Distinguishing between Macroeconomic and Financial Uncertainty: Classification Search in Stochastic Volatility in Mean VARs

Preliminary Draft

Sharada Nia Davidson* Chenghan Hou† Gary Koop‡

September 24, 2021

Abstract

Stochastic Volatility in Mean Vector Autoregressive models (SVMVARs) are a popular tool for measuring macroeconomic and financial uncertainty and their economic impacts. SVMVARs estimate macroeconomic (financial) uncertainty using a large set of macroeconomic (financial) variables. But what if there is uncertainty regarding whether variables are classified as macroeconomic or financial? We address this question, developing scalable Markov chain Monte Carlo algorithms for classification search in large SVMVARs with unclassified variables. Using time-invariant or time-varying classification, the algorithm determines whether each unclassified variable should be treated as macroeconomic or financial. We show that allowing for data-driven classification improves model fit. Our results also suggest that without data-driven classification, macroeconomic uncertainty, its adverse effects and its contribution to fluctuations in economic variables tend to be underestimated. Financial uncertainty is also underestimated but its effects on headline macroeconomic variables tend to be overestimated.

Keywords: Bayesian VAR, Uncertainty, Stochastic Volatility, Big Data

JEL Codes: C11, C55, E32, E44.

*University of Strathclyde, Department of Economics, UK. Email: sharada.n.davidson@strath.ac.uk.

†Hunan University, Center for Economics, Finance and Management Studies, China. Email: chenghan.hou@hotmail.com

‡University of Strathclyde, Department of Economics, UK. Email: gary.koop@strath.ac.uk.

1 Introduction

Macroeconomists and policymakers have begun to investigate the different roles played by macroeconomic and financial uncertainty in business cycle fluctuations. But can we easily distinguish between macroeconomic and financial uncertainty? A growing body of literature attempts to do so using a two step approach (see Jurado, Ludvigson and Ng, 2015; Ludvigson, Ma and Ng, forthcoming; Carriero, Clark and Marcellino, 2018, 2021; and Redl, 2020). In the first step, subjective expert judgement is used to classify variables as either macroeconomic or financial. In the second step, these macroeconomic and financial series are used to econometrically estimate macroeconomic and financial uncertainty.

Estimates of macroeconomic and financial uncertainty play a vital role in current research with those developed by Jurado, Ludvigson and Ng (2015) being used, for example, to forecast US recessions (Ercolani and Natoli, forthcoming), disentangle news and financial uncertainty shocks (Cascaldi-Garcia and Galvao, forthcoming) and consider whether macroeconomic uncertainty influences household consumption choices (Nam, Lee and Jeon, forthcoming). However, despite the importance of these estimates, it remains unclear how dozens of series should be classified in the first step. These include money supply, credit extension, exchange rates, interest rates and stock price indices. If 10 such variables are included in analysis, this means there are 10^2 possible classification schemes which the researcher must choose between. This has led to different studies classifying key variables in different ways, examples of which are provided in Table 1.

Table 1: Federal Funds Rate and US Share Price Index Classification in Selected Studies

	Federal Funds Rate	US Share Price Index
Ludvigson, Ma and Ng (forthcoming)	Macroeconomic	Macroeconomic
Carriero, Clark and Marcellino (2018)	Macroeconomic	Financial
Redl (2020)	Financial	Financial

We address the issues above by developing Bayesian econometric methods which can jointly classify variables and produce estimates of uncertainty and its impact on the economy. Our starting point is the popular Stochastic Volatility in Mean Vector Autoregression (SVMVAR). VARs are commonly used by economists to jointly model macroeconomic and financial variables. Additionally, in the SVMVAR, the volatility of each variable in the model has a common component and idiosyncratic component both of which are time-varying. Macroeconomic uncertainty is modelled as the common component driving the volatilities of all macroeconomic variables. Similarly, financial uncertainty is modelled as the common component driving the volatilities of all financial variables. Importantly, both macroeconomic and financial uncertainty affect the levels of all included variables. It is this feature which allows for uncertainty and its impact on the economy to be jointly estimated.

In our paper, we determine the appropriate specification of large SVMVARs by allowing the classification of selected variables to be estimated within the model. Our novel Monte Carlo Markov Chain (MCMC) algorithm has three distinctive features. First, we distinguish between macroeconomic variables, financial variables and unclassified variables, allowing the algorithm to determine whether each unclassified variable should be included in the macroeconomic block or financial block. Second, in the full unrestricted version of our SVMVAR model, for each unclassified variable, we allow for Markov switching time-varying classification. In doing so, we recognise that a variables classification may change over time as the structure of the economy evolves or experiences major crises. Third, our computationally efficient algorithm which extends Cross, Hou, Koop and Poon (2021) is scalable, allowing us to consider a larger cross-section of variables than previously possible. Although our paper disentangles macroeconomic and financial uncertainty, our flexible model can be used in any context where it is difficult to distinguish between different types of uncertainty.

Other studies which deploy SVMVAR methods include Cross, Hou and Poon (2018) who examine the effects of domestic and foreign uncertainty in three small open economies and Carriero, Clark and Marcellino (2020) who examine comovements in macroeconomic uncertainty across advanced economies. The latter uses two datasets with one consisting of 67 quarterly variables for the US, Euro Area and UK. Notably, measures of money supply, stock price indices and interest rates are all treated as macroeconomic variables when estimating macroeconomic uncertainty. The study most closely related to ours is Carriero, Clark and Marcellino (2018, CCM) who develop state-of-the-art Bayesian methods to estimate a 30 variable SVMVAR to measure the effect of macroeconomic and financial uncertainty on the US economy. This time, the federal funds rate is treated as a macroeconomic variable while the stock price index and credit spread are treated as financial variables. Our study builds on this literature by investigating how the classification of variables influences the measurement of macroeconomic and financial uncertainty.

Our empirical work contrasts different model specifications. For models 1-3 we use the same 30 variable dataset as CCM but extend it to December 2020, including the COVID-19 pandemic in our sample. Model 1, the SVMVAR-TVC, is the full unrestricted SVMVAR model with time-varying classification. Model 2, the SVMVAR-TIC, is the same as model 1 but imposes the restriction that we have time-invariant classification. Model 3, the SVM-VAR, does not allow for classification at all and simply imposes the restriction that all variables are classified by the researcher as either macroeconomic or financial as in CCM.

Our results suggest that without data-driven classification, macroeconomic uncertainty tends to be underestimated particularly during the global financial crisis. The adverse effect of macroeconomic uncertainty on real activity and the stock market also tend to be underestimated. The responses of the policy rate and credit spread are also stronger when we introduce classification of variables. We also find that financial uncertainty tends to be underestimated without data-driven classification. However, its effects on headline macroeconomic variables such as industrial production and the unemployment rate tend to be overestimated. Like CCM, we also find that uncertainty is not the dominant driver of fluctuations in macro variables or asset returns, however, when we introduce data-driven classification, the relative contribution of macroeconomic uncertainty in particular increases during crisis periods.

The paper is organized as follows. Section 2 introduces our SVMVAR model which can distinguish between macroeconomic and financial uncertainty. Section 3 discusses our empirical analysis using the CCM dataset. Section 4 concludes. An online supplementary appendix contains additional details on the data, MCMC algorithm and results.

2 A Model to Distinguish between Macroeconomic and Financial Uncertainty

Bloom (2014) refers to uncertainty as an amorphous concept. It is undoubtedly important for the economy, but hard to measure in practice. Many measures of uncertainty have been proposed including those based on: asset market variables such as the VIX (Bloom, 2007); the dispersion of surveys of businesses (Bachmann, Elstner and Sims, 2013) and forecasters (Scotti, 2016); and the number of terms relating to uncertainty in newspapers (Baker, Bloom and Davis, 2016). However, a growing literature equates uncertainty with Stochastic Volatility (SV), the time-varying second moments of time series variables (see CCM; Berger, Grabert and Kempa, 2016; Mumtaz and Theodoridis, 2017; Mumtaz and Musso, 2019 among many others). Within this strand of the literature, the SVMVAR is a popular model. The SVMVAR jointly estimates uncertainty (through the SV in the errors) and produces an estimate of their impact on the economy (by allowing the SVs to enter the conditional mean of the VAR). In this section, we describe a new SVMVAR which allows for uncertainty in the way variables are classified and develops Bayesian MCMC methods for its estimation.

2.1 The SVMVAR

Let $y_t^m = (y_{1,t}^m, \ldots, y_{n_m,t}^m)$ be an $n_m \times 1$ vector of macroeconomic variables, $y_t^f = (y_{1,t}^f, \ldots, y_{n_f,t}^f)$ be an $n_f \times 1$ vector of financial variables, and $y_t^u = (y_{1,t}^u, \ldots, y_{n_u,t}^u)$ be an $n_u \times 1$ vector of unclassified variables that might belong to either the macro or financial block. The main contribution of this paper lies in the treatment of these unclassified variables. If there are no unclassified variables then we obtain a conventional SVMVAR similar to, for example, Cross, Hou, Koop and Poon (2021).

We consider the following SVMVAR model, denoting $\mathbf{y}_t = (y_t^{m'}, y_t^{u'}, y_t^{f'})'$ and $n = n_m + n_f + n_u$ and :

$$\mathbf{B}_0\mathbf{y}_t = \sum_{i=1}^{p} \mathbf{B}_i\mathbf{y}_{t-i} + \sum_{j=0}^{q} \mathbf{A}_j\mathbf{h}_{t-j} + \boldsymbol{\epsilon}_t^y, \quad \boldsymbol{\epsilon}_t^y \sim \mathcal{N}(\mathbf{0}, \mathbf{U}_t), \tag{1}$$

where $\mathbf{B}_0$ is an $n \times n$ lower triangular matrix with ones on its diagonal[1], $\mathbf{B}_1, \ldots, \mathbf{B}_p$ are $n \times n$ coefficient matrices and $\mathbf{h}_t = (h_{m,t}, h_{f,t})'$ is a vector of common log-volatilities for the macro and financial variables which are described below. In our application, we will set $p = 6$ and $q = 2$ as in CCM. The direct impacts of the common log-volatilities on all n variables are captured by the $n \times 2$ coefficient matrices $\mathbf{A}_0, \ldots, \mathbf{A}_q$.

The covariance matrix $\mathbf{U}_t$ is specified as a diagonal matrix:

$$\mathbf{U}_t = \begin{pmatrix} \Omega_{m,t} & & \\ & \Omega_{u,t} & \\ & & \Omega_{f,t} \end{pmatrix}, \tag{2}$$

where the volatilities of the macro, financial and unclassified variables are respectively defined as $\Omega_{m,t} = \text{diag}(e^{\omega_{1,t}^m}, \ldots, e^{\omega_{n_m,t}^m})$, $\Omega_{f,t} = \text{diag}(e^{\omega_{1,t}^f}, \ldots, e^{\omega_{n_f,t}^f})$ and $\Omega_{u,t} = \text{diag}(e^{\omega_{1,t}^u}, \ldots, e^{\omega_{n_u,t}^u})$.

For variables in the macro and financial blocks, the time-varying log-volatilities are specified as:

$$\omega_{i,t}^m = \eta_{i,t}^m + h_{m,t}, \quad i = 1, \ldots, n_m, \tag{3}$$

$$\omega_{i,t}^f = \eta_{i,t}^f + h_{f,t}, \quad i = 1, \ldots, n_f, \tag{4}$$

[1]We stress that we are working in structural VAR form to speed up computation, not in order to impose structural identification. Carriero, Clark and Marcellino (2019) shows how this allows for equation-by-equation estimation which greatly speeds up computation. Section 3.1 of their paper discusses the role of variable ordering and shows that the MCMC draws of VAR coefficients are invariant to ordering.

where the variables $\eta_{i,t}^m$ and $\eta_{i,t}^f$ capture idiosyncratic volatility components associated with the ith macro and financial variables, respectively. The common log-volatilities for all the variables in the macro and financial blocks are given by $h_{m,t}$ and $h_{f,t}$. These are our measures of macro and financial uncertainty.

For the unclassified variables, the log-volatilities are specified as:

$$\omega_{i,t}^u = \eta_{i,t}^u + h_{s_{i,t},t}, \quad i = 1, \ldots, n_u, \tag{5}$$

where $s_{i,t} \in \{m, f\}$ is the indicator variable for the ith unclassified variable which follows a Markov switching process with transition probability $p(s_{i,t} = k|s_{i,t-1} = l) = p_{l,k}^i$, $k, l \in \{m, f\}$. Note that the ith unclassified variable is again defined as a sum of two components. The first component is the idiosyncratic component denoted as $\eta_{i,t}^u$, and the second component is determined by the indicator variable $s_{i,t}$ as either the common log-volatility of the macro block or of the financial block. For example, if $s_{i,t} = m$, then $h_{s_{i,t},t} = h_{m,t}$, which indicates that the ith unclassified variable belongs to the macro block at time t. This specification not only allows for the classification of each uncertainty variable to either the macro or financial block, but does so in a time varying fashion. So it is possible that a variable switches from the financial block to the macro block (or vice versa). This allows us to investigate a range of interesting possibilities. For instance, the volatility of a variable may appear like a financial volatility in normal times but like a macro volatility in times of crisis.

We follow CCM in assuming that our measures of uncertainty depend not only on past uncertainty but also past values of the variables themselves. That is, we assume the common log-volatilities evolve as follows:

$$h_{m,t} = \sum_{i=1}^{p^m} \phi_{m,i} h_{m,t-i} + \delta_m' y_{t-1}^m + \epsilon_{m,t}, \quad \epsilon_{m,t} \sim \mathcal{N}(0, \omega_m^2) \tag{6}$$

$$h_{f,t} = \sum_{i=1}^{p^f} \phi_{f,i} h_{f,t-i} + \delta_f' y_{t-1}^f + \epsilon_{f,t}, \quad \epsilon_{f,t} \sim \mathcal{N}(0, \omega_f^2) \tag{7}$$

The idiosyncratic log-volatilities are assumed are standard SV processes:

$$\eta_{i,t}^m = \mu_{m.i} + \rho_{m,i} \eta_{i,t-1}^f + \epsilon_{i,t}^m, \quad \epsilon_{i,t}^m \sim \mathcal{N}(0, \sigma_{m,i}^2), \quad i = 1, \ldots, n_m, \tag{8}$$

$$\eta_{i,t}^f = \mu_{f.i} + \rho_{f,i} \eta_{i,t-1}^f + \epsilon_{i,t}^f, \quad \epsilon_{i,t}^f \sim \mathcal{N}(0, \sigma_{f,i}^2), \quad i = 1, \ldots, n_f, \tag{9}$$

$$\eta_{i,t}^u = \mu_{u.i} + \rho_{u,i} \eta_{i,t-1}^u + \epsilon_{i,t}^u, \quad \epsilon_{i,t}^u \sim \mathcal{N}(0, \sigma_{u,i}^2), \quad i = 1, \ldots, n_u, \tag{10}$$

2.2 Priors

2.2.1 Prior for $(\mathbf{B}_1, \ldots, \mathbf{B}_p)$ and $(\mathbf{A}_0, \ldots, \mathbf{A}_q)$

We use a Minnesota-type prior for the VAR coefficients $(\mathbf{B}_1, \ldots, \mathbf{B}_p)$. It is Gaussian with the prior means of the VAR coefficients being set to 0. The prior variance of the coefficient corresponding to the lag l of variable j in equation i is $\frac{\pi_1}{l^2}$ for $i = j$ and $\frac{\pi_1 \pi_2}{l^2} \frac{d_i}{d_j}$ otherwise, for $l = 1, \ldots, p$. This reflects the usual Minnesota prior property that own lags are likely to be more important than other lags and more recent lags are likely to be more important than more distant lags. The latter property is also reasonable for the coefficients on the uncertainty measures. Accordingly, we adopt the same Minnesota prior form for them except that we assume the same prior variance for the coefficients on the financial and macro uncertainty measures. In particular, we set the prior means of the coefficient matrices of the log-volatilities $(\mathbf{A}_0, \ldots, \mathbf{A}_q)$ to 0. The prior variances of the coefficients for the lag l uncertainty measures in equation i are set to $\frac{\pi_3 d_i}{(l+1)^2}$, for $l = 0, \ldots, q$. Following standard Minnesota prior practice, we set the prior hyperparameters d_i to be the residual variances of AR(p) models for variable i for $i = 1, .., n$. In our application, we set the global shrinkage hyperparameter $\pi_1 = 0.1^2$, the cross-variable shrinkage hyperparameter $\pi_2 = 0.5^2$, and the overall shrinkage hyperparameter for the impact matrices of the volatility $\pi_3 = 10$. For each intercept in $\mathbf{b}$, we set its prior mean to 0 and variance to 100. These choices are similar to those used by CCM.

2.2.2 Prior for the other parameters

For the parameters in the volatility processes we use relatively non-informative priors. In particular, for the log-volatility processes defined in equations (8)-(10), we set each initial condition to have a truncated Gaussian prior with mean of 0 and variance of 5. For each AR coefficient, we assume a Gaussian prior with mean 0.9 and variance 0.2^2 restricted to the interval $(-1, 1)$ to guarantee stationarity. For the variance of each disturbance term, we use an inverse-gamma prior that has a mean 0.03 and degree of freedom 10. The AR intercepts are Gaussian distributed with means set at $d_1, \dots, d_n$ respectively and variance 10 for each.

For the common log-volatilities defined in (6) and (7), each of the initial conditions has a Gaussian prior and we set its mean and variance at 0 and 5 respectively. For the AR coefficients, we use Gaussian priors and set the mean of the first-order lag coefficient at 0.9, and the second-order lag coefficient at 0. The prior variances of the AR coefficients are set at 0.2^2. We assume Gaussian priors for coefficients (δ_m, δ_f) with means 0 and variances 0.4^2. The variances of the disturbance terms follow inverse-gamma distributions with mean 0.01 and degrees of freedom 10.

The indicator variables $s_{i,t}$ depend on the Markov transition probabilities that are assumed to have Dirichlet priors:

$$(p^i_{m,m}, p^i_{m,f}) \sim \mathcal{D}(\alpha^i_{m,m}, \alpha^i_{m,f}), \quad (p^i_{f,m}, p^i_{f,f}) \sim \mathcal{D}(\alpha^i_{f,m}, \alpha^i_{f,f}), i = 1, \dots, n_u.$$

We make relatively noninformative choices for the prior hyperparameters of $(\alpha^i_{m,m}, \alpha^i_{f,f}) = (10, 10)$ and $(\alpha^i_{m,f}, \alpha^i_{f,m}) = (1, 1)$, $i = 1, \dots, n_u$.

2.3 Posterior Inference in the SVMVAR

In this sub-section we provide an informal description of our MCMC algorithm which allows for Bayesian inference in our large SVMVARs. Full details of the algorithm are provided in the Technical Appendix.

Bayesian inference in VARs with stochastic volatility is typically done using MCMC methods involving the auxiliary mixture sampler of Kim, Shephard and Chib (1998). However, once stochastic volatility is added to the mean the auxiliary mixture sampler can no longer be used. Instead papers such as CCM use particle filtering and, in particular, the particle Gibbs with ancestor sampling algorithm of Lindsten, Jordan and Schon (2014). However, particle filtering can be computationally burdensome in large models and can suffer from particle degeneracy problems. These points demonstrated in Cross, Hou, Koop and Poon (2021) who develop an MCMC algorithm which involves a Metropolis-Hastings step to draw the log-volatilities. This involves a Gaussian candidate generating density with variance depending on the Hessian of the conditional posterior of the log-volatilities. It exploits the fact that this Hessian is block-banded. Band and sparse matrix algorithms can be exploited to allow for efficient computation even in large SVMVARs. This opens the door to Bayesian estimation of very large SVMVARs such as those considered in this paper.

In the present paper, we extend the methods of Cross, Hou, Koop and Poon (2021) to allow for variables whose classification is uncertain. Conditional on knowing the way the variables are classified (i.e. conditional on $s_{i,t}$ for $i = 1, \ldots, n_u$), we can reorder the equations of the unclassified variables to group them appropriately with either the macro or financial variables. The methods of Cross, Hou, Koop and Poon (2021) can then be applied directly to sample the log-volatilities $(\mathbf{h}_1, \ldots, \mathbf{h}_T)$. Given $(\mathbf{h}_1, \ldots, \mathbf{h}_T)$, draws of the indicator variable $s_{i,t}$ and the Markov transition probabilities $(p^i_{m,m}, p^i_{m,f}, p^i_{f,m}, p^i_{f,f})$ can be obtained using the algorithm of Chib (1996). More details about the MCMC sampler are given in the Technical Appendix.

In our empirical results, we present evidence on the importance of adding the log-volatilities to the conditional mean of the VAR equations, recalling that Stochastic Volatility in the Mean (SVM) allows uncertainty to directly affect the economy. We do this by computing Bayes factors which quantify the extent to which one model is favoured over another. Specifically, we compute the Bayes factor in favour of a VAR model without SVM against its counterpart with SVM using the Savage-Dickey density ratio:

$$\mathrm{BF} = \frac{p\left(\mathbf{A}=0|\mathbf{y}\right)}{p\left(\mathbf{A}=0\right)},$$

which can be calculated using the MCMC draws.

We also present evidence on the importance of allowing for classification uncertainty for the selected unclassified variables. We do this by computing the Bayes factor in favour of the SVMVAR model against the SVMVAR-TIC and SVMVAR-TVC using the Savage-Dickey density ratio:

$$\mathrm{BF} = \frac{p\left(s_{1,t}=m, s_{2,t}=f, s_{3,t}=f, t=1,\ldots,T|\mathbf{y}\right)}{p\left(s_{1,t}=m, s_{2,t}=f, s_{3,t}=f, t=1,\ldots,T\right)},$$

which again can be calculated using the MCMC draws.

3 Empirical Analysis Using the CCM Dataset

3.1 The Data

We use an updated version of the dataset considered in CCM. The data consists of 30 monthly US variables and the sample spans January 1960 to December 2020. The macro and financial variables are selected as in CCM, except that we treat the federal funds rate (FEDFUNDS), S&P 500 and the credit spread (Baa-10y Treasury, BAAT10Y) as unclassified variables. Therefore, we have 17 macro variables, 10 financial variables and 3 unclassified variables. Our choice of unclassified variables is motivated by CCM. Although they classify the federal funds rate as a macro variable and the S&P 500 and credit spread as financial variables, they acknowledge some uncertainty over this choice since "this specification reflects some choice as to what constitutes a macroeconomic variable rather than a financial variable" (page 804). They then argue that the federal funds rate should be treated as a macro variables since it is the instrument of monetary policy. That said, studies as recent as Redl (2020) instead consider the policy rate a financial variable. For the S&P 500 and credit spread "the distinction between macro and finance is admittedly less clear" (page 805) but ultimately CCM place them in the group of financial variables. The complete list of variables and their transformations are given in the Data Appendix. Like CCM, the model is estimated with standardised data.

3.2 The Importance of Classification

The unrestricted version of our model, the SVMVAR-TVC, which allows for time varying classification is given in (1) through (10). We also consider a version of our model, the SVMVAR-TIC, which does not allow for this time variation in classification. This is obtained by restricting $p^i_{l,k} = 0$ for $l \neq k$. Finally we consider a model, the SVMVAR, where the variables are classified as in CCM. The latter two models are restricted versions of our model and their priors are the same as in the unrestricted model, conditional on the restriction holding.

We first turn to our attention to our Bayes factors presented in Table 2. Focussing on the top section of the table (rows 1-4), note that the smaller the Bayes factor, the greater the statistical evidence in favour of models with SVM. The results strongly suggest that the addition of SVM is important and that time-varying classification adds additional benefits. Turning to the bottom section of the table (rows 5-7), the smaller the Bayes factor, the greater the statistical evidence in favour of models with data-driven classification. The results provide strong statistical evidence that introducing data-driven classification improves model fit.

Table 2: The Importance of SVM and Classification

Bayes factor comparing VARs to counterparts with SVM	
CCM classification scheme	0.00
Time invariant classification	0.00
Time varying classification	0.00
Bayes factor comparing SVMVAR to counterparts with estimated classification	
SVMVAR-TIC	0.00
SVMVAR-TVC	6.41×10^{-31}

Having established the importance of data-driven classification, we now examine our results in further detail. Using the SVMVAR-TIC, we find that the federal funds rate and S&P 500 are classified as financial variables with posterior probability near 1 while the credit spread is classified as a macroeconomic variable. This does not coincide with the classification scheme selected by CCM or others in the literature including Ludvigson, Ma and Ng (forthcoming) and Redl (2020).

As shown in Figure 1, using the SVMVAR-TVC, however, we also find evidence which suggests that the classification of variables evolves over time. Focussing first on the federal funds rate, we find that our model successfully detects that classification should become stable after the zero lower bound is reached as we can see following the global financial crisis. We also uncover that the probability that the federal funds rate should be included as a macroeconomic variable rises during significant periods of monetary policy tightening and loosening. For example, the probability that it should be classified as a macroeconomic variable rises significantly during the early 1980s when the Federal Reserve engaged in monetary policy tightening to fight rising inflation. Similarly, during the coronavirus pandemic, as monetary policy hit the zero lower bound following a period of normalisation we again see a stark change. The probability that the federal funds rate should be classified as a macroeconomic variable peaks as monetary policy loosens and again reaches the zero lower bound.

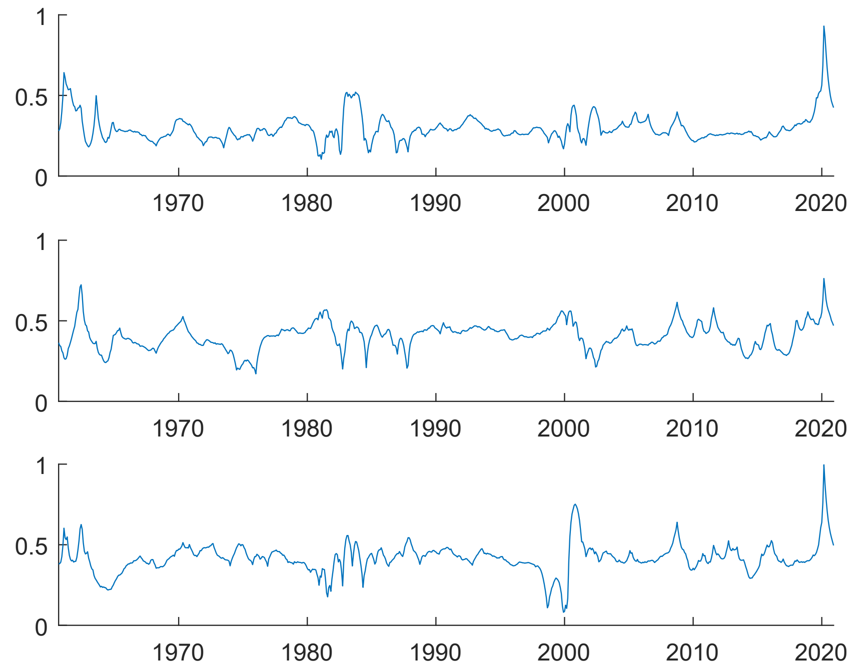

Figure 1: The estimated posterior probability, $p(s_{i,t} = m|\mathbf{y})$, that the federal funds rate (top), S&P 500 (middle) and credit spread (bottom) are classified as macro variables

If we now consider the S&P 500, as expected, the probability that it is classified as a financial variable is high during times of financial turmoil such as the 1962 flash crash, 1973-1974 crash, Black Monday, the 2001 dot com bust and the global financial crisis from 20072008. Last, if we turn to the credit spread, the probability that it is classified as a macroeconomic variable is high through the early 1980s through to the mid 1990s as the economy experiences positive credit spread shocks (Prieto, Eickmeier, Marcellino, 2016) fuelled by financial deregulation, giving businesses and households increased credit access, leading to increased economic performance (Justiniano and Primiceri, 2008). After this period of change, the probability that the credit spread is classified as a macroeconomic variable declines until the recession of 2001. It's importance as a macroeconomic variable then surges, something which is also seen during the global financial crisis and coronavirus pandemic. These episodes reflect businesses increased borrowing cost which, in turn, reduces investment and economic growth (Arrelano, Bai and Kehoe, 2019; Christiano, Motto and Rostagno, 2014; Gilchrist, Sim and Zakrajek, 2014).

3.3 Estimates of Macro and Financial Uncertainty

In this sub-section we present and discuss our estimates of macro and financial uncertainty obtained from our three models. In particular, we emphasise the differences observed when we introduce data-driven classification. Our estimates are presented in Figure 2. Since our estimates of uncertainty are considered random variables we can also plot credible intervals around them. These are presented in the Supplementary Figures Appendix, however, we find that the degree of uncertainty around our estimates does not change substantially across specifications.

While the broad trends in the estimates are similar across models, there are important differences in the magnitude of our uncertainty measures particularly during economic crises when uncertainty is high. We also find that throughout the sample the SVMVAR-TVC provides slightly higher estimates of macro and financial uncertainty than the SVMVAR-TIC.

We find that without data-driven classification, macroeconomic uncertainty tends to be underestimated during crises and subsequent recessions. This is most apparent during the global financial crisis and great recession from 2007-2009 as well as during the coronavirus pandemic of 2020. However, the underestimation of macroeconomic uncertainty can also be seen during other recessionary periods including the mid 1970s, mid 1980s, 1990 and the early 2000s. In contrast, without data-driven classification financial uncertainty is consistently over-estimated throughout the sample. Prominent examples of this include the 1980s recession and the global financial crisis. We can also see that in some cases when data-driven classification is not used, financial uncertainty can peak too early. This occurs during the mid 1970s recession and the global financial crisis.

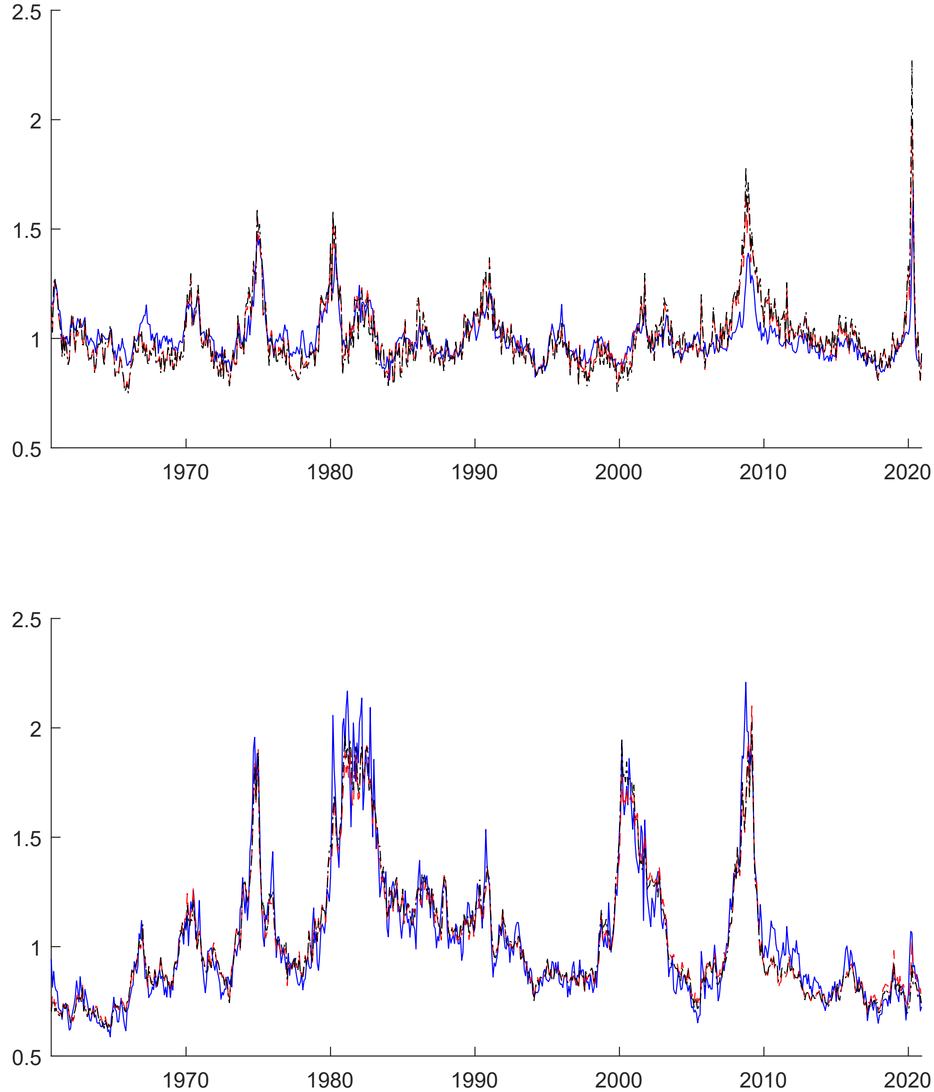

Figure 2: Posterior medians of macro uncertainty $e^{0.5h_{m,t}}$ (top panel) and financial uncertainty $e^{0.5h_{m,t}}$ (bottom panel) for SVMVAR (blue solid lines), SVMVAR-TIC (red dashed lines) and SVMVAR-TVC (black dashed dotted lines)

3.4 Impulse Responses and Historical Decompositions

To compute our impulse response functions, we follow standard practise (CCM and Banbura, Giannone and Reichlin, 2010 among many others) and adopt a Cholesky identification scheme where macroeconomic and financial variables are considered to be slow- and fast-moving respectively. We then reverse standardisation and accumulate the impulse response functions to obtain the responses in levels or log levels. Since our focus is on the importance of classification, for brevity, we compare the posterior medians of the impulse response functions obtained from our three different models. However, impulse response functions with credible intervals can be found in the Supplementary Figures Appendix. Reassuringly, these indicate that we achieve the same level of precision when using the SVMVAR-TVC and SVMVAR-TIC as when using the SVMVAR.

In Figures 3 and **??** we consider the responses of our different variables to a one standard deviation macroeconomic uncertainty shock and financial uncertainty shock respectively. While we find that the sign of the responses are similar across models, there are important differences in the magnitude and thus relative importance of macroeconomic and financial uncertainty shocks.

While our results are broadly in line with CCM, they find that the effects of macroeconomic uncertainty on consumer prices and financial indicators are muted and imprecisely estimated. Across models, however, we detect a non-zero price decline in the PCE price index, reflecting a fall in the prices of goods and services purchased by consumers. We also detect a non-zero decline in the S&P 500, excess CRSP return and a range of industry returns.

Without data-driven classification, we tend to find that the effects of macroeconomic uncertainty are underestimated. This is particularly pronounced if we consider the macro variables employment, real consumer spending and real manufacturing and trade sales. If we consider our three unclassified variables, we also see a more pronounced monetary policy loosening, stock price decline and increase in the cost of borrowing (reflected in a more hump shaped credit spread) if we use data-driven classification. We see smaller differences if we consider real indicators such as housing starts and permits, the ISM index of new orders and weekly hours worked, variables which have a quick and sizable response to macroeconomic uncertainty.

Unlike the SVMVAR-TIC, results from the SVMVAR-TVC suggest that the adverse effect of macroeconomic uncertainty on industrial production, capacity utilisation and asset returns are also underestimated if data-driven classification is not used. The ability of the SVMVAR-TVC to detect this and uncover larger differences relative to the SVMVAR-TIC arises, in part, from allowing the federal funds rate to be classified as a macroeconomic variable during key events.

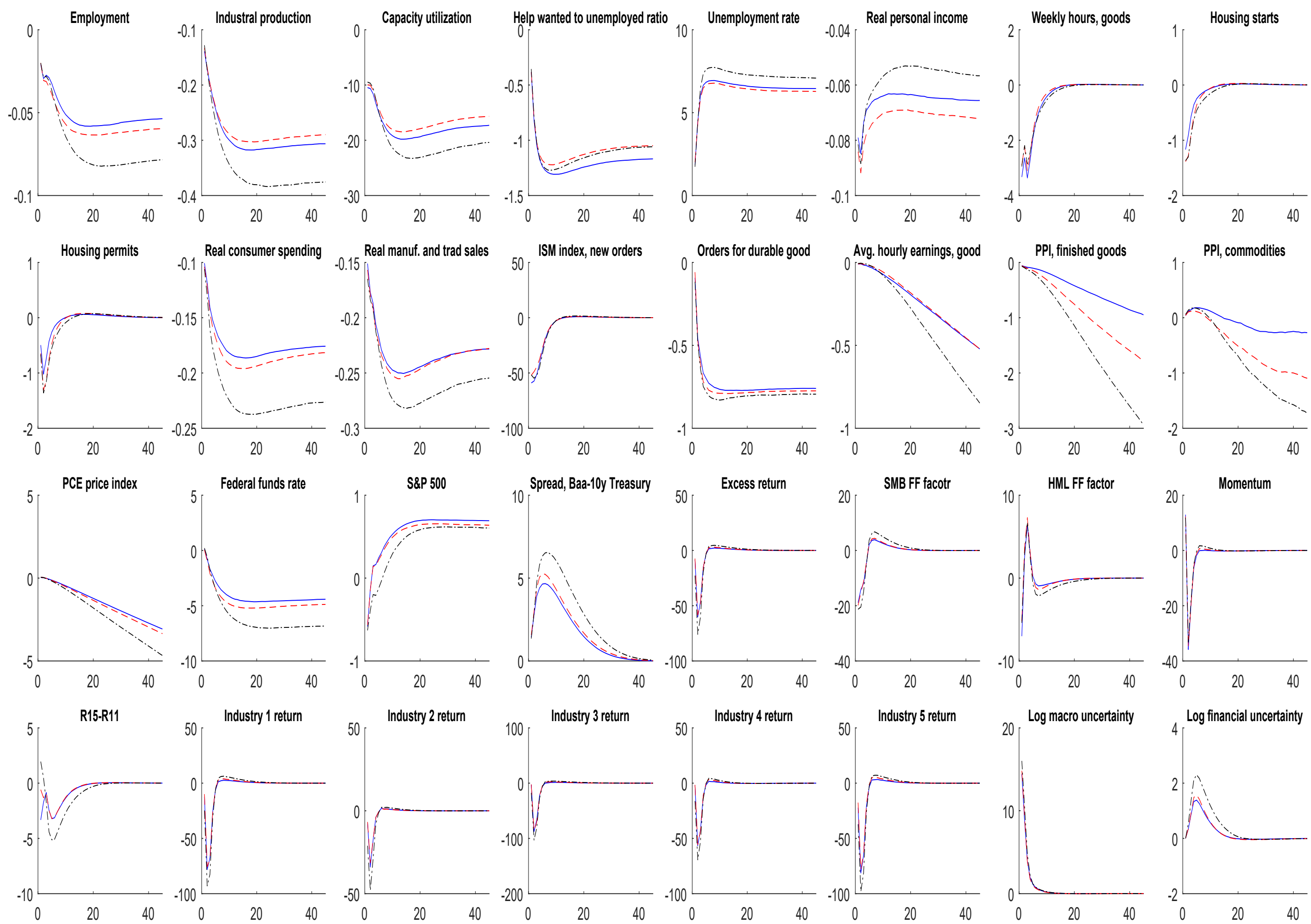

Figure 3: Posterior medians of impulse responses for one standard deviation shock to macro uncertainty: SVMVAR (blue solid lines), SVMVAR-TIC (red dashed lines) and SVMVAR-TVC (black dashed-dotted lines).

Turning to financial uncertainty, without data-driven classification, we tend to find that the effects of financial uncertainty on headline macroeconomic variables such as industrial production and the unemployment rate are overestimated. This is more pronounced if we consider the SVMVAR-TIC where the federal funds must be classified as a financial variable throughout the sample despite its importance as a macroeconomic variable when significant monetary policy changes take place. A similar trend is seen if we consider the ISM index of new orders. If we consider the effect of financial uncertainty on average hourly earnings, real consumer spending and, to a lesser extent, prices data-driven classification detects a larger adverse effect if a financial uncertainty shock occurs. If we consider our unclassified variables, data-driven classification does not conclusively point towards over or under estimation of the effects of financial uncertainty, however, if we favour the SVMVAR-TVC, it appears that financial uncertainty may have a larger effect on the federal funds rate and credit spread.

Before concluding, we consider the effect of data-driven classification on historical decompositions between January 2003 and December 2020 in Figures 4, 5 and 6 . Like CCM, our charts start in 2003 to ensure readability and to capture the role of uncertainty during the global financial crisis as well as the coronavirus pandemic. For brevity, we focus on our unclassified variables but historical decompositions for all variables can be found in the Supplementary Figures Appendix. Like CCM we find that uncertainty is not the dominant driver of fluctuations in any of our variables but does contribute significantly to variation in real indicators of economic activity. Uncertainty also contributes sizably to the credit spread, federal funds rate and, to a lesser extent, inflation. We find that the S&P 500 and other asset returns are least affected by uncertainty.

Without data-driven classification, we find that the relative contribution of macroeconomic uncertainty, particularly during the Great Recession and, to some extent, the coronavirus pandemic, is underestimated with the SVMVAR-TVC detecting a slightly larger contribution than the SVMVAR-TIC. This is most pronounced if we consider the credit spread and real indicators of economic activity such as housing starts, housing permits and the ISM index of new orders. In contrast, the contribution of financial uncertainty is not substantially altered using data-driven classification.

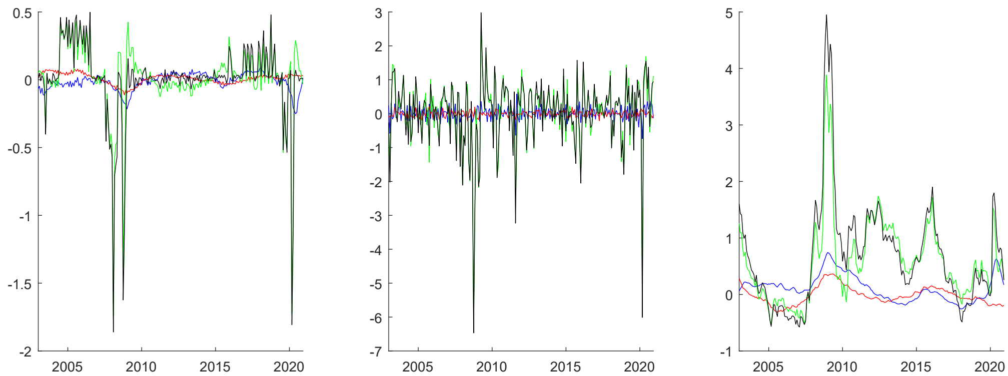

Figure 4: SVMVAR: Historical decompositions of FEDFUNDS (left), S&P 500 (middle) and BAAT10Y (right) for each model for 2003M1 - 2020M12: the actual data (black lines), macro shocks (blue lines), financial shocks (red lines) and VAR shocks (green lines).

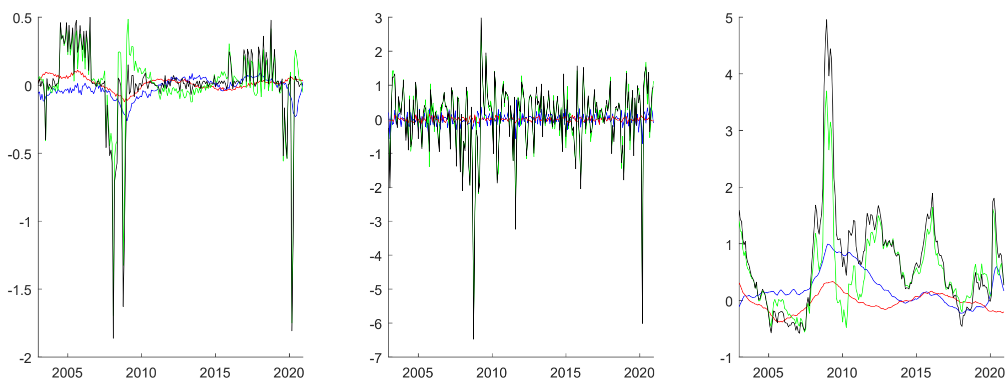

Figure 5: SVMVAR-TIC: Historical decompositions of FEDFUNDS (left), S&P 500 (middle) and BAAT10Y (right) for each model for 2003M1 - 2020M12: the actual data (black lines), macro shocks (blue lines), financial shocks (red lines) and VAR shocks (green lines).

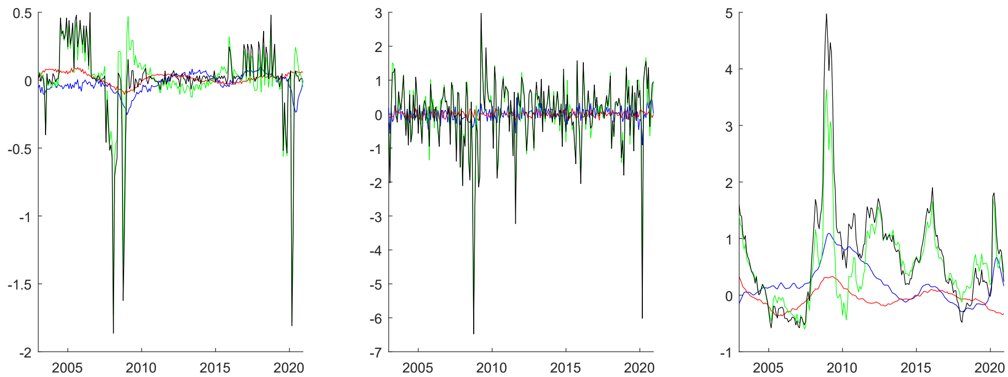

Figure 6: SVMVAR-TVC: Historical decompositions of FEDFUNDS (left), S&P 500 (middle) and BAAT10Y (right) for each model for 2003M1 - 2020M12: the actual data (black lines), macro shocks (blue lines), financial shocks (red lines) and VAR shocks (green lines).

4 Conclusion

In this paper, we determine the appropriate specification of large Stochastic Volatility in Mean Vector Autoregressions (SVMVARs). Our novel MCMC algorithm has three distinctive features. First, we distinguish between macroeconomic variables, financial variables and unclassified variables, allowing the algorithm to determine whether each unclassified variable should be included in the macroeconomic block or financial block. Second, in the full unrestricted version of our SVMVAR model, for each unclassified variable, we allow for Markov switching time-varying classification. In doing so, we recognise that a variables classification may change over time as the structure of the economy evolves or experiences major crises. Third, our computationally efficient algorithm is scalable, allowing us to consider a larger cross-section of variables than previously possible. Although our paper disentangles macroeconomic and financial uncertainty, our flexible model can be used in any context where it is difficult to distinguish between different types of uncertainty.

Our results suggest that without data-driven classification, macroeconomic uncertainty tends to be underestimated particularly during the global financial crisis. The adverse effect of macroeconomic uncertainty on real activity and the stock market also tend to be underestimated. The responses of the policy rate and credit spread are also stronger when we introduce classification of variables. We also find that financial uncertainty tends to be underestimated without data-driven classification. However, its effects on headline macroeconomic variables such as industrial production and the unemployment rate tend to be overestimated. Like previous studies, we also find that uncertainty is not the dominant driver of fluctuations in real activity or asset returns, however, when we introduce data-driven classification, the relative contribution of macroeconomic uncertainty, in particular, increases during crisis periods.

References

[1] Arellano, C., Bai, Y. and Kehoe, P.J., 2019. Financial frictions and fluctuations in volatility. *Journal of Political Economy*, 127(5), pp.2049-2103.

[2] Bachmann, R., Elstner, S. and Sims, E.R., 2013. Uncertainty and economic activity: Evidence from business survey data. *American Economic Journal: Macroeconomics*, 5(2), pp.217-49.

[3] Baker, S.R., Bloom, N. and Davis, S.J., 2016. Measuring economic policy uncertainty. *The quarterly journal of economics*, 131(4), pp.1593-1636.

[4] Banbura, M., Giannone, D. and Reichlin, L., 2010. Large Bayesian vector auto regressions. *Journal of applied Econometrics*, 25(1), pp.71-92.

[5] Berger, T., Grabert, S. and Kempa, B., 2016. Global and countryspecific output growth uncertainty and macroeconomic performance. *Oxford Bulletin of Economics and Statistics*, 78(5), pp.694-716.

[6] Bloom, N., 2014. Fluctuations in uncertainty. *Journal of Economic Perspectives*, 28, 153-176.

[7] Bloom, N., 2007. Uncertainty and the Dynamics of R& D. *American Economic Review*, 97(2), pp.250-255.

[8] Chib, S., 1996. Calculating posterior distributions and modal estimates in Markov mixture models. *Journal of Econometrics*, 75(1): 79-97.

[9] Carriero, A., Clark, T.E. and Marcellino, M., 2018. Measuring uncertainty and its impact on the economy. *Review of Economics and Statistics*, 100(5), pp.799-815.

[10] Carriero, A., Clark, T.E. and Marcellino, M., 2019. Bayesian vector autoregressions with stochastic volatility and non-conjugate priors. *Journal of Econometrics*, 212, 137-154.

[11] Carriero, A., Clark, T.E. and Marcellino, M., 2020. Assessing international commonality in macroeconomic uncertainty and its effects. *Journal of Applied Econometrics*, 35(3), pp. 73-293.

[12] Carriero, A., Clark, T.E., Marcellino, M.G. and Mertens, E., 2021. *Measuring uncertainty and its effects in the COVID-19 era.* CEPR Discussion Paper No. DP15965.

[13] Cascaldi-Garcia, D. and Galvao, A.B., forthcoming. News and uncertainty shocks. *Journal of Money, Credit and Banking.*

[14] Christiano, L.J., Motto, R. and Rostagno, M., 2014. Risk shocks. *American Economic Review*, 104(1), pp.27-65.

[15] Cross, J.L., Hou, C. and Poon, A., 2018. *International Transmission of Macroeconomic Uncertainty in Small Open Economies: An Empirical Approach.*

[16] Cross, J.L., Hou, C., Koop, G. and Poon, A., 2021. *Macroeconomic Forecasting with Stochastic Volatility in Mean VARs.*

[17] Ercolani, V. and Natoli, F., 2020. Forecasting US recessions: The role of economic uncertainty. Economics letters, 193, p.109302.

[18] Gilchrist, Simon, JaeW. Sim, and Egon Zakrajsek, *Uncertainty, Financial Frictions, and Investment Dynamics*, NBER working paper 20038 (2014).

[19] Jurado, K., Ludvigson, S.C. and Ng, S., 2015. Measuring uncertainty. *American Economic Review*, 105(3), pp.1177-1216.

[20] Justiniano, A. and Primiceri, G.E., 2008. The time-varying volatility of macroeconomic fluctuations. *American Economic Review*, 98(3), pp.604-41.

[21] Kim, S. and Shephard, N. and Chib, S., 1998. Stochastic volatility: likelihood inference and comparison with ARCH models. *The Review of Economic Studies*, 65, 361-393.

[22] Lindsten, F., Jordan, M. and Schon, T., 2014, Particle Gibbs with ancestor sampling, *Journal of Machine Learning Research*, 15, 2145-2184.

[23] Ludvigson, S.C., Ma, S. and Ng, S., forthcoming. Uncertainty and business cycles: exogenous impulse or endogenous response? *American Economic Journal: Macroeconomics*

[24] Mumtaz, H. and Theodoridis, K., 2017. Common and country specific economic uncertainty. *Journal of International Economics*, 105,

[25] Mumtaz, H. and Musso, A., 2019. The evolving impact of global, region-specific, and country-specific uncertainty. *Journal of Business & Economic Statistics*, pp.1-16.

[26] Nam, E.Y., Lee, K. and Jeon, Y., forthcoming. Macroeconomic uncertainty shocks and households consumption choice. *Journal of Macroeconomics.*

[27] Prieto, E., Eickmeier, S. and Marcellino, M., 2016. Time variation in macrofinancial linkages. *Journal of Applied Econometrics*, 31(7), pp.1215-1233.

[28] Redl, C., 2020. Uncertainty matters: evidence from close elections. *Journal of International Economics*, 124, p.103296.

[29] Scotti, C., 2016. Surprise and uncertainty indexes: Real-time aggregation of real-activity macro-surprises. *Journal of Monetary Economics*, 82, pp.1-19.

Online Appendices

A. Data Appendix

Table 3: Variables in Models 1-3

Macroeconomic	Financial	Unclassified
All employees: total nonfarm (Δ ln)	Excess return	S&P 500 (Δ ln, F)
Industrial production index (Δ ln)	SMB FF factor	Spread, Baa-10y Treasury (F)
Capacity utilization: manufacturing (Δ)	HML FF factor	Federal funds rate (Δ, M)
Help wanted-to-unemployed ratio (Δ)	Momentum factor	
Unemployment rate (Δ)	R15_R11	
Real personal income (Δ ln)	Industry 1 return	
Weekly hours: goods producing	Industry 2 return	
Housing starts (ln)	Industry 3 return	
Housing permits (ln)	Industry 4 return	
Real consumer spending (Δ ln)	Industry 5 return	
Real manufacturing trade sales (Δ ln)		
ISM: new orders index		
Orders for durable goods (Δ ln)		
Avg. hourly earnings, goods producing (Δ^2 ln)		
PPI, finished goods (Δ^2 ln)		
PPI, commodities (Δ^2 ln)		
PPI, price index (Δ^2 ln)		

For the variables transformed before inclusion in the model, we indicate the transformation in parentheses. For model 3, we classify the federal funds rate, credit spread and S&P 500 as in CCM as indicated in parentheses.

B. Technical Appendix

In this appendix we provide the key estimation steps for the SVMVAR-TIC and SVMVAR-TVC. Estimations for the other parameters are straightforward and readily available in Carriero, Clark and Marcellino (2018, 2020) and Cross, Hou, Koop and Poon (2021).

Drawing log-volatilities $(\mathbf{h}_1, \ldots, \mathbf{h}_T)$

We modify the methods proposed by Cross, Hou, Koop and Poon (2021) (CHKP) to sample the log-volatility. To illustrate the algorithm, we consider the following simplified version of our model:

$$\mathbf{y}_t = \sum_{i=0}^{q} \mathbf{A}_i \mathbf{h}_{t-q} + \boldsymbol{\epsilon}_t^y, \quad \boldsymbol{\epsilon}_t^y \sim \mathcal{N}(\mathbf{0}, \mathbf{U}_t). \tag{11}$$

Furthermore, we set the idiosyncratic log volatility components $\eta_{i,t}^k = 0$ for all i, j and $k \in \{m, f, u\}$ for expositional simplicity. Conditional on the indicator variables $s_{i,t}$, we can group the unclassified variables appropriately with the other macro and financial variables. For example, suppose we have $n_m = 2$, $n_f = 2$ and $n_u = 3$, then $n = n_m + n_f + n_u = 7$. Given the indicator variable $(s_{1,t}, s_{2,t}, s_{3,t}) = (f, f, m)$, we can construct a $n \times n$ selection matrix

$$\mathbf{Q}_t = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

and reorder equation (11) as

$$\widetilde{\mathbf{y}}_t = \sum_{i=0}^{q} \widetilde{\mathbf{A}}_i \mathbf{h}_{t-q} + \widetilde{\boldsymbol{\epsilon}}_t^y, \quad \boldsymbol{\epsilon}_t^y \sim \mathcal{N}(\mathbf{0}, \widetilde{\mathbf{U}}_t), \tag{12}$$

where $\widetilde{\mathbf{y}}_t = \mathbf{Q}_t\mathbf{y}_t$, $\widetilde{\mathbf{A}}_i = \mathbf{Q}_t\mathbf{A}_i$, $i = 0, \ldots, q$ and $\widetilde{\mathbf{U}}_t = \mathbf{Q}_t\mathbf{U}_t\mathbf{Q}_t'$. In equation (12), we can easily check that equations of the unclassified variables are reordered such that the first block of variables in $\widetilde{\mathbf{y}}_t$ are all macro variables, and the second block of variables are all financial variables. Let $\mathbf{h}_m = (h_{m,1}, \ldots, h_{m,T})$ and $\mathbf{h}_f = (h_{f,1}, \ldots, h_{f,T})$ and suppose we would like to sample $\mathbf{h}_m = (h_{m,1}, \ldots, h_{m,T})$. For notational simplicity, we will suppress the other conditioning arguments except for $\mathbf{h}_m$ and $\mathbf{h}_f$ in the following section. With the transformation given in equation (12), we can follow CHKP to derive the gradient $\mathbf{f}^m$ and the Hessian $\mathbf{G}^m$ of the log conditional likelihood $\log p(\mathbf{y}|\mathbf{h}_f, \mathbf{h_m})$.

Another component required for applying the method of CHKP is the gradient and the Hessian of the log prior density of $\mathbf{h}_m$. Given the prior specified in equation (6), it gives

$$\log p(\mathbf{h}_m) = -\frac{1}{2}(\mathbf{H}_m\mathbf{h}_m - \widetilde{\mathbf{h}}_m)'\mathbf{S}_m^{-1}(\mathbf{H}_m\mathbf{h}_m - \widetilde{\mathbf{h}}_m) + c_1,$$

where c_1 is the normalizing constant, $\widetilde{\mathbf{h}}_m$ is a $T \times 1$ vector with $\delta_m' y_{t-1}$ being its tth element, and

$$\mathbf{H}_m = \begin{pmatrix} 1 & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ -\phi_{m,1} & 1 & \ddots & \ddots & \cdots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & \ddots & \vdots & \vdots \\ -\phi_{m,p^m} & \ddots & -\phi_{m,1} & 1 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & -\phi_{m,p^m} & \cdots & -\phi_{m,1} & 1 \end{pmatrix}, \mathbf{S}_m = \begin{pmatrix} V_m & 0 & \cdots & 0 \\ 0 & \omega_m^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \omega_m^2 \end{pmatrix},$$

where V_m is the variance of the initial state and we set $V_m = 5$ in our empirical study. Thus, it gives

$$\frac{\partial \log p(\mathbf{h}_m)}{\partial \mathbf{h}_m} = -\mathbf{H}_m'\mathbf{S}_m^{-1}\mathbf{H}_m\mathbf{h}_m + \mathbf{H}_m'\mathbf{S}_m^{-1}\widetilde{\mathbf{h}}_m, \tag{13}$$

$$\frac{\partial^2 \log p(\mathbf{h}_m)}{\partial \mathbf{h}_m^2} = -\mathbf{H}_m'\mathbf{S}_m^{-1}\mathbf{H}_m. \tag{14}$$

Combining the results of the gradients and the Hessians of the log conditional likelihood $\log p(\mathbf{y}|\mathbf{h}_m, \mathbf{h}_f)$ and the log prior density $\log p(\mathbf{h}_m)$, the methods proposed by CHKP can be applied for efficiently drawing $\mathbf{h}_m$. As $\mathbf{h}_m$ and $\mathbf{h}_f$ appear in the model symmetrically, similar sampling procedure can also be applied for drawing $\mathbf{h}_f$. We refer the readers to Cross, Hou, Koop and Poon (2021) for more details.

Drawing the indicator variables for SVMVAR-TVC

Let $\mathbf{s}_i = (s_{i,1}, \dots, s_{i,T})$, $i = 1, \dots, n_u$. It can be seen that the indicator variables $(\mathbf{s}_1, \dots, \mathbf{s}_{n_u})$ are conditional independent and we can sample them one by one. Denote $\mathbf{y}_{1:k,t} = (y_{1,t}, \dots, y_{k,t})'$. For the ith unclassified variable observed at time t, we have

$$y_{i,t}^u = \sum_{j=1}^{p} \mathbf{b}_{j,n_m+i}\mathbf{y}_{t-j} + \sum_{k=0}^{q} \mathbf{a}_{k,n_m+i}\mathbf{h}_{t-k} - \mathbf{b}_{0,n_m+i}\mathbf{y}_{1:n_m+i-1,t} + \epsilon_{n_m+i,t}, \quad \epsilon_{n_m+i,t} \sim \mathcal{N}(0, e^{\omega_{i,t}^f}),$$

where $\mathbf{b}_{j,i}$ and $\mathbf{a}_{k,i}$ are the ith row of $\mathbf{B}_j$ and $\mathbf{A}_k$ respectively, and $\mathbf{b}_{0,n_m+i} = (b_{0,n_m+i,1}, \dots, b_{0,n_m+i,n_m+i-1})$ with $b_{0,j,k}$ is the (j,k)th element of $\mathbf{B}_0$. Thus, it leads to

$$y_{i,t}^u = \mathbf{x}_{i,t}\boldsymbol{\theta}_i + \epsilon_{n_m+i,t}, \quad \epsilon_{n_m+i,t} \sim \mathcal{N}(0, e^{\omega_{i,t}^f}), \tag{15}$$

where $\mathbf{x}_{i,t} = (\mathbf{y}_{t-1}', \dots, \mathbf{y}_{t-p}', \mathbf{h}_{t-1}', \dots, \mathbf{h}_{t-q}', \mathbf{y}_{1:n_m+i-1,t}')'$ and $\boldsymbol{\theta}_i$ is the parameter vector associated to $\mathbf{x}_{i,t}$. Given the $s_{i,t}$ and the log-volatility

$$\omega_{i,t}^u = \eta_{i,t}^u + h_{s_{i,t},t},$$

the conditional likelihood can be obtained by noting that $y_{i,t}^u$ is Normally distributed, which gives

$$p(y_{i,t}^u|s_{i,t},\boldsymbol{\theta}_i,h_{m,t},h_{f,t}) \propto e^{-\frac{1}{2}h_{s_{i,t}}} \exp\left(-\frac{1}{2}e^{-(\eta_{i,t}^u+h_{s_{i,t},t})}\left(y_{i,t}-\mathbf{x}_{i,t}\boldsymbol{\theta}_i\right)^2\right).$$

Here we have suppressed the other conditional arguments except for s_i, $h_{m,t}$,$h_{f,t}$ and $\boldsymbol{\theta}_i$ for notational convenience. Given the conditional likelihood and the Markov transition probability $p_{l,k}^i$, $k,l \in \{m,f\}$, the forward-backward algorithm of Chib (1996) can be applied for drawing the indicator variables $(s_{i,1},\dots,s_{i,T}), i=1,\dots,n_u$.

Drawing the transition probability

Given the indicator variable $\mathbf{s}_i = (s_{i,t},\dots,s_{i,T})$, $i=1,\dots,n_u$, and the Dirichlet priors on $(p_{m,m}^i,p_{m,f}^i)$ and $(p_{f,m}^i,p_{f,f}^i)$, it follows that

$$\begin{aligned}(p_{m,m}^i,p_{m,f}^i|\mathbf{s}_i) &\sim \mathcal{D}(\alpha_{m,m}^i+N_{m.m},\alpha_{m,f}^i+N_{m,f}),\\(p_{f,m}^i,p_{f,f}^i|\mathbf{s}_i) &\sim \mathcal{D}(\alpha_{f,m}^i+N_{f.m},\alpha_{f,f}^i+N_{f,f}),\end{aligned}$$

where $N_{k,l}=\sum_{j=1}^{T-1}\mathbf{1}(s_j=k,s_{j+1}=l)$, $k,l\in\{m,f\}$.

Drawing the indicator variables for SVMVAR-TIC

Note that the parameters in $(\mathbf{B}_0,\dots,\mathbf{B}_p,\mathbf{A}_0,\dots,\mathbf{A}_q)$ are conditionally independent across equations, thus we can sample the indicator variables $(s_1,\dots,s_{n_u})$ one by one. Furthermore, as the indicator variables for the SVMVAR-TIC are time-invariant, we can improve the MCMC efficiency by integrating out the parameters $(\mathbf{B}_0,\dots,\mathbf{B}_p,\mathbf{A}_0,\dots,\mathbf{A}_q)$. To see this, we first stack the ith unclassified variable over time and rewrite equation (15) as

$$\mathbf{y}_i^u = \mathbf{X}_i\boldsymbol{\theta}_i+\boldsymbol{\epsilon}_i, \quad \boldsymbol{\epsilon}_i \sim \mathcal{N}(\mathbf{0},\boldsymbol{\Sigma}_{s_i}),$$

where $\mathbf{y}_i^u = (y_{i,1}^u, \ldots, y_{i,T}^u)'$, $\mathbf{X}_i = (\mathbf{x}_{i,1}', \ldots, \mathbf{x}_{i,T}')'$, $\boldsymbol{\epsilon}_i = (\epsilon_{n_m+i,1}, \ldots, \epsilon_{n_m+i,T})'$ and $\boldsymbol{\Sigma}_{s_i} = \mathrm{diag}(e^{\eta_{i,1}^u + h_{s_i,1}}, \ldots, e^{\eta_{i,T}^u + h_{s_i,T}})$.

Note that $\boldsymbol{\theta}_i$ has a Normal prior, denoted as $\boldsymbol{\theta}_i \sim \mathcal{N}(\boldsymbol{\theta}_{0,i}, \mathbf{V}_{\theta_i})$. We can integrate out $\boldsymbol{\theta}_i$ as follows:

$$
\begin{aligned}
&p(s_i|\mathbf{y}_i^u) \propto \int p(\mathbf{y}_i^u|\boldsymbol{\theta}_i, s_i)p(\boldsymbol{\theta}_i)p(s_i)d\boldsymbol{\theta}_i \\
&\propto p(s_i) \int |\boldsymbol{\Sigma}_{s_i}|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(\mathbf{y}_i^u - \mathbf{X}_i\boldsymbol{\theta})'\boldsymbol{\Sigma}_{s_i}^{-1}(\mathbf{y}_i^u - \mathbf{X}_i\boldsymbol{\theta})\right) \exp\left(-\frac{1}{2}(\boldsymbol{\theta}_i - \boldsymbol{\theta}_{0,i})'\mathbf{V}_{\theta_i}^{-1}(\boldsymbol{\theta}_i - \boldsymbol{\theta}_{0,i})\right) d\boldsymbol{\theta}_i \\
&\propto p(s_i)|\boldsymbol{\Sigma}_{s_i}|^{-\frac{1}{2}} \exp\left(\frac{1}{2}\widehat{\boldsymbol{\theta}}_i'\mathbf{K}_{\theta_i}\widehat{\boldsymbol{\theta}}_i\right) \int \exp\left(-\frac{1}{2}\left(\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i\right)'\mathbf{K}_{\theta_i}\left(\boldsymbol{\theta}_i - \widehat{\boldsymbol{\theta}}_i\right)\right) d\boldsymbol{\theta}_i \\
&= p(s_i)|\boldsymbol{\Sigma}_{s_i}|^{-\frac{1}{2}}|\mathbf{K}_{\theta_i}|^{-\frac{1}{2}} \exp\left(\frac{1}{2}\widehat{\boldsymbol{\theta}}_i'\mathbf{K}_{\theta_i}\widehat{\boldsymbol{\theta}}_i\right),
\end{aligned}
$$

where $\mathbf{K}_{\theta_i} = \mathbf{X}'\boldsymbol{\Sigma}_{s_i}^{-1}\mathbf{X} + \mathbf{V}_{\theta_i}^{-1}$ and $\widehat{\boldsymbol{\theta}}_i = \mathbf{K}_{\theta_i}^{-1}\left(\mathbf{X}_i'\boldsymbol{\Sigma}_{s_i}^{-1}\mathbf{y}_i + \mathbf{V}_{\theta_i}^{-1}\boldsymbol{\theta}_{0,i}\right)$. As we assume uniform prior for the indicator variable, i.e., $p(s_i = m) = p(s_i = f) = 0.5$, for the SVMVAR-TIC model, then we have

$$
p(s_i|\mathbf{y}_i^u) \propto |\boldsymbol{\Sigma}_{s_i}|^{-\frac{1}{2}}|\mathbf{K}_{\theta_i}|^{-\frac{1}{2}} \exp\left(\frac{1}{2}\widehat{\boldsymbol{\theta}}_i'\mathbf{K}_{\theta_i}\widehat{\boldsymbol{\theta}}_i\right).
$$

Thus it follows that

$$
Pr(s_i = k|\mathbf{y}_i^u) = \frac{p(s_i = k|\mathbf{y}_i^u)}{p(s_i = m|\mathbf{y}_i^u) + p(s_i = f|\mathbf{y}_i^u)}, \text{ for } k \in \{m, f\}.
$$

C. Supplementary Figures

Model without classification

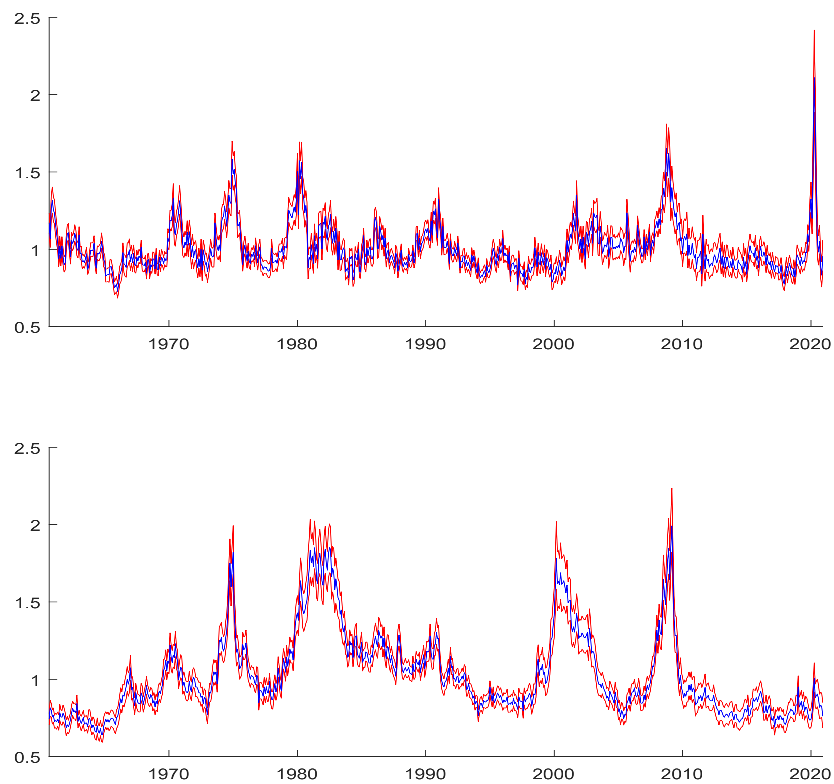

Figure 7: Uncertainty estimates: posterior medians (blue line) and 15%/85% quantiles (red lines), with macro uncertainty ($e^{0.5h_{m,t}}$) in the top panel and financial uncertainty ($e^{0.5h_{f,t}}$) in the bottom panel.

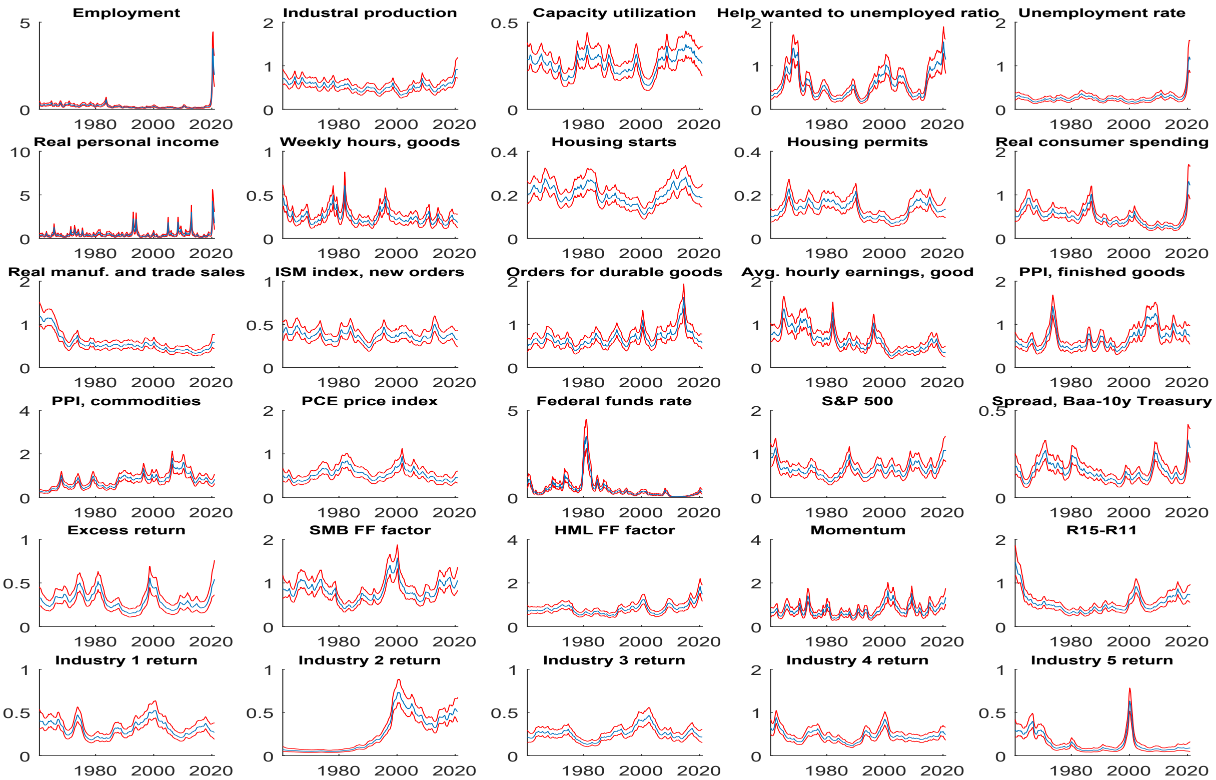

Figure 8: Idiosyncratic volatility estimates ($e^{0.5\eta_{i,t}}$): posterior medians (blue line) and 15%/85% quantiles (red lines).

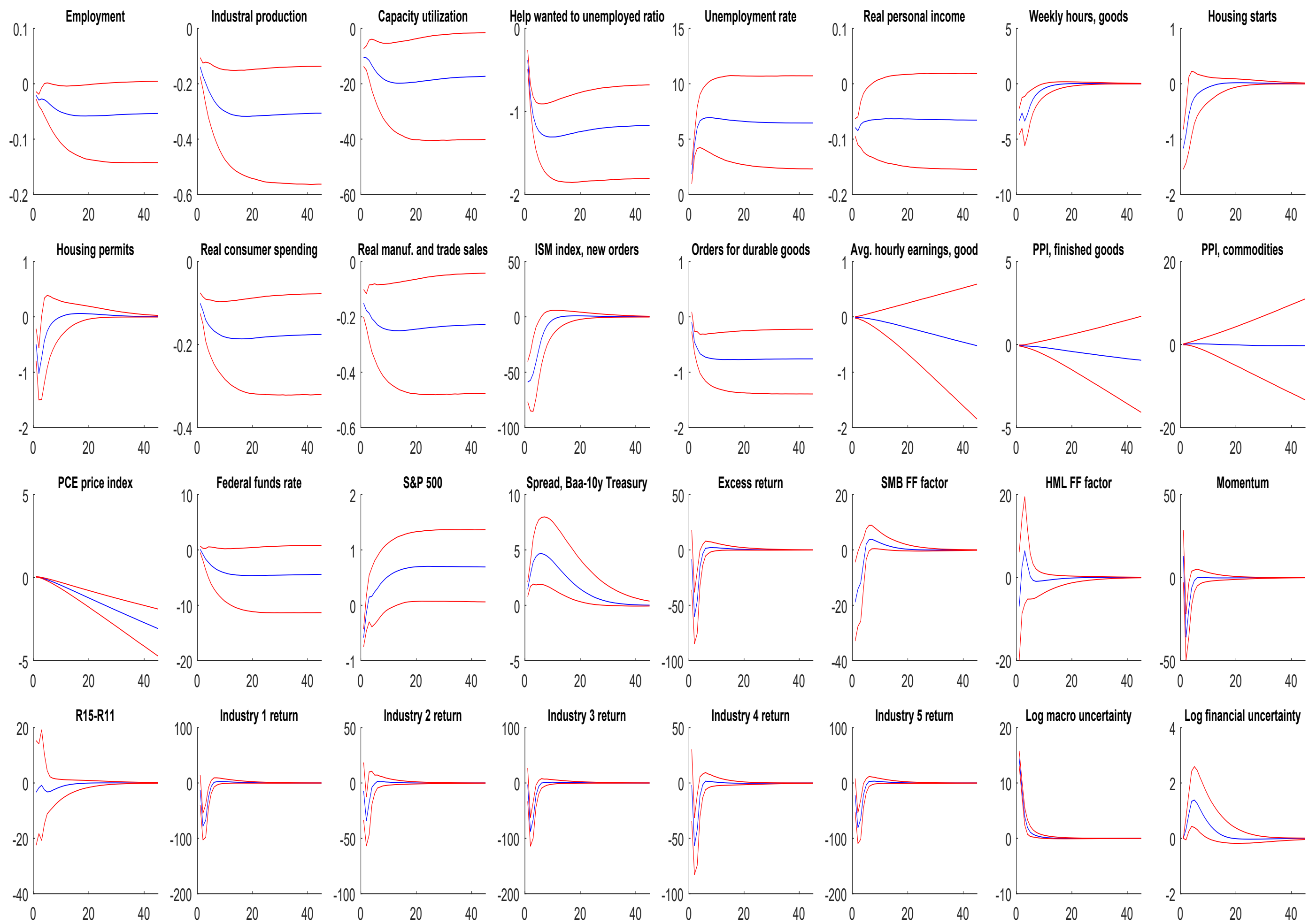

Figure 9: Impulse responses for one standard deviation shock to macro uncertainty: posterior medians (blue lines) and 15%/85% quantiles (red lines).

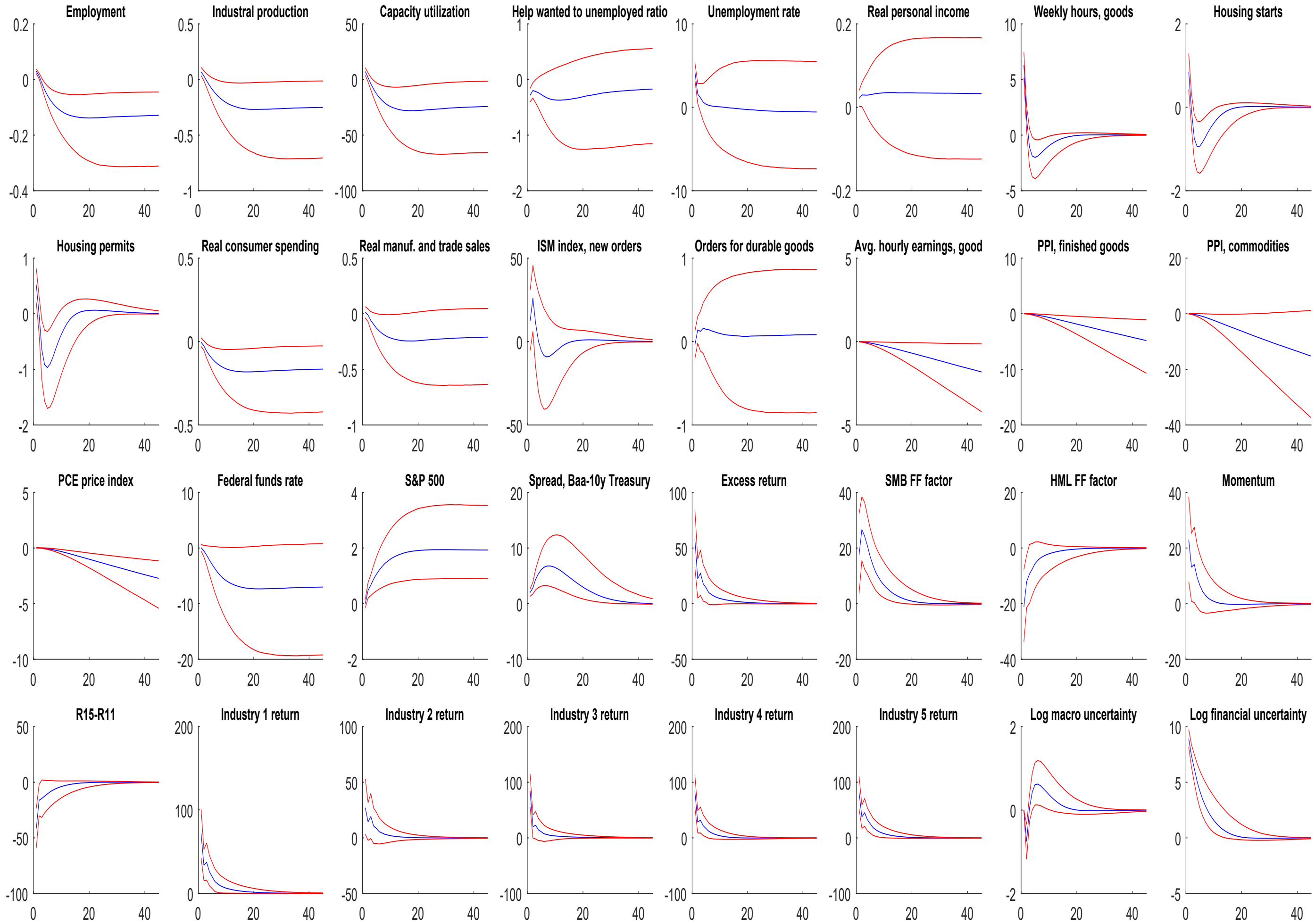

Figure 10: Impulse responses for one standard deviation shock to financial uncertainty: posterior medians (blue lines) and 15%/85% quantiles (red lines).

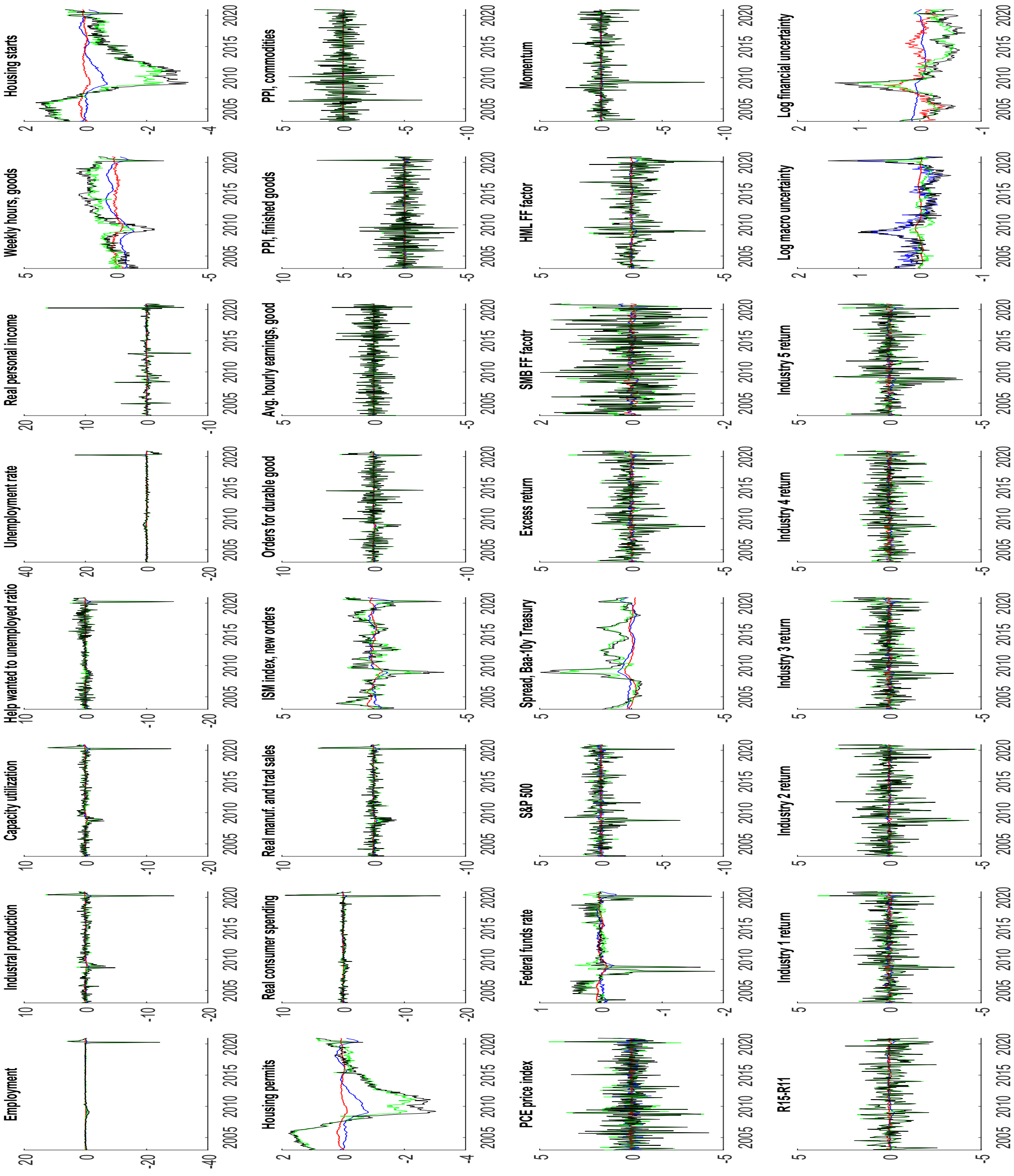

Figure 11: Historical decomposition of SVMVAR for 2003M1 - 2020M12: actual data series (black lines), macro shocks (blue lines), financial shocks (red lines) and VAR shocks (green lines).

Model with time-invariant classification (SVMVAR-TIC)

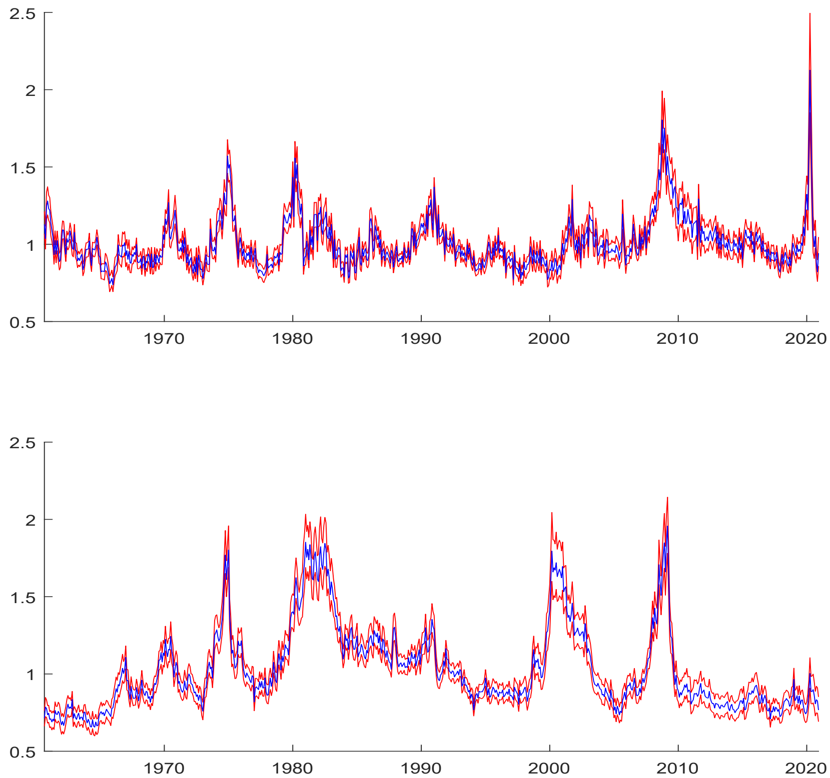

Figure 12: Uncertainty estimates: posterior medians (blue line) and 15%/85% quantiles (red lines), with macro uncertainty ($e^{0.5h_{m,t}}$) in the top panel and financial uncertainty ($e^{0.5h_{f,t}}$) in the bottom panel.

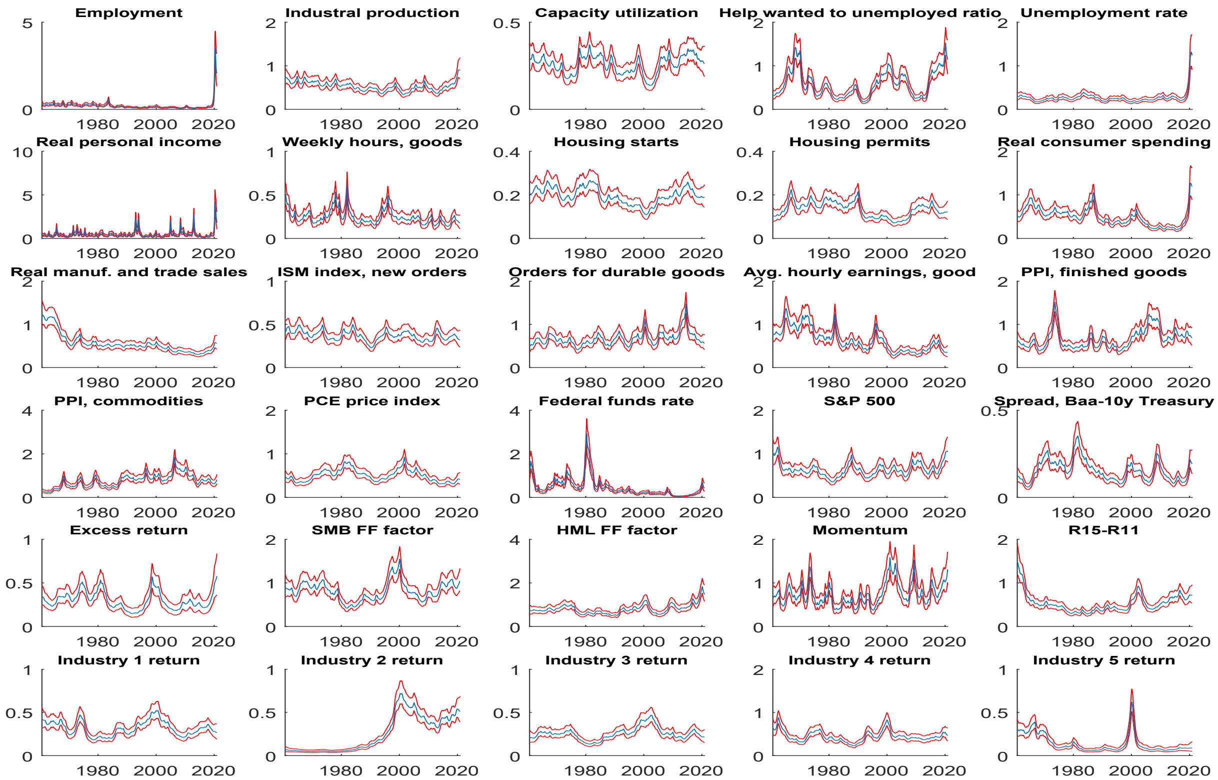

Figure 13: Idiosyncratic volatility estimates ($e^{0.5\eta_{i,t}}$): posterior medians (blue line) and 15%/85% quantiles (red lines).

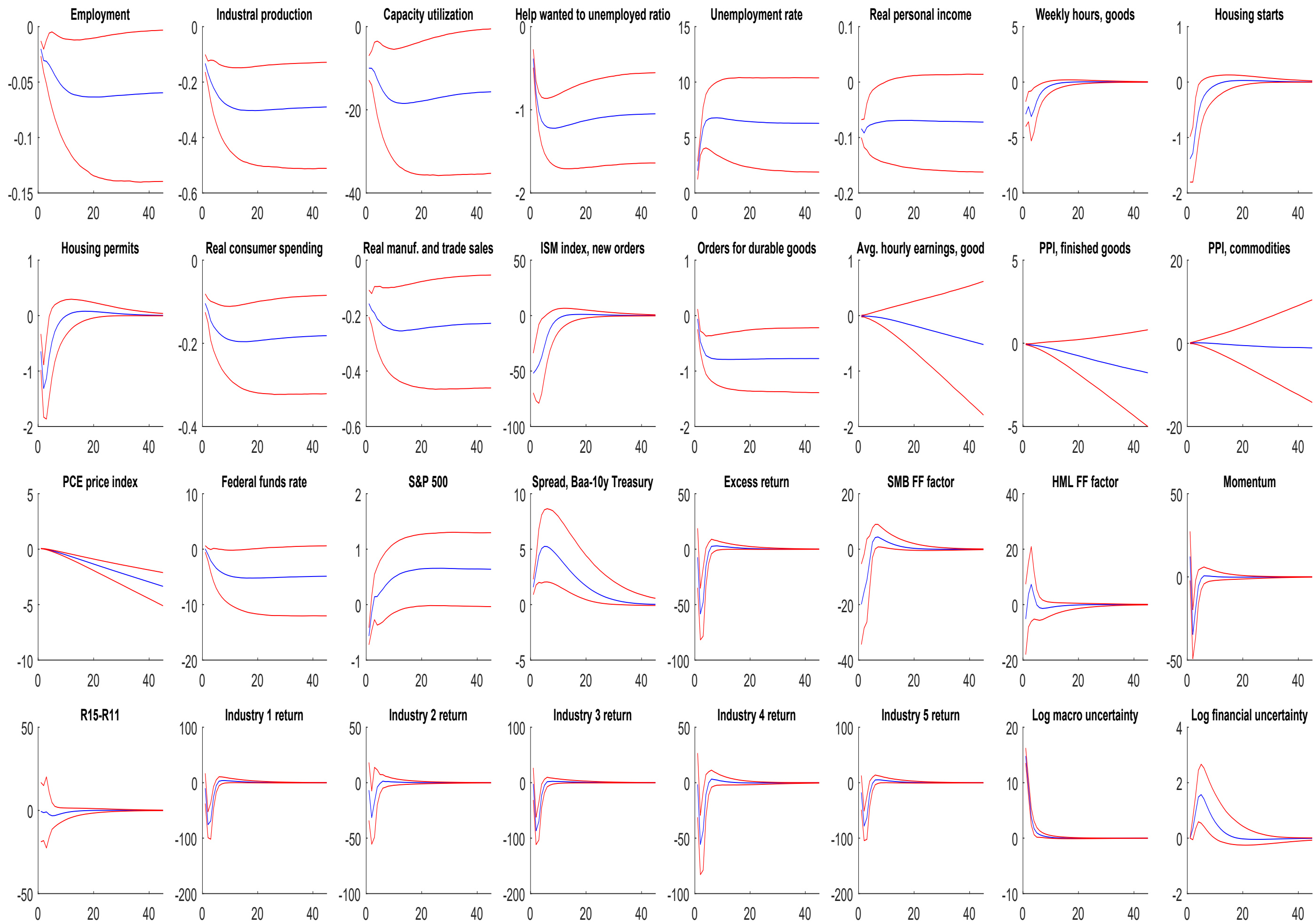

Figure 14: Impulse responses for one standard deviation shock to macro uncertainty: posterior medians (blue lines) and 15%/85% quantiles (red lines).

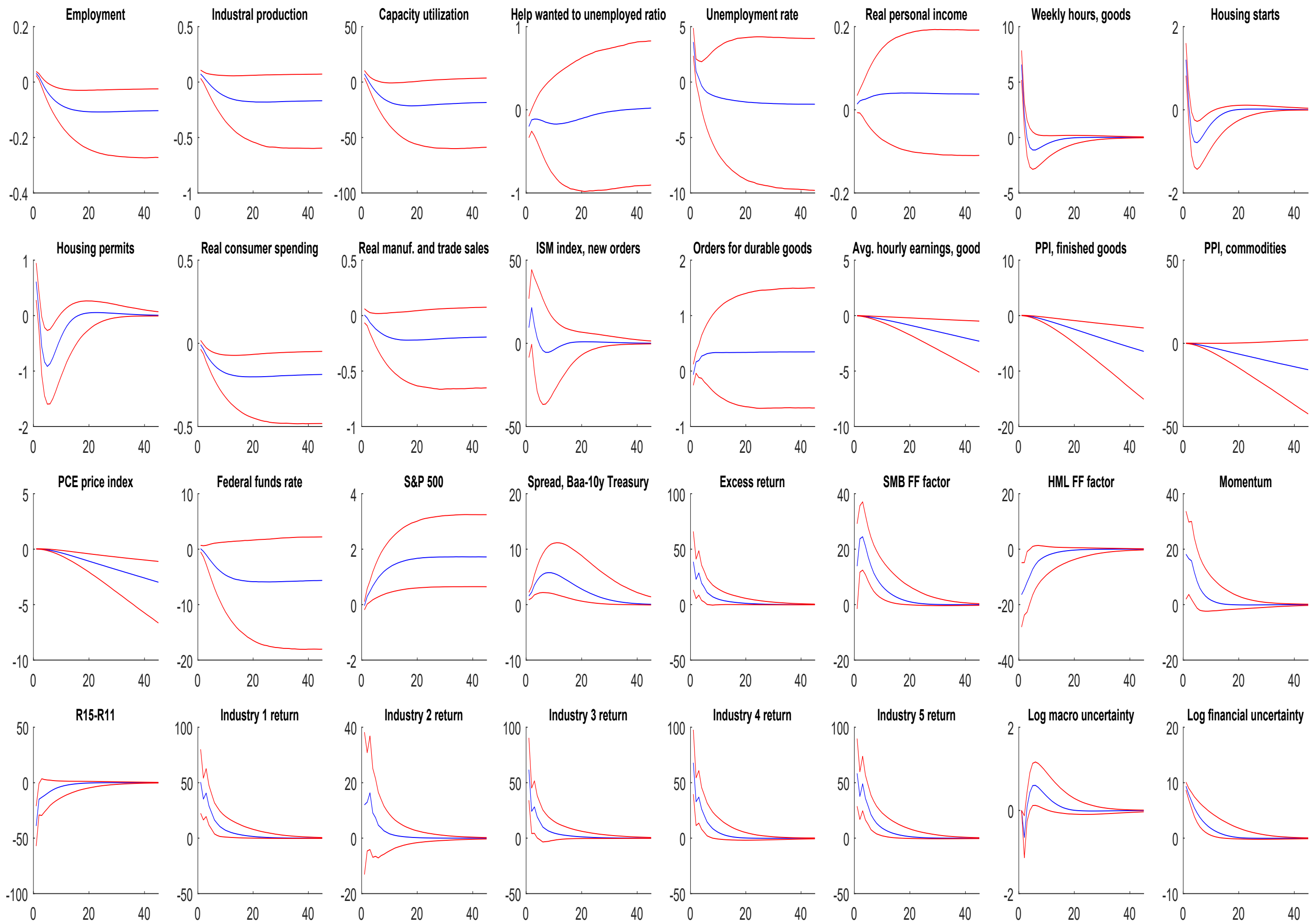

Figure 15: Impulse responses for one standard deviation shock to financial uncertainty: posterior medians (blue lines) and 15%/85% quantiles (red lines).

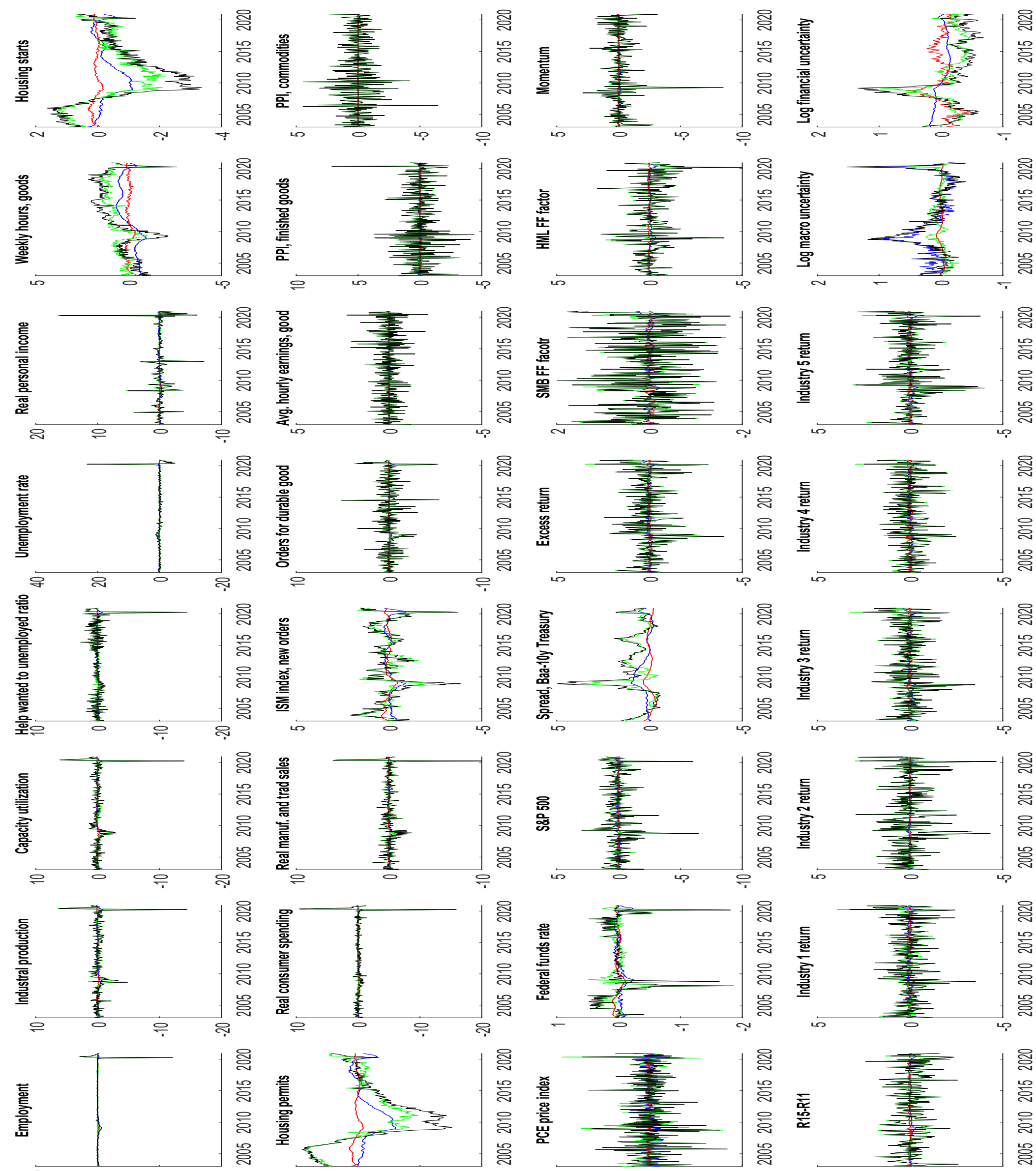

Figure 16: Historical decomposition of SVMVAR-TIC for 2003M1 - 2020M12: actual data series (black lines), macro shocks (blue lines), financial shocks (red lines) and VAR shocks (green lines).

Model with time-varying classification (SVMVAR-TVC)

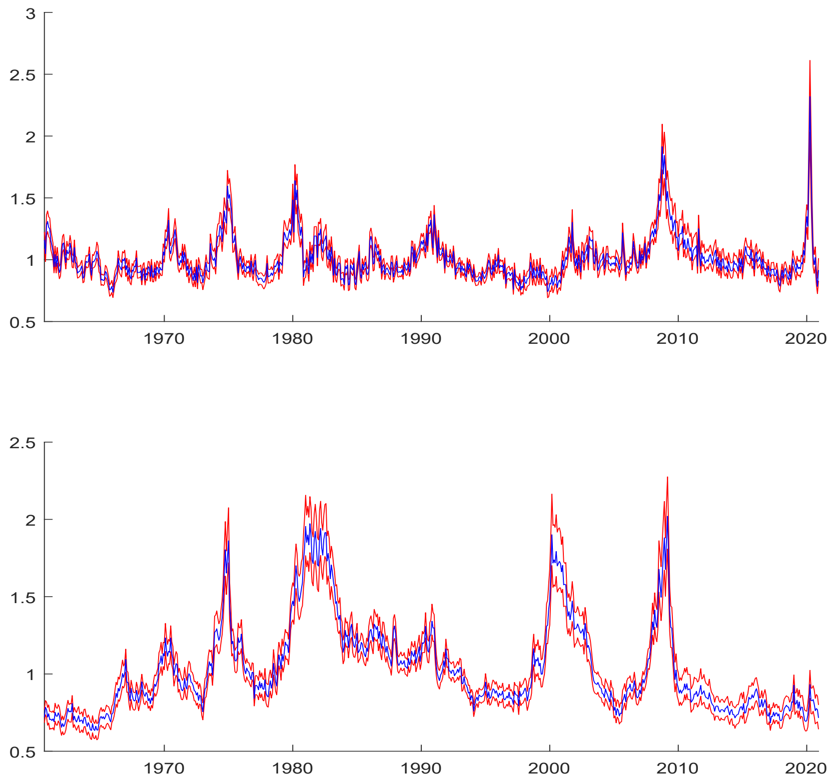

Figure 17: Uncertainty estimates: posterior medians (blue line) and 15%/85% quantiles (red lines), with macro uncertainty ($e^{0.5h_{m,t}}$) in the top panel and financial uncertainty ($e^{0.5h_{f,t}}$) in the bottom panel.

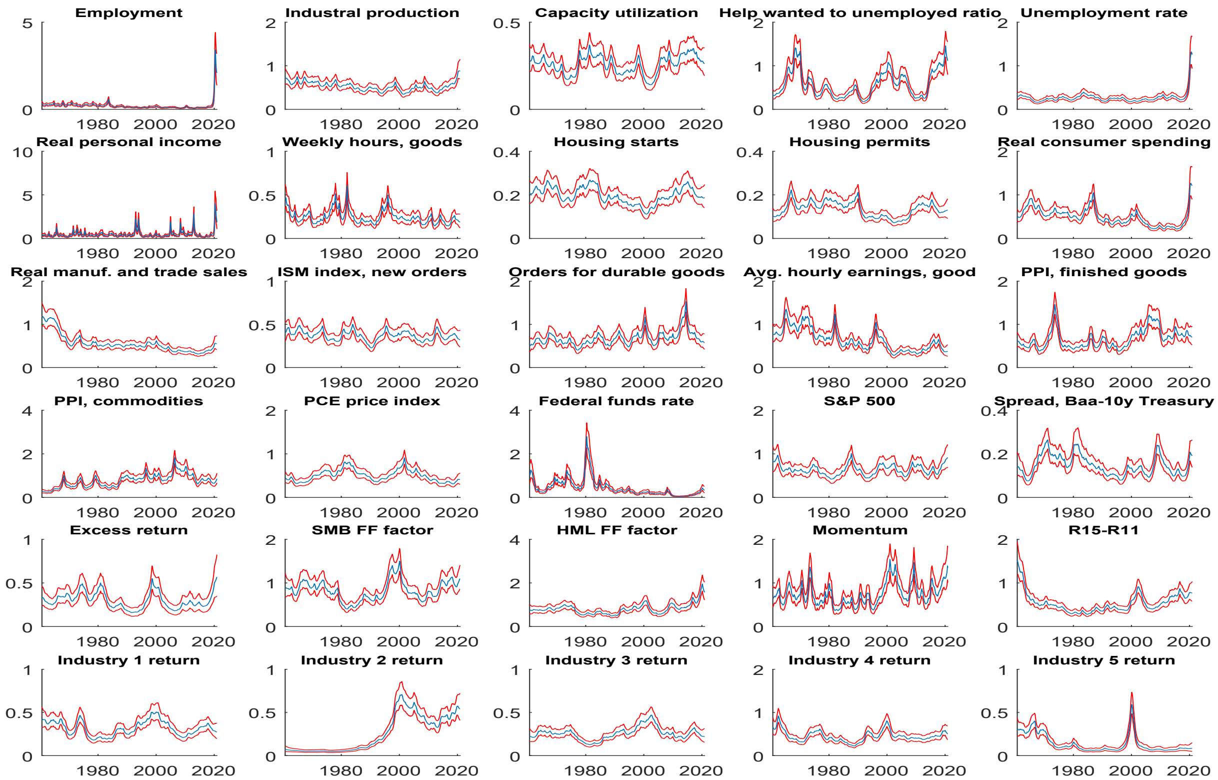

Figure 18: Idiosyncratic volatility estimates ($e^{0.5\eta_{i,t}}$): posterior medians (blue line) and 15%/85% quantiles (red lines).

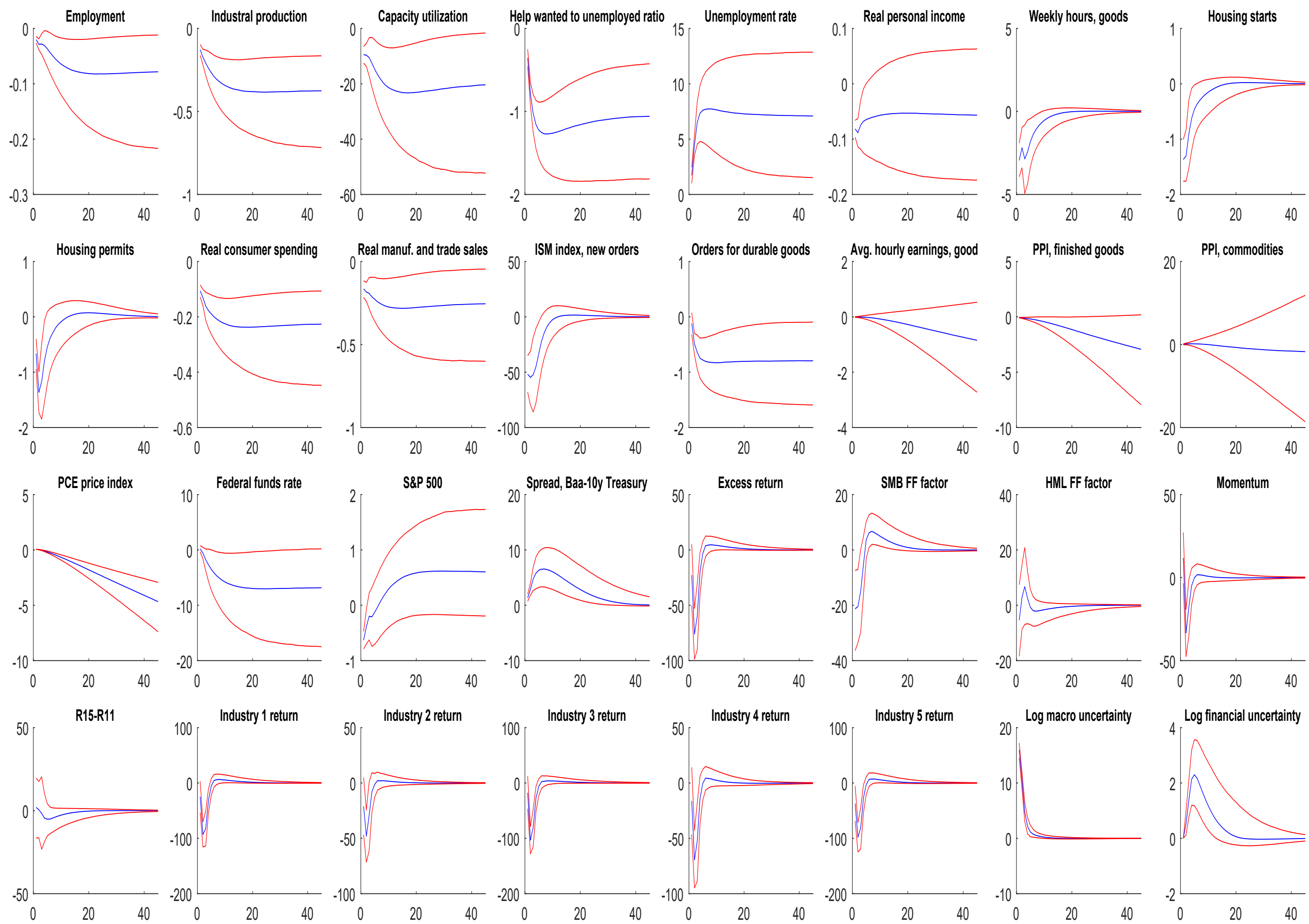

Figure 19: Impulse responses for one standard deviation shock to macro uncertainty: posterior medians (blue lines) and 15%/85% quantiles (red lines).

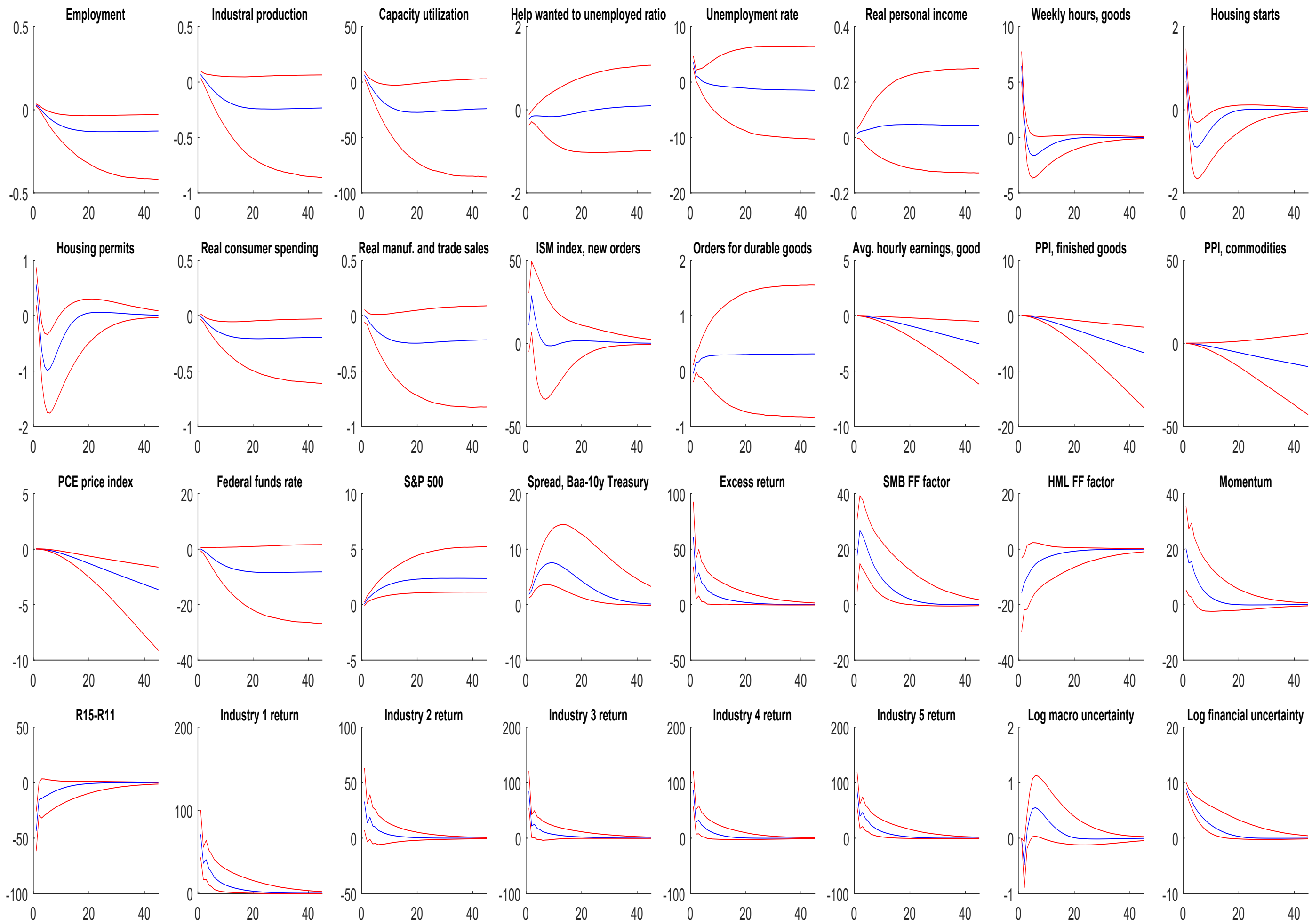

Figure 20: Impulse responses for one standard deviation shock to financial uncertainty: posterior medians (blue lines) and 15%/85% quantiles (red lines).

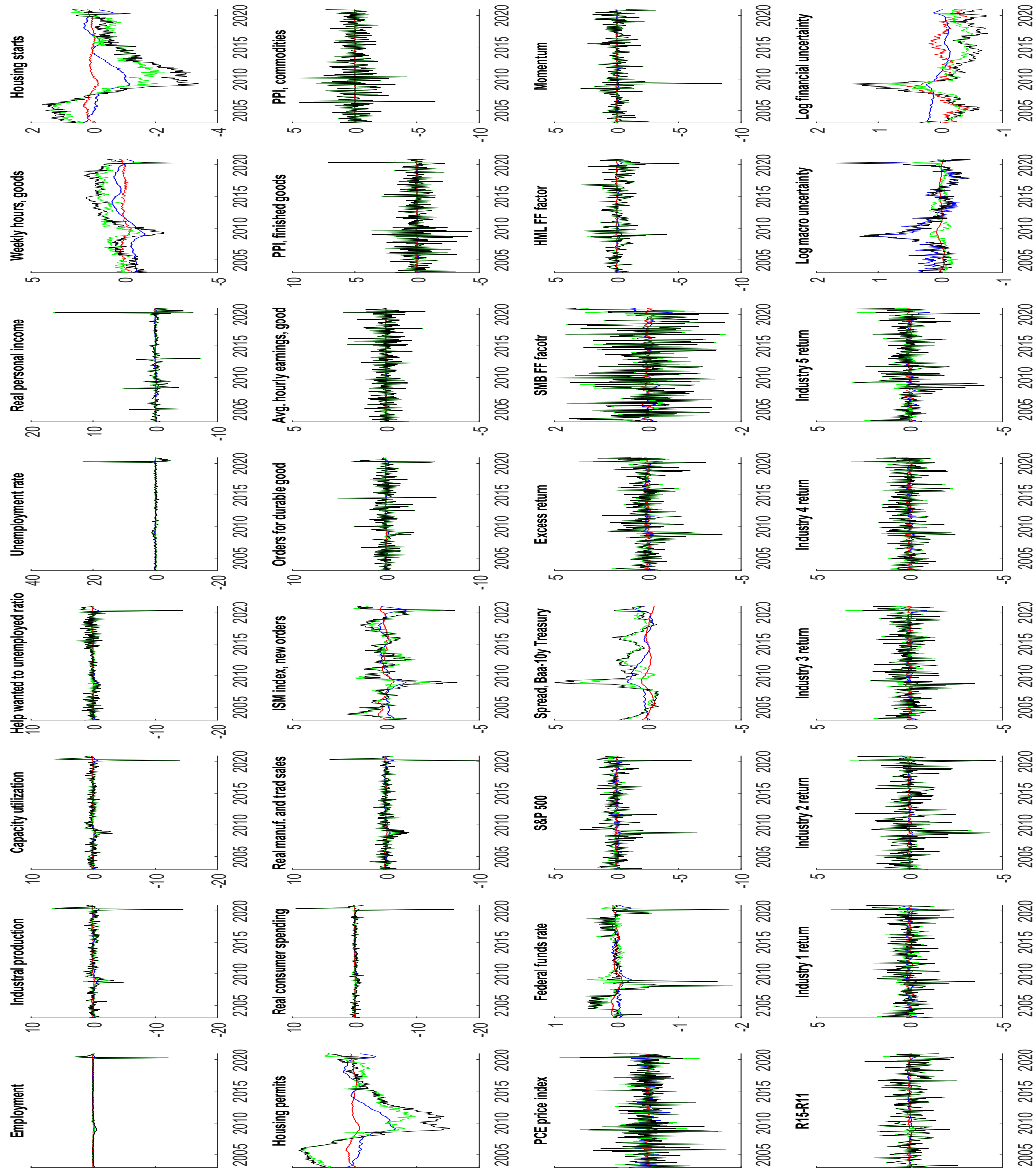

Figure 21: Historical decomposition of SVMVAR-TVC for 2003M1 - 2020M12: actual data series (black lines), macro shocks (blue lines), financial shocks (red lines) and VAR shocks (green lines).