t$-distributed errors estimated up to December 2007.
Figure 8. (2/4) Impulse responses to the ten orthogonal structural shocks in a model with t-distributed errors estimated up to December 2007.
Figure 9. (3/4) Impulse responses to the ten orthogonal structural shocks in a model with $t$-distributed errors estimated up to December 2007.
Figure 10. (4/4) Impulse responses to the ten orthogonal structural shocks in a model with $t$-distributed errors estimated up to December 2007.
FIGURE 11. Impulse responses to the 6th ordered (monetary policy) shock, with 68% (dark blue) and 90% (light blue) posterior uncertainty regions. Scaled to an “average” period with unit scale. Variables are in the order listed in Table 1.
**Figure 12.** Impulse responses to the 9th (bond spread) and 10th (inter-bank spread) ordered shocks, with 68% (dark blue) and 90% (light blue) posterior uncertainty regions. Scaled to an “average” period with unit scale. Variables are in the order listed in Table 1.
FIGURE 13. Impulse responses to the 3rd (household credit) and 4th (firm credit) ordered shocks, with 68% (dark blue) and 90% (light blue) posterior uncertainty regions. Scaled to an “average” period with unit scale. Variables are in the order listed in Table I.
FIGURE 14. 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 scale. Variables are in the order listed in Table 1.
Figure 15. 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 16. Posterior mode forecasts, from a model without the two credit aggregates, estimated up to points in and around the Great Recession.
Figure 17. Posterior mode forecasts, from a model without the three credit spreads, estimated up to points in and around the Great Recession.
Figure 18. 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 19. 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 20. 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 21. Impulse responses to candidate credit growth shocks in a (Gaussian errors) model with 3 year growth rates of credit aggregates.
**Figure 22.** 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 23. 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 24. 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.