PhD THESIS DECLARATION

I, the undersigned

FAMILY NAME	*Corsello*
NAME	*Francesco*
Student ID no.	*1268754*

Thesis title:

Essays in Large Time-Varying Vector Auto Regressions: Theory and Applications

PhD in	*Economics and Finance*
Cycle	*28^{th}*
Student's Tutor	*Professor Massimiliano Marcellino*
Year of thesis defence	*2018*

DECLARE

under my responsibility:

1. that, according to Italian Republic Presidential Decree no. 445, 28^{th} December 2000, mendacious declarations, falsifying records and the use of false records are punishable under the Italian penal code and related special laws. Should any of the above prove true, all benefits included in this declaration and those of the temporary "embargo" are automatically forfeited from the beginning;
2. that the University has the obligation, according to art. 6, par. 11, Ministerial Decree no. 224, 30^{th} April 1999, to keep a copy of the thesis on deposit at the "Bilioteche Nazionali Centrali" (Italian National Libraries) in Rome and Florence, where consultation will be permitted, unless there is a temporary "embargo" protecting the rights of external bodies and the industrial/commercial exploitation of the thesis;
3. that the Bocconi Library will file the thesis in its "Archivio Istituzionale ad Accesso Aperto" (Institutional Registry) which permits online consultation of the complete text (except in cases of temporary "embargo");
4. that, in order to file the thesis at the Bocconi Library, the University requires that the thesis be submitted online by the student in unalterable format to Società

NORMADEC (acting on behalf of the University), and that NORMADEC will indicate in each footnote the following information:

- PhD thesis: Essays in Large Time-Varying Vector Auto Regressions: Theory and Applications
- by Corsello Francesco
- defended at Università Commerciale "Luigi Bocconi" – Milano in the year 2018;
- the thesis is protected by the regulations governing copyright (Italian law no. 633, 22nd April 1941 and subsequent modifications). The exception is the right of Università Commerciale "Luigi Bocconi" to reproduce the same, quoting the source, for research and teaching purposes;

5. that the copy of the thesis submitted online to Normadec is identical to the copies handed in/sent to the members of the Thesis Board and to any other paper or digital copy deposited at the University offices, and, as a consequence, the University is absolved from any responsibility regarding errors, inaccuracy or omissions in the contents of the thesis;
6. that the contents and organization of the thesis is an original work carried out by the undersigned and does not in any way compromise the rights of third parties (Italian law, no. 633, 22nd April 1941 and subsequent integrations and modifications), including those regarding security of personal details; therefore the University is in any case absolved from any responsibility whatsoever, civil, administrative or penal, and shall be exempt from any requests or claims from third parties;
7. that the thesis is not subject to "embargo", i.e. that it is not the result of work included in the regulations governing industrial property; it was not written as part of a project financed by public or private bodies with restrictions on the diffusion of the results; is not subject to patent or protection registrations.

30 January 2018

Francesco Corsello

ABSTRACT

This PhD dissertation is composed of four research papers, introducing various novel large and time-varying vector autoregressive (VAR) models, their estimation by Bayesian methods, and their application to forecasting and structural analysis.

Chapter 1 ("Rank Reduction in BVARs with Time-Varying Coefficients and Stochastic Volatility: Specification and Estimation") reviews and amends the Gibbs Sampler to estimate a reduced rank time-varying coefficients VAR with stochastic volatility. The adopted rank reduction assumes singular covariance matrices of coefficients and volatilities innovations. Eventually, a number of alternative VAR models that allow for time variation are presented, and their Bayesian estimation strategies are sketched.

Chapter 2 ("Labor Market and Financial Shocks: a Time-Varying Analysis") assesses the interactions of financial shocks and the labor market, by means of a time-varying structural VAR with stochastic volatility, estimated using US data and the Bayesian methodology described in Chapter 1. Results show that the identified financial shock has an asymmetric impact on the unemployment rate in the last three decades. Moreover, there is strong evidence of time variation in the financial shock's volatility, but not in the systematic transmission mechanism, consistently with recent literature.

Chapter 3 ("The Economic Drivers of Volatility and Uncertainty") proposes a Multivariate Autoregressive Index (MAI) model featuring Stochastic Volatility (SV) for a large set of real, nominal and financial indicators for the US economy. Within this framework, structural demand, supply and monetary/financial shocks can be identified from the VAR of the observable factors, using common techniques. The estimated volatility of each variable is then decomposed into common and idiosyncratic components, with the former being substantial, in general more than 50% of the overall volatility, and driven by the structural shocks.

Chapter 4 ("The Global Component of Inflation Volatility") introduces a novel model (MAI-AR-SV) that combines reduced rank restrictions, autoregressive terms and stochastic volatility, for which a new Bayesian MCMC estimation algorithm is set up. The model is used to decompose the national inflation rates of 20 OECD countries into a common component, driven by a single global factor, which typically captures the slow-moving trend, a country-specific component, and an error term featuring, in turn, common and idiosyncratic time-varying volatility. A single common factor explains on average about 70% of the variability of all inflation rates. Moreover, there is substantial and time-varying commonality in the inflation volatilities, and it has increased in the last two decades. While various sources can be behind the global inflation factor, since the early '90s it is strongly correlated with Chinese PPI and Oil inflation. Finally, point and density forecast evaluation shows that the model performs very well against a set of multivariate and univariate competitors.

I più cari ringraziamenti vanno a Massimiliano Marcellino, il Prof., non solo per avermi concesso la preziosa opportunità di lavorare con lui, ma anche per il sostegno e la supervisione, incessanti, sicuramente cruciali per la costruzione di questo elaborato, e fondamentali per valorizzare l'intero percorso di studi.

Ringrazio anche Andrea Carriero, che mi ha sempre spronato all'efficienza e al pragmatismo necessari per dosare al meglio le risorse disponibili.

Un importantissimo grazie anche a Giorgio Primiceri, che ha contribuito a far nascere la curiosità e l'entusiamo per la mia ricerca.

Un grandissimo grazie a Marco Taboga e Marco Bernardini, per i preziosi e utilissimi consigli, e anche per le ricche chiacchierate che ci regaliamo spesso. Grazie anche a tutti gli altri colleghi della divisione Analisi Finanziaria che mi sopportano e supportano quotidianamente.

Dovuti e sentiti ringraziamenti vanno anche alla Direzione e ai colleghi del servizio Gestione dei Rischi, tra cui il Dr. Gelsomino, e gli indimenticabili Antonio R. e Tommaso P., che mi hanno consentito di iniziare e continuare questo percorso di studi e di ricerca.

Un enorme grazie ai ragazzi di via Röntgen confluiti successivamente in via Nazionale, Valerio N.L., Tommaso e Alessandro, con cui abbiamo condiviso, tra le altre cose, croci e delizie della transizione. Valerio va ringraziato anche per tutto il percorso sacro di consigli tecnici scambiati in questi anni, non solo al fine di scrivere insieme un lavoro di ricerca.

Un caloroso grazie a tutti gli altri amici e colleghi che mi hanno accolto a Roma, impreziosendo il mio quotidiano, tra cui Valerio il Prince, Elisa, Maddalena, Ilaria e Riccardo.

Non posso dimenticare l'affetto dei grandi amici e compagni di sempre, Andrea, Giorgio, e Salvo, sempre vicini nonostante la distanza.

Un profondo grazie a mio fratello Antonio, che tiene sempre vivace in me il prezioso seme della contestazione e dell'autocritica.

Un fondamentale grazie a Giulia, il mio conforto e la mia pace.

Infine, un incommensurabile grazie a Mamma e Papà, per l'inestimabile valore dei loro insegnamenti e per l'inesauribile e incondizionato supporto. Senza di loro, nulla di tutto questo sarebbe stato possibile. A loro, e solo a loro, è dedicata questa tesi.

Contents

Disclaimer

Part of the research contained in this dissertation was carried out while the author worked at the Bank of Italy.
The views expressed herein are the authors' only and do not necessarily reflect those of the Bank of Italy.

Introduction

In the last decade, the dynamics of macro-financial variables have been pronounced, somehow ending the era of Great Moderation. In this complex and evolving scenario, the assumption of parameter stability in econometric modeling of long time series seems inappropriate. On the other hand, largely parametrized multivariate models like Vector Auto Regressions suffer the curse of dimensionality even more when parameter stability is broken. However, considering the important accomplishments recently achieved by computer science and multivariate statistical modeling, the macro-financial econometrics literature is more and more producing sophisticated strategies to break the assumption of constant coefficients and volatilities in empirical models, even in case of large amounts of parameters. Bayesian methodologies are particularly of help on this respect, definitely because of the joint work of prior distributions as shrinkage devices and of Monte Carlo Markov Chain techniques as means to address the distributional complexity of posterior objects. However, considering the need of making inference on a number of elements that increases with the availability of larger dataset and longer samples, the substantial computational costs may become huge. Machines alone cannot resolve the complexity arising from all these directions, humans should drive their rising potential, addressing the trade-off between efficiency and sophistication. For these reasons, a fundamental challenge for researchers is constituted by the design of effective computational infrastructures and efficient estimation strategies, in order to cope with the curse of dimensionality of models sophisticated enough to fit the data.

This PhD dissertation is composed of four research papers proposing and applying multivariate models (unrestricted and restricted VARs) that take into account parameter instability, estimated through Bayesian methodologies tailored for highly parametrized empirical applications. The following paragraphs describe the single chapters in greater detail.

Chapter 1 ("Rank Reduction in BVARs with Time-Varying Coefficients and Stochastic

Volatility: Specification and Estimation") reviews and amends the Gibbs Sampler to estimate a reduced rank time-varying coefficients VAR with stochastic volatility, adopting the rank reduction strategy presented by de Wind and Gambetti (2014). This strategy reduces the rank of time variation assuming singular covariance matrices of coefficients and volatilities innovations. The model builds upon the literature started by Cogley and Sargent (2001), who introduced time-varying coefficients in VARs. After Sims (2001) pointed that time variation measured in coefficients may be mistaken for drifting volatility of the residuals, Cogley and Sargent (2005) and Primiceri (2005) show how to estimate VAR models with both coefficients and volatility changing over time. However, the resulting VAR models were highly subject to the curse of dimensionality when specifications included a non-small amount of variables and lags. Moreover, the assumption that each VAR coefficient changes over time following an autonomous component seems conceptually and economically pretentious in some cases, apart from being operationally expensive. For these reasons, rank reduction strategies constitute a reasonable way to attenuate the computational burden, reducing the number of components changing over time. Indeed, in most cases, reducing the dimensionality of time-varying states may leave the model still able to capture some time variation in the data. This chapter introduces also a new algorithm to draw efficiently from the Singular Inverse Wishart, and implements the $\log \chi_1^2$ approximation of Omori et al. (2007). Eventually, a number of alternative VAR models that allow for time variation are presented, and their Bayesian estimation strategies are sketched.

Chapter 2 ("Labor Market and Financial Shocks: a Time-Varying Analysis", co-authored with V. Nispi Landi) assesses the interactions of financial shocks and the labor market, by means of a time-varying VAR with stochastic volatility estimated using US data. Studying the linkages between financial and labor markets is crucial to gain a better comprehension of business cycle fluctuations, especially after the Great Financial Crisis and in light of the debate regarding financial stability policies. Most of the literature typically used a standard homoskedastic VAR model to identify the impact of a financial shock, using sign or short-run restrictions. However, ignoring changing volatility seems quite restrictive when the object of interest is a financial shock. Indeed, since financial time series are typically asymmetric and exhibit changes in volatility over time (the financial shock's size is likely to be high in periods of financial turbulence), a standard VAR may overestimate the effect of a financial innovation across the sample. To this end, a time-varying coefficients VAR with stochastic volatility is estimated over the sample period 1973-2016, using the Bayesian methodology described in Chapter 1. The specification includes eight

US quarterly time series, encompassing labor-market and macroeconomic variables along with the National Financial Condition Index. The financial shock is identified through a short-run restriction, and constitutes an unanticipated exogenous increase in the NFCI. Results show that financial shocks have had an asymmetric impact on the unemployment rate in the last three decades: while negative financial shocks have been responsible for the high unemployment rates in the early 1980s and during the Great Recession, no relevant contributions of financial shocks are found throughout expansion periods. Moreover, the analysis exhibits strong evidence of time variation in the financial shock volatility, but weak evidence of time variation in the systematic transmission mechanism, consistently with some recent literature (see for instance Justiniano and Primiceri, 2008, Stock and Watson, 2012 and Prieto et al., 2016). Therefore, the asymmetric response of the unemployment rate seems to be driven entirely by the financial shock's drifting volatility.

Chapter 3 ("The Economic Drivers of Volatility and Uncertainty", co-authored with A. Carriero and M. Marcellino) proposes a Multivariate Autoregressive Index (MAI) model featuring Stochastic Volatility (SV) for a large set of real, nominal and financial indicators regarding the US economy. The MAI is a particular reduced rank VAR where each variable is driven by (lags of) linear combinations of groups of variables, and the latter can be interpreted as observable factors, or "Indexes", see e.g. Reinsel (1983), Geweke (1996) and Carriero et al. (2016b) for details. In this model the indexes follow themselves a VAR, which is an implication of the MAI model rather than a priori assumption as in factor models. Within this framework, structural demand, supply and monetary/financial shocks can be identified from the VAR of the observable factors, using common techniques. A common feature of papers modeling volatilities as stochastic and time-varying is that different shocks can drive the levels and the volatilities of the variables, but the nature of shocks driving volatilities is left unexplained. By applying a decomposition similar to that introduced by Johansen (1995) in the context of cointegrated models, the estimated volatility of each variable is decomposed into two orthogonal components: one being common to all the variables, the other instead variable specific. Moreover, the common volatility can be further decomposed into three orthogonal components driven by volatilities of the structural demand, supply and monetary / financial shocks, which are the same shocks driving the common factors and hence the conditional means of all variables. Making use of several efficient computational strategies, as for instance the triangularization scheme of Carriero et al. (2016a), a Gibbs Sampling algorithm for (Bayesian) inference is built for the MAI-SV model, allowing for time variation both in variances and covariances across shocks. The model is estimated using 20 macroeconomic and financial variables for the US, over a sample covering from 1964 to 2016. The identified demand, supply and monetary / financial shocks have effects on the conditional expectations of the variables, qualitatively in line with economic theory and with the results in Bańbura et al. (2010) and Carriero et al. (2016b), even though in this context the actual size of the shocks, and therefore their effects, are changing over time due to drifting volatilities. Moreover, the common component of volatility is substantial, in line with the results reported in Carriero et al. (2017), though there are differences across variables. Overall, the structural demand, supply and financial shocks appear as important drivers not only of the level of macroeconomic and financial variables but also of their changing conditional volatility. Common volatility explains, in general, more than 50% of the overall volatility, with their relative importance being variable dependent and, sometimes, changing over time.

Chapter 4 ("The Global Component of Inflation Volatility", co-authored with A. Carriero and M. Marcellino) introduces a novel Multivariate Autoregressive Index (MAI) model, with Stochastic Volatility (SV), and autoregressive (AR) terms, to model jointly the national inflation rates of 20 OECD countries. Cubadda and Guardabascio (2017) included autoregressive terms in a MAI model to improve the forecasting performance, even though they assume constant volatility and operate in a classical estimation framework. This chapter, evolving from the model presented in Chapter 3, combines reduced rank restrictions, autoregressive terms and stochastic volatility into the MAI-AR-SV model, for which a new Bayesian MCMC estimation algorithm is set up. The global dynamics are captured by a single index (a linear combination of all the national inflation rates), representing the global inflation factor. This choice allows to decompose national inflation rates into a common component driven by the single global factor, which typically captures the slow moving trend, a country-specific component and an error term featuring, in turn, common and idiosyncratic time-varying volatility. A decomposition of the stochastic volatilities, similar to the one proposed in chapter 3, allows to measure the time-varying contribution of the single global component relatively to the idiosyncratic components. Papers such as Borio and Filardo (2007) and Ciccarelli and Mojon (2010) illustrate that global developments play an important role in the determination of inflation rates, finding that a substantial amount of variation in a large set of national inflation rates can be explained by global factors. Engle (1982) and Stock and Watson (2007) pointed out the importance of allowing for conditional time-varying volatility when modelling inflation. This literature motivates the choice of an econometric model where the relative contribution of global and country-specific factors as drivers of inflation developments is estimated and can vary over time and across countries. The analysis is carried out using data for 20

OECD countries on headline (all items) inflation rates, over the period 1960-2016, and on core (non-Food and non-Energy items) inflation rates, over the period 1979-2016. Results for headline inflation show that a single common factor explains on average about 70% of the variability of all inflation rates. Moreover, there is also substantial commonality in the inflation volatilities, and it has increased in the last two decades. The average (across countries) share of stochastic volatility explained by the global component spans from 20% to 65% throughout the sample, peaking in the years 2008-2009. While various sources can be behind the global (headline) inflation factor, it turns out that since the early '90s it is strongly correlated with Chinese PPI and Oil inflation. The same analysis on core inflation rates finds a smaller but non-negligible degree of commonality. The global core inflation factor explains roughly 25% of the variability of core CPI inflation levels and the average (across countries) share of stochastic volatility explained by the global component spans from 10% to 20% throughout the sample, without displaying sizable variation over time. The remaining substantial national components of core inflation level and volatility leave scope for national monetary policies. Finally, point and density forecast evaluation shows that the MAI-AR-SV model has very good out of sample properties for inflation rates, when compared with a set of multivariate and univariate competitors, and the SV specification is particularly relevant for the proper calibration of density forecasts.

Bibliography

Bańbura, M., Giannone, D., and Reichlin, L. (2010). Large bayesian vector auto regressions. *Journal of Applied Econometrics*, 25(1):71–92.

Borio, C. E. and Filardo, A. J. (2007). Globalisation and inflation: New cross-country evidence on the global determinants of domestic inflation. *BIS Working Papers 227, Bank for International Settlements.*

Carriero, A., Clark, T. E., and Marcellino, M. (2016a). Large vector autoregressions with stochastic volatility and flexible priors. *Federal Reserve Bank of Cleveland Working Paper, no. 16-17.*

Carriero, A., Clark, T. E., and Marcellino, M. (2017). Measuring uncertainty and its impact on the economy. *The Review of Economics and Statistics.*

Carriero, A., Kapetanios, G., and Marcellino, M. (2016b). Structural analysis with multivariate autoregressive index models. *Journal of Econometrics*, 192(2):332–348.

Ciccarelli, M. and Mojon, B. (2010). Global inflation. *The Review of Economics and Statistics*, 92(3):524–535.

Cogley, T. and Sargent, T. J. (2001). Evolving post-world war ii us inflation dynamics. *NBER macroeconomics annual*, 16:331–373.

Cogley, T. and Sargent, T. J. (2005). Drifts and volatilities: monetary policies and outcomes in the post wwii us. *Review of Economic dynamics*, 8(2):262–302.

Cubadda, G. and Guardabascio, B. (2017). Representation, estimation and forecasting of the multivariate index-augmented autoregressive model. *CEIS Research Paper 397, Tor Vergata University, CEIS.*

de Wind, J. and Gambetti, L. (2014). Reduced-rank time-varying vector autoregressions. *CPB Netherlands Bureau for Economic Policy Analysis, No. 270.*

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007.

Geweke, J. (1996). Bayesian reduced rank regression in econometrics. *Journal of econometrics*, 75(1):121–146.

Johansen, S. (1995). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press on Demand.

Justiniano, A. and Primiceri, G. E. (2008). The Time-Varying Volatility of Macroeconomic Fluctuations. *American Economic Review*, 98(3):604–41.

Omori, Y., Chib, S., Shephard, N., and Nakajima, J. (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics*, 140(2).

Prieto, E., Eickmeier, S., and Marcellino, M. (2016). Time Variation in Macro-Financial Linkages. *Journal of Applied Econometrics*.

Primiceri, G. E. (2005). Time Varying Structural Vector Autoregressions. *The Review of Economic Studies*, 72(3):821–852.

Reinsel, G. (1983). Some results on multivariate autoregressive index models. *Biometrika*, 70(1):145–156.

Sims, C. A. (2001). [evolving post-world war ii us inflation dynamics]: Comment. *NBER macroeconomics annual*, 16:373–379.

Stock, J. H. and Watson, M. W. (2007). Why has us inflation become harder to forecast? *Journal of Money, Credit and banking*, 39(s1):3–33.

Stock, J. H. and Watson, M. W. (2012). Disentangling the channels of the 2007-2009 recession. *Brookings Papers on Economic Activity*, 2012(1):81–135.

Chapter 1

Rank Reduction in BVARs with Time-Varying Coefficients and Stochastic Volatility: Specification and Estimation

1 Introduction

The macro-financial instability and the nonlinearity in policy variables displayed in the last decades have casted serious concerns regarding the assumption of parameter stability when modeling aggregate variables. Empirical models like VARs, which are not based on fully-fledged macrotheoretical skeletons, may be particularly impaired by perturbations in economic patterns. Indeed, for even medium samples of recent data, assuming constant VAR coefficients and time invariant VAR residual volatility shouldn't be advisable. For these reasons, after the huge steps made by computer science and multivariate statistical modeling in the last two decades, the literature started to include sophisticated and refined attempts to break the assumption of constant coefficients and volatilities in multivariate empirical models characterized by non-small parametrization, like VARs. Cogley and Sargent (2001) started to assume and estimate time-varying coefficients in a monetary VAR, and found non-trivial amount of time variation. Sims (2001) questioned that the time variation measured could be attributed entirely to VAR coefficients and not, at least partly, to residuals' volatility. Cogley and Sargent (2005) and Primiceri (2005) enriched the debate by presenting models in which both coefficients and volatility were free to move over time. These papers adopted Bayesian strategies and Markov Chain Monte Carlo methods to tackle the distributional complexity of posterior objects of interest. Notwithstanding, the time-varying VAR models remain characterized by large parametrization and considerable computational costs, since inference has to be made for each set of coefficients of each time period. In medium VAR models with more than 2 lags the number of time-varying forces starts to be considerably high. This is why strategies of rank reduction may be useful to reduce the number of components whose changes over time has to be analyzed. For instance, consider a VAR model with $n = 5$ variables and $p = 4$ lags including a constant term: we reach a total $k = n(np + 1) = 105$ coefficients of the VAR. The assumption that all 105 coefficients change over time following an autonomous component seems conceptually and economically pretentious, apart from being operationally expensive. Indeed, in most cases, reducing substantially the dimensionality of the time-varying states is clearly beneficial in terms of computational costs and may leave the model still able to capture the time variation in the data.

Building upon the Primiceri (2005) framework, de Wind and Gambetti (2014) propose and present a strategy to reduce the number of time-varying forces by assuming rank deficient covariance matrices of coefficients and volatilities innovations. This work reviews entirely and amends the Gibbs Sampler proposed by de Wind and Gambetti (2014). In terms of methodological novelties, a new algorithm to draw from the Singular Inverse Wishart is proposed and the $\log \chi^2_1$ approximation of Omori et al. (2007) is implemented. Moreover, a number of alternative models that allow for time variation are presented and their Bayesian estimation strategy is sketched.

Section 2 presents in detail the benchmark model, the alternatives and cuts through the proposed estimation strategies in detail. Section 3 briefly comments about the algorithm. Section 4 concludes and indicates some avenues for future research in the field.

2 The models

2.1 TV-VAR-SV

We consider the following VAR with p lags and n variables, time-varying coefficients and stochastic volatility, which represents the framework of Primiceri (2005):

$$\underbrace{y_t}_{n\times 1} = c_t + \sum_{\ell=1}^{p} \underbrace{B_{\ell,t}}_{n\times n} y_{t-\ell} + u_t, \qquad u_t \overset{i}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{\Omega_t}_{n\times n}\right).$$

After having defined the vector β_t containing all the coefficients of period t

$$\underbrace{\beta_t}_{k\times 1} \equiv vec\left(\begin{bmatrix} c_t & B_{1,t} & \dots & B_{p,t}\end{bmatrix}'\right), \qquad k \equiv n + pn^2$$

we assume that the law of motion for the $(\beta_t)_{t=1}^{T}$ is given by:

$$\beta_t = \beta_{t-1} + \nu_{\beta,t}, \qquad \nu_{\beta,t} \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{Q_\beta}_{k\times k}\right).$$

Considering now the reduced form covariance matrix Ω_t, and its triangular decomposition:

$$\Omega_t = A_t^{-1}\Sigma_t\Sigma_t\left(A_t^{-1}\right)',$$

$$\underbrace{A_t}_{n\times n} = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ \alpha_{1,t} & 1 & \ddots & \ddots & \vdots \\ \alpha_{2,t} & \alpha_{3,t} & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ \alpha_{r-n+2,t} & \alpha_{r-n+3,t} & \dots & \alpha_{r,t} & 1 \end{bmatrix}, \qquad \underbrace{\alpha_t}_{r\times 1} \equiv \begin{bmatrix} \alpha_{1,t} \\ \alpha_{2,t} \\ \vdots \\ \alpha_{r,t} \end{bmatrix}, \qquad r \equiv \frac{n\,(n-1)}{2},$$

$$\underbrace{\Sigma_t}_{n\times n} = \begin{bmatrix} \sigma_{1,t} & 0 & \dots & 0 \\ 0 & \sigma_{2,t} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sigma_{n,t} \end{bmatrix}, \qquad \underbrace{\sigma_t}_{n\times 1} \equiv \begin{bmatrix} \sigma_{1,t} \\ \sigma_{2,t} \\ \vdots \\ \sigma_{n,t} \end{bmatrix},$$

we assume that the law of motion for the elements in the lower triangular matrix A_t is:

$$\alpha_t = \alpha_{t-1} + \nu_{\alpha,t}, \qquad \nu_{\alpha,t} \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{Q_\alpha}_{r\times r} \right),$$

and that the law of motion for the drifting standard deviations σ_t is:

$$\log \sigma_t = \log \sigma_{t-1} + \nu_{\sigma,t}, \qquad \nu_{\sigma,t} \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{Q_\sigma}_{n\times n} \right).$$

Eventually, constructing the matrix X_t that contains all variables' lags:

$$\underbrace{X_t}_{k\times n} \equiv I_n \ \otimes \begin{bmatrix} 1 \\ y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \end{bmatrix},$$

and decomposing the reduced form error as

$$u_t = A_t^{-1}\Sigma_t\varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ I_n \right),$$

we can stack the VAR in the form

$$y_t = X_t'\beta_t + A_t^{-1}\Sigma_t\varepsilon_t.$$

2.2 Rank Reduction

The strategy of de Wind and Gambetti (2014) lays in the rank deficiency imposed to some or all of the covariance matrices of the innovations $(Q_\beta, Q_\alpha, Q_\sigma)$. The strength of this strategy comes from reducing the dimensionality of time-varying components. Indeed, assume for example that a multivariate stochastic process $\underbrace{z_t}_{m\times 1}$ has the following law of motion:

$$z_t = z_{t-1} + \nu_{z,t}, \qquad \underbrace{\nu_{z,t}}_{m\times 1} \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{Q_z}_{m\times m}\right).$$

If Q_z is a reduced rank matrix with rank $q_z < m$, then Q_z can be decomposed in:

$$Q_z = \underbrace{\Lambda_z}_{m\times q_z} \Lambda_z',$$

so that the law of motion can be transformed, becoming

$$z_t = z_{t-1} + \Lambda_z\ \widetilde{\nu}_{z,t}, \qquad \underbrace{\widetilde{\nu}_{z,t}}_{q_z\times 1} \overset{iid}{\sim} \mathcal{MN}\left(\ \mathbf{0},\ I_{q_z}\ \right),$$

where a smaller number of shocks are actually responsible for the time variation of z_t. Indeed, by projecting z_t onto the column space of Λ_z, we can decompose z_t as:

$$z_t = \underbrace{\Lambda_z\left(\Lambda_z'\Lambda_z\right)^{-1}\Lambda_z'}_{P_z}\ z_t + \underbrace{\left(I_m - \Lambda_z\left(\Lambda_z'\Lambda_z\right)^{-1}\Lambda_z'\right)}_{M_z}\ z_t.$$

The latter addendum of the right hand-side depends just on the starting condition z_0. Indeed, premultiplying the law of motion by M_z, we obtain:

$$\begin{aligned}
M_z z_t &= M_z z_{t-1} + M_z \; \Lambda_z \; \widetilde{\nu}_{z,t} \\
&= M_z z_{t-1} + \left(I_m - \Lambda_z \left(\Lambda_z' \Lambda_z \right)^{-1} \Lambda_z' \right) \; \Lambda_z \; \widetilde{\nu}_{z,t} \\
&= M_z z_{t-1} + \left(\Lambda_z - \Lambda_z \right) \widetilde{\nu}_{z,t} \\
&= M_z z_{t-1} + \mathbf{0} \\
&= M_z z_0.
\end{aligned}$$

Moreover, defining

$$\begin{aligned}
R_z &\equiv \left(\Lambda_z' \Lambda_z \right)^{-1} \Lambda_z', \\
\widetilde{z}_t &\equiv R_z z_t,
\end{aligned}$$

and considering that

$$\begin{aligned}
P_z &= \Lambda_z \left(\Lambda_z' \Lambda_z \right)^{-1} \Lambda_z' = \Lambda_z R_z, \\
R_z \Lambda_z &= \left(\Lambda_z' \Lambda_z \right)^{-1} \Lambda_z' \Lambda_z = I_{q_z}, \\
P_z z_t &= \Lambda_z R_z z_t \\
&= \Lambda_z \widetilde{z}_t,
\end{aligned}$$

then, premultiplying by R_z both sides of the law of motion of z_t, we obtain the law of motion of $\widetilde{z}_t$:

$$\widetilde{z}_t = \widetilde{z}_{t-1} + \widetilde{\nu}_{z,t}, \qquad \underbrace{\widetilde{\nu}_{z,t}}_{q_z \times 1} \overset{iid}{\sim} \mathcal{MN} \left(\; \mathbf{0}, \; I_{q_z} \; \right),$$

and the decomposition of z_t can be rewritten as:

$$\begin{aligned}
\underbrace{z_t}_{m \times 1} &= P_z z_t + M_z z_t \\
&= \Lambda_z \underbrace{\widetilde{z}_t}_{q_z \times 1} + M_z z_0.
\end{aligned}$$

Therefore, we decomposed the m-dimensional z_t in a set of linear combinations of a smaller dimensional vector of time-varying component $\widetilde{z}_t$ and in a "residual" non time-varying part that just depends on initial condition. There is not a unique decomposition $Q_z = \Lambda_z \Lambda_z'$ with reduced rank q_z. In this framework we set:

$$\underbrace{\Lambda_z}_{m \times q_z} = \underbrace{V_z}_{m \times q_z} \underbrace{D_z}_{q_z \times q_z},$$

where D_z is the diagonal matrix containing the square roots of the q_z non-zero eigenvalues of Q_z, while V_z is the matrix whose columns are the associated eigenvectors (normalized to unit length).

de Wind and Gambetti (2014) propose the rank reduction in all the main time-varying elements of the model, which constitutes the most general case. They assume for each set of vectors among $(\beta_t,\ \alpha_t,\ \log \sigma_t)_{t=1}^T$ a law of motion as the one of z_t above, and set a different reduced rank $(q_\beta, q_\alpha, q_\sigma)$ for each covariance matrix of the innovations $(Q_\beta, Q_\alpha, Q_\sigma)$. The law of motions of this framework are the following:

$$\underbrace{\beta_t}_{k \times 1} = \Lambda_\beta \widetilde{\beta}_t + M_\beta \beta_0, \qquad \underbrace{\widetilde{\beta}_t}_{q_\beta \times 1} = \widetilde{\beta}_{t-1} + \widetilde{\nu}_{\beta,t}, \quad \underbrace{\widetilde{\nu}_{\beta,t}}_{q_\beta \times 1} \overset{iid}{\sim} \mathcal{MN}\left(\ \mathbf{0},\ I_{q_\beta}\ \right), \quad Q_\beta = \underbrace{\Lambda_\beta}_{k \times q_\beta} \Lambda_\beta',$$

$$\underbrace{\alpha_t}_{r \times 1} = \Lambda_\alpha \widetilde{\alpha}_t + M_\alpha \alpha_0, \qquad \underbrace{\widetilde{\alpha}_t}_{q_\alpha \times 1} = \widetilde{\alpha}_{t-1} + \widetilde{\nu}_{\alpha,t}, \quad \underbrace{\widetilde{\nu}_{\alpha,t}}_{q_\alpha \times 1} \overset{iid}{\sim} \mathcal{MN}\left(\ \mathbf{0},\ I_{q_\alpha}\ \right), \quad Q_\alpha = \underbrace{\Lambda_\alpha}_{r \times q_\alpha} \Lambda_\alpha',$$

$$\underbrace{\log \sigma_t}_{n \times 1} = \Lambda_\sigma \widetilde{\sigma}_t + M_\sigma \log \sigma_0, \quad \underbrace{\widetilde{\sigma}_t}_{q_\sigma \times 1} = \widetilde{\sigma}_{t-1} + \widetilde{\nu}_{\sigma,t}, \quad \underbrace{\widetilde{\nu}_{\sigma,t}}_{q_\sigma \times 1} \overset{iid}{\sim} \mathcal{MN}\left(\ \mathbf{0},\ I_{q_\sigma}\ \right), \quad Q_\sigma = \underbrace{\Lambda_\sigma}_{n \times q_\sigma} \Lambda_\sigma'.$$

2.3 The Prior

The prior distribution for the histories $(\beta_t,\ \alpha_t,\ \sigma_t)_{t=1}^T$ is defined via the chain rule, by exploiting the laws of motion of the time-varying coefficients and the associated distributions of the $\nu_{z,t}$. In addition, independent priors for the covariance matrices of the innovations

$(Q_\beta, Q_\alpha, Q_\sigma)$ are included to obtain the joint prior of all the model's elements:

$$f\left(\{(\beta_t,\ \alpha_t,\ \sigma_t)\}_{t=1}^T,\ Q_\beta, Q_\alpha, Q_\sigma\right) = f(\beta_0) f(\alpha_0) f(\sigma_0) f(Q_\beta) f(Q_\alpha) f(Q_\sigma) \cdot \prod_{t=1}^T \left[f_{\mathcal{N}}(\beta_t|\ \beta_{t-1},\ Q_\beta) f_{\mathcal{N}}(\alpha_t|\ \alpha_{t-1},\ Q_\alpha) \cdot f_{log\mathcal{N}}(\sigma_t|\ \sigma_{t-1},\ Q_\sigma)\right].$$

Then, it is crucial to define priors for the initial conditions of histories $(\beta_0,\ \alpha_0,\ \sigma_0)$ and for the covariance matrices $(Q_\beta, Q_\alpha, Q_\sigma)$. The prior distributions for such elements are assumed to be the following:

$$\beta_0 \sim \mathcal{MN}\left(\ \bar{\beta},\ \bar{P}_\beta\ \right), \qquad \alpha_0 \sim \mathcal{MN}\left(\ \bar{\alpha},\ \bar{P}_\alpha\ \right), \qquad \ln \sigma_0 \sim \mathcal{MN}\left(\ \bar{\sigma},\ \bar{P}_\sigma\ \right),$$
$$Q_\beta \sim \mathcal{SIW}_k^{q_\beta}\left(\bar{Q}_\beta,\ \tau_{\beta,0}\right), \quad Q_\alpha \sim \mathcal{SIW}_r^{q_\alpha}\left(\bar{Q}_\alpha,\ \tau_{\alpha,0}\right), \quad Q_\sigma \sim \mathcal{SIW}_n^{q_\sigma}\left(\bar{Q}_\sigma,\ \tau_{\sigma,0}\right).$$

2.4 The Gibbs Sampler

This section provides a step by step description of the Gibbs Sampler for the most general case in which rank reduction is implemented for all the time-varying states of the model. The algorithm has to be initialized at some starting point $(\beta_t^0,\ \alpha_t^0,\ \sigma_t^0)_{t=0}^T,\ Q_\beta^0, Q_\alpha^0, Q_\sigma^0$.

2.4.1 Step 1: draw a history of TV-VAR coefficients $(\beta_t^i)_{t=1}^T$

Given the previous step of the Markov chain $\left(\beta_t^{i-1},\ \alpha_t^{i-1},\ \sigma_t^{i-1}\right)_{t=1}^T,\ Q_\beta^{i-1}, Q_\alpha^{i-1}, Q_\sigma^{i-1}$, we want to draw from:

$$f\left((\beta_t)_{t=1}^T \middle| \left\{\left(\alpha_t^{i-1},\ \sigma_t^{i-1}\right)\right\}_{t=1}^T,\ Q_\beta^{i-1}, Q_\alpha^{i-1}, Q_\sigma^{i-1}\right).$$

The distribution of β_t depends on α_t and σ_t only via the reduced form covariance matrix of the residuals Ω_t, and does not depend on the innovations covariances (Q_α, Q_σ) but just on the covariance matrix Q_β. In order to exploit the benefits from the rank reduction, we draw separately the transformed time-varying components $\Lambda_\beta \widetilde{\beta}_t$ and the non-TV component $M_\beta \beta_0$, since

$$\beta_t = \Lambda_\beta \widetilde{\beta}_t + M_\beta \beta_0.$$

2.4.1.1 SubStep: draw a history of TV components $\left(\widetilde{\beta}_t^i\right)_{t=1}^T$

We are now conditioning on the elements from the previous step of the chain:

$$\Big(M_\beta \beta_0,\ \Lambda_\beta, (u_t, \Omega_t, y_t, X_t)_{t=1}^T\Big).$$

Then, from the stacked version of the VAR and the decomposition of β_t we can write

$$\begin{aligned} y_t &= X_t' \beta_t + u_t \\ &= X_t' \Lambda_\beta \widetilde{\beta}_t + X_t' M_\beta \beta_0 + A_t^{-1} \Sigma_t \varepsilon_t, \\ \underbrace{y_t - X_t' M_\beta \beta_0}_{y_t^*} &= \underbrace{X_t' \Lambda_\beta}_{X_t^*} \widetilde{\beta}_t + u_t. \end{aligned}$$

Adding the law of motion of $\widetilde{\beta}_t$, we obtain a Gaussian state space representation:

$$\begin{aligned} y_t^* &= X_t^* \widetilde{\beta}_t + u_t, & u_t &\overset{iid}{\sim} \mathcal{MN}\left(\ \mathbf{0},\ \Omega_t\ \right), \\ \underbrace{\widetilde{\beta}_t}_{q_\beta \times 1} &= \widetilde{\beta}_{t-1} + \widetilde{\nu}_{\beta,t}, & \underbrace{\widetilde{\nu}_{\beta,t}}_{q_\beta \times 1} &\overset{iid}{\sim} \mathcal{MN}\left(\ \mathbf{0},\ I_{q_\beta}\ \right). \end{aligned}$$

Using this representation, we can apply the FFBS (Forward Filtering Backward Sampling) procedure described in Carter and Kohn (1994). The FF part is based on the following Kalman filter recursion. After having set the initial condition of the states borrowing from the prior of β_0, whose first moment and covariance matrix are respectively $\bar{\beta}$ and $\bar{P}_\beta$, we can set

$$\widetilde{\beta}_{0|0} = R_\beta \bar{\beta}, \qquad P_{\beta,0|0} = R_\beta \bar{P}_\beta R_\beta'.$$

Then, given $\left(\widetilde{\beta}_{t-1|t-1},\ P_{\beta,t-1|t-1}\right)$, we compute:

$$\begin{aligned} P_{\beta,t|t-1} &= P_{\beta,t-1|t-1} + I_{q_\beta}, \\ K_{\beta,t} &= P_{\beta,t|t-1} X_t^{*\prime} \left(X_t^* P_{\beta,t|t-1} X_t^{*\prime} + \Omega_t\right)^{-1}, \\ \widetilde{\beta}_{t|t} &= \widetilde{\beta}_{t-1|t-1} + K_{\beta,t} \left(y_t^* - X_t^{*\prime} \widetilde{\beta}_{t-1|t-1}\right), \\ P_{\beta,t|t} &= P_{\beta,t|t-1} - K_{\beta,t} X_t^{*\prime} P_{\beta,t|t-1}. \end{aligned}$$

With an entire set of updating and prediction steps $\left(\widetilde{\beta}_{t|t}, P_{\beta,\,t|t}, P_{\beta,\,t|t-1}\right)_{t=1}^{T}$, we start to sample backward, beginning by sampling $\widetilde{\beta}_T$ from $\mathcal{MN}\left(\ \widetilde{\beta}_{T|T},\ P_{\beta,\,T|T}\ \right)$, and then for each $t\ \in \{T-1, T-2, \ldots, 2, 1\}$ we sample recursively each $\widetilde{\beta}_t$ from $\mathcal{MN}\left(\widetilde{\beta}_{t|t+1}, P_{\beta,\,t|t+1}\right)$ where:

$$
\begin{aligned}
\widetilde{\beta}_{t|t+1} &= \widetilde{\beta}_{t|t} + P_{\beta,\,t|t}P^{-1}_{\beta,\,t+1|t}\left(\widetilde{\beta}_{t+1} - \widetilde{\beta}_{t|t}\right), \\
P_{\beta,\,t|t+1} &= P_{\beta,\,t|t} - P_{\beta,\,t|t}P^{-1}_{\beta,\,t+1|t}P_{\beta,\,t|t}.
\end{aligned}
$$

2.4.1.2 SubStep: draw a vector of time invariant component $M_\beta\beta_0$ of β_t

Conditioning at the previous step of the chain and at the new time-varying components:

$$
\left(\Lambda_\beta, \left(\widetilde{\beta}_t,\ u_t, \Omega_t, y_t, X_t\right)_{t=1}^{T}\right),
$$

we can draw $M_\beta\beta_0$ using the following representation:

$$
\begin{aligned}
y_t &= X_t'\beta_t + u_t \\
&= X_t'\Lambda_\beta\widetilde{\beta}_t + X_t'M_\beta\beta_0 + A_t^{-1}\Sigma_t\varepsilon_t, \\
\underbrace{y_t - X_t'\Lambda_\beta\widetilde{\beta}_t}_{y_t^{\bullet}} &= X_t'\ \underbrace{M_\beta\beta_0}_{\mu_{\beta,0}} + u_t, \\
y_t^{\bullet} &= X_t'\ \mu_{\beta,0} + u_t, \qquad u_t \overset{i}{\sim} \mathcal{MN}\left(\ \mathbf{0},\ \Omega_t\ \right).
\end{aligned}
$$

This can be seen as a linear regression model, but we have still to impose some restriction on $\underbrace{\mu_{\beta,0}}_{k\times 1}$. Indeed, since the following holds

$$
\begin{aligned}
\underbrace{R_\beta}_{q_\beta\times k}\,\underbrace{\mu_{\beta,0}}_{k\times 1} &= \underbrace{R_\beta}_{q_\beta\times k}\,\underbrace{M_\beta}_{k\times k}\,\underbrace{\beta_0}_{k\times 1} \\
&= \left(\Lambda_\beta'\Lambda_\beta\right)^{-1}\Lambda_\beta'\left(I_k - \Lambda_\beta\left(\Lambda_\beta'\Lambda_\beta\right)^{-1}\Lambda_\beta'\right)\beta_0 \\
&= \left[\left(\Lambda_\beta'\Lambda_\beta\right)^{-1}\Lambda_\beta' - \left(\Lambda_\beta'\Lambda_\beta\right)^{-1}\Lambda_\beta'\Lambda_\beta\left(\Lambda_\beta'\Lambda_\beta\right)^{-1}\Lambda_\beta'\right]\beta_0 \\
&= \left[\left(\Lambda_\beta'\Lambda_\beta\right)^{-1}\Lambda_\beta' - \left(\Lambda_\beta'\Lambda_\beta\right)^{-1}\Lambda_\beta'\right]\beta_0 \\
&= \mathbf{0}_{q_\beta\times k}\;\beta_0 \\
&= \mathbf{0}_{q_\beta\times 1},
\end{aligned}
$$

we can impose these $q_\beta < k$ restrictions to the k coefficients $\mu_{\beta,0}$ through the matrix R_β. Our linear regression can now be written in the stacked form:

$$
\begin{bmatrix} y_1^\bullet \\ y_2^\bullet \\ \vdots \\ y_T^\bullet \end{bmatrix} = \begin{bmatrix} X_1' \\ X_2' \\ \vdots \\ X_T' \end{bmatrix}\mu_{\beta,0} + \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_T \end{bmatrix},
$$

$$
\underbrace{Y^\bullet}_{nT\times 1} = \underbrace{X^\bullet}_{nT\times k}\,\mu_{\beta,0} + \underbrace{U^\bullet}_{nT\times 1},
$$

where $U^\bullet$ is multivariate normal with a block diagonal covariance matrix where the blocks are the reduced form covariance matrices $\{\Omega_t\}_{t=1}^T$:

$$
\underbrace{U^\bullet}_{nT\times 1} \sim \mathcal{MN}\left(\,\mathbf{0}_{nT\times 1},\;\Omega\,\right), \qquad \underbrace{\Omega}_{nT\times nT} = \begin{bmatrix} \underbrace{\Omega_1}_{n\times n} & \mathbf{0}_{n\times n} & \dots & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \underbrace{\Omega_2}_{n\times n} & \ddots & \mathbf{0}_{n\times n} \\ \vdots & \ddots & \ddots & \mathbf{0}_{n\times n} \\ \mathbf{0}_{n\times n} & \dots & \mathbf{0}_{n\times n} & \underbrace{\Omega_T}_{n\times n} \end{bmatrix}.
$$

To sum up, our restricted regression model can be written as:

$$\underbrace{Y^\bullet}_{nT\times 1} = \underbrace{X^\bullet}_{nT\times k} \mu_{\beta,0} + \underbrace{U^\bullet}_{nT\times 1}, \qquad U^\bullet \sim \mathcal{MN}\left(\mathbf{0}_{nT\times 1},\ \Omega \right)$$

$$\underbrace{R_\beta}_{q_\beta\times k} \underbrace{\mu_{\beta,0}}_{k\times 1} = \mathbf{0}_{q_\beta\times 1}.$$

The Bayesian approach of this restricted regression model is well documented in the Appendix of de Wind and Gambetti (2014). We just report that, given the normal prior distribution of β_0:

$$\beta_0 \sim \mathcal{MN}\left(\bar{\beta},\ \bar{P}_\beta \right),$$

then the posterior distribution of $\mu_{\beta,0}$ is conjugate.
In particular, given the unrestricted standard Bayesian regression moments:

$$\Psi_\beta = \left(X^{\bullet\prime}\Omega^{-1}X^\bullet + \bar{P}_\beta^{-1}\right)^{-1},$$

$$\psi_\beta = \Psi_\beta \left(X^{\bullet\prime}\Omega^{-1}Y^\bullet + \bar{P}_\beta^{-1}\bar{\beta}\right),$$

then the sampling posterior distribution for $\mu_{\beta,0}$ is given by:

$$\mu_{\beta,0} \sim \mathcal{MN}\left(\widetilde{\mu}_{\beta,0},\ P_{\mu_\beta} \right),$$

$$\widetilde{\mu}_{\beta,0} = \Psi_{\mu_\beta}\psi_\beta,$$

$$P_{\mu_\beta} = \Psi_{\mu_\beta}\Psi_\beta,$$

$$\Psi_\mu = I_k - \Psi_\beta R'_\beta \left(R_\beta \Psi_\beta R'_\beta\right)^{-1} R_\beta.$$

2.4.1.3 SubStep: sum the TV and non-TV draws of β_t

We can simply sum the draws obtained in the two previous substeps

$$\forall t \in \{1, \ldots, T\}, \qquad \beta_t = \Lambda_\beta \widetilde{\beta}_t + M_\beta \beta_0,$$

to obtain a draw for the history of TV-VAR coefficients $(\beta_t^i)_{t=1}^T$.

2.4.2 Step 2: draw a reduced rank covariance matrix Q_β

Given the new histories of betas, that is conditioning on $(\beta_t^i)_{t=0}^T$, we can draw the reduced rank covariance matrix Q_β. Indeed, recalling that:

$$\beta_t = \beta_{t-1} + \nu_{\beta,t}, \qquad \nu_{\beta,t} \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{Q_\beta}_{k\times k}\right),$$

and given the random walk law of motion, having a history of betas is equivalent to have a complete sequence of innovations $\nu_{\beta,t}$. Indeed, stacking the $\nu_{\beta,t}$ across time, we get:

$$\underbrace{\nu_\beta^*}_{k\times T} = \begin{bmatrix} \nu_{\beta,1} & \nu_{\beta,2} & \dots & \nu_{\beta,T} \end{bmatrix},$$

and we can easily compute the innovations sum of squares matrix:

$$\underbrace{S_\beta}_{k\times k} = \underbrace{\nu_\beta^*}_{k\times T} \underbrace{\nu_\beta^{*\prime}}_{T\times k}.$$

By construction, the rank of S_β would be $q_\beta < k$. If the prior on matrix Q_β is a $k \times k$ Singular Inverse Wishart with rank $q_\beta < k$, scale matrix $\bar{Q}_\beta$ and degrees of freedom $\tau_{\beta,0}$, that is

$$Q_\beta \sim \mathcal{SIW}_k^{q_\beta}\left(\bar{Q}_\beta,\ \tau_{\beta,0}\right),$$

then the posterior is conjugate and given by:

$$Q_\beta \big| \left(\beta_t^i\right)_{t=0}^T \sim \mathcal{SIW}_k^{q_\beta}\left(S_\beta + \bar{Q}_\beta,\ \tau_{\beta,0} + T\right).$$

2.4.2.1 Zoom in: draw from a Singular Inverse Wishart distribution

Consider the $m \times m$ matrix Ξ having the following $\mathcal{SIW}$ distribution:

$$\Xi \sim \mathcal{SIW}_m^{\,q}\left(\underbrace{S}_{m\times m},\ d\right),$$

where S is a scale matrix, d the degrees of freedom, and q is the rank of both Ξ and S.

Following Bodnar and Okhrin (2008) and Dıaz-Garcıa et al. (1997), and applying some modifications tailored at improving the algorithm efficiency, it is possible to draw from the distribution of Ξ through the following steps:

1. Construct the Moore–Penrose pseudoinverse S^+ of the scale matrix S. Moore Penrose computation follows the highly efficient Qrginv algorithm proposed by Ataei (2014).

2. Construct the diagonal matrix $\underbrace{\Lambda_S}_{q\times q}$ containing the non-zero eigenvalues of S^+ in the main diagonal, and the matrix $\underbrace{U_S}_{m\times q}$ whose columns are the associated eigenvectors.

3. Draw d independent vectors $\underbrace{\widetilde{z}}_{q\times 1}$ from the multivariate normal $\mathcal{MN}\left(\ \mathbf{0},\ \Lambda_S\ \right)$ and stack them as columns in the matrix $\underbrace{\widetilde{Z}}_{q\times d}$.

4. Construct the matrices $\underbrace{Z}_{m\times q} = U_S\widetilde{Z}$ and $\underbrace{W}_{m\times m} = ZZ'$. The matrix W is a draw from a Singular Wishart distribution with scale matrix S^+ and d degrees of freedom.

5. Compute the Moore–Penrose pseudoinverse W^+ of W. The matrix $\widetilde{\Xi} = W^+$ is a draw from the Singular Inverse Wishart of interest.

2.4.3 Step 2: draw a history of off-diagonal elements $(\alpha_t^i)_{t=1}^T$

Given the previous step of the Markov chain $\left(\left(\alpha_t^{i-1},\ \sigma_t^{i-1}\right)_{t=1}^T, Q_\alpha^{i-1}, Q_\sigma^{i-1}\right)$ and the just drawn history of TV-VAR coefficients $(\beta_t^i)_{t=1}^T$ along with the drawn covariance matrix of coefficients innovation Q^i_β, we want to draw from:

$$f\left((\alpha_t)_{t=1}^T \middle| \left(\beta_t^i,\ \sigma_t^{i-1}\right)_{t=1}^T,\ Q^i_\beta, Q^{i-1}_\alpha, Q^{i-1}_\sigma \right).$$

Also in this case, we draw separately the transformed time-varying components $\Lambda_\alpha \widetilde{\alpha}_t$ and the non-TV component $M_\alpha \alpha_0$, since

$$\alpha_t = \Lambda_\alpha \widetilde{\alpha}_t + M_\alpha \alpha_0.$$

2.4.3.1 SubStep: draw a history of TV components $(\widetilde{\alpha}_t^i)_{t=1}^T$

At this step we can condition on:

$$\left(M_\alpha \alpha_0,\ \Lambda_\alpha, (\beta_t,\ \sigma_t,\ y_t,\ X_t)_{t=1}^T \right).$$

From the stacked version of the VAR we can write:

$$y_t = X_t' \beta_t + u_t = X_t' \beta_t + A_t^{-1} \Sigma_t \varepsilon_t,$$

$$\underbrace{y_t - X_t' \beta_t}_{\widehat{y}_t} = A_t^{-1} \Sigma_t \varepsilon_t,$$

$$A_t\ \widehat{y}_t = \Sigma_t \varepsilon_t.$$

Focusing on matrix A_t, we can rewrite it as sum of the following matrices:

$$\underbrace{A_t}_{n\times n} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ \alpha_{1,t} & 1 & 0 & \dots & \vdots \\ \alpha_{2,t} & \alpha_{3,t} & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \alpha_{r-j-1,t} & \alpha_{r-j,t} & \dots & \alpha_{r,t} & 1 \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}}_{I_n} + \underbrace{\begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ \alpha_{1,t} & 0 & 0 & \dots & 0 \\ \alpha_{2,t} & \alpha_{3,t} & 0 & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \alpha_{r-j-1,t} & \alpha_{r-j,t} & \dots & \alpha_{r,t} & 0 \end{bmatrix}}_{A_t^*} = I_n + A_t^*.$$

We can then rewrite the model as:

$$\begin{aligned} A_t\,\widehat{y}_t &= \Sigma_t \varepsilon_t, \\ (I_n + A_t^*)\,\widehat{y}_t &= \Sigma_t \varepsilon_t, \\ \widehat{y}_t &= -A_t^* \widehat{y}_t + \Sigma_t \varepsilon_t. \end{aligned}$$

Using the properties of the *vec* and Kronecker product, the following holds:

$$-\,I_n \underbrace{A_t^*}_{n\times n} \underbrace{\widehat{y}_t}_{n\times 1} = -\underbrace{(I_n \otimes \widehat{y}_t')}_{n\times n^2} \underbrace{vec\,(A_t^{*\prime})}_{n^2\times 1},$$

where the vector on the right hand-side is:

$$vec\,(A_t^{*\prime}) = \begin{bmatrix} 0 & \dots & 0 & \alpha_{1,t} & 0 & \dots & 0 & \alpha_{2,t} & \alpha_{3,t} & 0 & \dots & \dots & \alpha_{r-j-1,t} & \dots & \alpha_{r,t} & 0 \end{bmatrix}',$$

with the zeros having positions $[(i-1)\,n + j]_{j\in\{1,\dots,n\}}^{i\in\{1,\dots,n\}}$. By removing such zeros we obtain exactly the elements below the main diagonal of A_t gathered in the vector α_t. Removing the corresponding columns in $-\left(I_n \otimes \widehat{y}_t'\right)$, we define Z_t, which has the following form:

$$\underbrace{Z_t}_{n\times r} = \begin{bmatrix} 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \widehat{y}_{1,t} & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ 0 & \widehat{y}_{1,t} & \widehat{y}_{2,t} & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \vdots \\ 0 & 0 & 0 & \widehat{y}_{1,t} & \widehat{y}_{2,t} & \widehat{y}_{3,t} & 0 & \cdots & \cdots & \cdots & \cdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & \widehat{y}_{1,t} & \widehat{y}_{2,t} & \widehat{y}_{3,t} \;\cdots & \widehat{y}_{n-1,t} \end{bmatrix}.$$

We can finally rewrite the VAR as measurement equation in the following form:

$$\begin{aligned} \widehat{y}_t &= -A_t^* \widehat{y}_t + \Sigma_t \varepsilon_t, \\ \widehat{y}_t &= -\left(I_n \otimes \widehat{y}_t'\right) vec\left(A_t^{*\prime}\right) + \Sigma_t \varepsilon_t, \\ \widehat{y}_t &= Z_t \alpha_t + \Sigma_t \varepsilon_t. \end{aligned}$$

Considering the form of Z_t, this measurement equation is such that each equation of $\widehat{y}_{i,t}$ conditions on all observables $(\widehat{y}_{j,t})_{j=1}^{i-1}$ (for $i \geq 2$). This is where de Wind and Gambetti (2014) claim that running the Kalman filter equation by equation (as in Primiceri, 2005) is unnecessary, since the matrix Z_t on its own is able to feature the recursive conditioning (notice also that the first variable is not related to the states). A small numerical experiment executed using simulated data confirms their claim. In other words, the actual observables of the implemented Kalman filter will be:

$$\begin{bmatrix} \widehat{y}_{2,t} \,|\, \widehat{y}_{1,t} \\ \widehat{y}_{3,t} \,|\, \widehat{y}_{1,t},\; \widehat{y}_{2,t} \\ \vdots \\ \widehat{y}_{n,t} \,|\, \widehat{y}_{1,t},\; \widehat{y}_{2,t}, \ldots, \widehat{y}_{n-2,t},\; \widehat{y}_{n-1,t} \end{bmatrix}$$

Nevertheless, for notational simplicity, this cumbersome formulation won't be used. We can now use the decomposition of the alphas and the law of motion of the TV components:

$$\underbrace{\alpha_t}_{r\times 1} = \Lambda_\alpha \widetilde{\alpha}_t + M_\alpha \alpha_0, \quad \underbrace{\widetilde{\alpha}_t}_{q_\alpha \times 1} = \widetilde{\alpha}_{t-1} + \widetilde{\nu}_{\alpha,t}, \quad \underbrace{\widetilde{\nu}_{\alpha,t}}_{q_\alpha \times 1} \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\; I_{q_\alpha}\right), \quad Q_\alpha = \underbrace{\Lambda_\alpha}_{r\times q_\alpha} \Lambda_\alpha',$$

and transform the observable equation in:

$$\begin{aligned}
\widehat{y}_t &= Z_t \alpha_t + \Sigma_t \varepsilon_t, \\
\widehat{y}_t &= Z_t \left(\Lambda_\alpha \widetilde{\alpha}_t + M_\alpha \alpha_0 \ \right) + \Sigma_t \varepsilon_t, \\
\underbrace{\widehat{y}_t - Z_t M_\alpha \alpha_0}_{\widehat{y}_t^*} &= \underbrace{Z_t \Lambda_\alpha}_{Z_t^*} \ \widetilde{\alpha}_t + \Sigma_t \varepsilon_t.
\end{aligned}$$

The state space form in this case becomes:

$$\begin{aligned}
\widehat{y}_t^* &= Z_t^* \ \widetilde{\alpha}_t + \Sigma_t \varepsilon_t, & \varepsilon_t &\overset{iid}{\sim} \mathcal{MN} \left(\ \mathbf{0}, \ I_n \ \right), \\
\underbrace{\widetilde{\alpha}_t}_{q_\alpha \times 1} &= \widetilde{\alpha}_{t-1} + \widetilde{\nu}_{\alpha,t}, & \underbrace{\widetilde{\nu}_{\alpha,t}}_{q_\alpha \times 1} &\overset{iid}{\sim} \mathcal{MN} \left(\ \mathbf{0}, \ I_{q_\alpha} \ \right).
\end{aligned}$$

We use also in this case the FFBS procedure. Given the prior moments $\left(\bar{\alpha}, \ \bar{P}_\alpha\right)$ for α_0, we initizialize the filter at:

$$\widetilde{\alpha}_{0|0} = R_\alpha \bar{\alpha}, \qquad P_{\alpha,0|0} = R_\alpha \bar{P}_\alpha R_\alpha'.$$

Then, given $\left(\widetilde{\alpha}_{t-1|t-1}, \ P_{\alpha,t-1|t-1}\right)$, we compute:

$$\begin{aligned}
P_{\alpha,t|t-1} &= P_{\alpha,t-1|t-1} + I_{q_\alpha}, \\
K_{\alpha,t} &= P_{\alpha,t|t-1} Z_t^{*\prime} \left(Z_t^* P_{\alpha,t|t-1} Z_t^{*\prime} + \Sigma_t^2 \right)^{-1}, \\
\widetilde{\alpha}_{t|t} &= \widetilde{\alpha}_{t-1|t-1} + K_{\alpha,t} \left(\widehat{y}_t^* - Z_t^{*\prime} \ \widetilde{\alpha}_{t-1|t-1} \right), \\
P_{\alpha,t|t} &= P_{\alpha,t|t-1} - K_{\alpha,t} Z_t^* P_{\alpha,t|t-1}.
\end{aligned}$$

With an entire set of updating and prediction steps $\left(\widetilde{\alpha}_{\alpha,t|t}, \ P_{\alpha,t|t}, \ P_{t|t-1}\right)_{t=1}^T$, we start to sample backward, beginning by sampling $\widetilde{\alpha}_T$ from $\mathcal{MN} \left(\ \widetilde{\alpha}_{T|T}, \ P_{\alpha,T|T} \ \right)$, and then for each $t \ \in \{T-1, T-2, \ldots, 2, 1\}$ we sample recursively each $\widetilde{\alpha}_t$ from $\mathcal{MN} \left(\ \widetilde{\alpha}_{t|t+1}, \ P_{\alpha,t|t+1} \ \right)$ where:

$$\begin{aligned}
\widetilde{\alpha}_{t|t+1} &= \widetilde{\alpha}_{t|t} + P_{\alpha,t|t} P_{\alpha,t+1|t}^{-1} \left(\widetilde{\alpha}_{t+1} - \widetilde{\alpha}_{t|t} \right), \\
P_{\alpha,t|t+1} &= P_{\alpha,t|t} - P_{\alpha,t|t} P_{\alpha,t+1|t}^{-1} P_{\alpha,t|t}.
\end{aligned}$$

2.4.3.2 SubStep: draw a vector of time invariant component $M_\alpha \alpha_0$ of α_t

At this point, we can condition on the previous step of the chain and on the new draw $(\widetilde{\alpha}_t^i)_{t=1}^T$:

$$\left(\Lambda_\alpha, (\widetilde{\alpha}_t, \beta_t,\ \sigma_t,\ y_t,\ X_t)_{t=1}^T\right),$$

and transform the VAR model as:

$$\widehat{y}_t = Z_t \alpha_t + \Sigma_t \varepsilon_t,$$

$$\widehat{y}_t = Z_t \left(\Lambda_\alpha \widetilde{\alpha}_t + M_\alpha \alpha_0\ \right) + \Sigma_t \varepsilon_t,$$

$$\underbrace{\widehat{y}_t - Z_t \Lambda_\alpha \widetilde{\alpha}_t}_{\widehat{y}_t^\bullet} = Z_t \underbrace{M_\alpha \alpha_0}_{\mu_{\alpha,0}} + \underbrace{\Sigma_t \varepsilon_t}_{\widehat{\varepsilon}_t},$$

$$\widehat{y}_t^\bullet = Z_t \mu_{\alpha,0} + \widehat{\varepsilon}_t, \qquad \widehat{\varepsilon}_t \overset{iid}{\sim} \mathcal{MN}\left(\ \mathbf{0},\ \Sigma_t^{\cdot 2}\ \right).$$

As we did for $M_\beta \beta_0$ we now impose the $q_\alpha < r$ restrictions given by:

$$\underbrace{R_\alpha}_{q_\alpha \times r} \underbrace{\mu_{\alpha,0}}_{r \times 1} = \mathbf{0}_{q_\alpha \times 1}.$$

As explained and done for the betas, we can stack across the time dimension and write the restricted regression model as:

$$\underbrace{\widehat{Y}^\bullet}_{nT \times 1} = \underbrace{Z}_{nT \times r} \mu_{\alpha,0} + \underbrace{\xi}_{nT \times 1}, \qquad \xi \sim \mathcal{MN}\left(\ \mathbf{0}_{nT \times 1},\ \Sigma^{\cdot 2}\ \right),$$

$$\underbrace{R_\alpha}_{q_\alpha \times r} \underbrace{\mu_{\alpha,0}}_{r \times 1} = \mathbf{0}_{q_\alpha \times 1}.$$

Then, given the prior for α_0 and the unrestricted Bayesian regression moments:

$$\alpha_0 \sim \mathcal{MN}\left(\ \bar{\alpha},\ \bar{P}_\alpha\ \right),$$

$$\Psi_\alpha = \left(Z' \Sigma^{\cdot -2} Z + \bar{P}_\alpha^{-1}\right)^{-1},$$

$$\psi_\alpha = \Psi_\alpha \left(Z' \Sigma^{\cdot -2}\ \widehat{Y}^\bullet + \bar{P}_\alpha^{-1} \bar{\alpha}\right),$$

then the sampling posterior distribution for $\mu_{\alpha,0}$ is given by:

$$\mu_{\alpha,0} \sim \mathcal{MN}\left(\ \widetilde{\mu}_{\alpha,0}\ ,\ P_{\mu_\alpha}\ \right),$$

$$\begin{aligned}
\widetilde{\mu}_{\alpha,0} &= \Psi_{\mu_\alpha}\psi_\alpha, \\
P_{\mu_\alpha} &= \Psi_{\mu_\alpha}\Psi_\alpha, \\
\Psi_{\mu_\alpha} &= I_r - \Psi_\alpha R'_\alpha \left(R_\alpha \Psi_\alpha R'_\alpha\right)^{-1} R_\alpha.
\end{aligned}$$

2.4.3.3 SubStep: sum the TV and non-TV draws of α_t

Summing the draws obtained in the two previous substeps

$$\forall t \in \{1, \ldots, T\}, \qquad \alpha_t = \Lambda_\alpha \widetilde{\alpha}_t + M_\alpha \alpha_0,$$

we obtain a history of TV off-diagonal elements $(\alpha_t^i)_{t=1}^T$.

2.4.4 Step 4: draw a reduced rank covariance matrix Q_α

Conditioning on the new $(\alpha_t^i)_{t=0}^T$, we can draw the reduced rank covariance matrix Q_α. Indeed, recalling their law of motion:

$$\alpha_t = \alpha_{t-1} + \nu_{\alpha,t}, \qquad \nu_{\alpha,t} \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{Q_\alpha}_{r\times r}\right),$$

having a complete history of the alphas is equivalent to have a complete histories of innovations $\nu_{\alpha,t}$. Stacking the $\nu_{\alpha,t}$ across time, we get:

$$\underbrace{\nu_\alpha^*}_{r\times T} = \begin{bmatrix} \nu_{\alpha,1} & \nu_{\alpha,2} & \ldots & \nu_{\alpha,T} \end{bmatrix},$$

and we can easily compute the innovations sum of squares matrix:

$$\underbrace{S_\alpha}_{r\times r} = \underbrace{\nu_\alpha^*}_{r\times T}\underbrace{\nu_\alpha^{*\prime}}_{T\times r}.$$

By construction, the rank of S_α would be $q_\alpha < r$. Then, if the prior on matrix Q_α is a $r \times r$ singular inverse-wishart with rank $q_\alpha < r$, scale matrix $\bar{Q}_\alpha$ and degrees of freedom

$\tau_{\alpha,0}$:

$$Q_\alpha \sim \mathcal{SIW}_r^{q_\alpha}\left(\bar{Q}_\alpha,\ \tau_{\alpha,0}\right),$$

then the posterior is conjugate and given by:

$$Q_\alpha \big| \left(\alpha_t^i\right)_{t=0}^T \sim \mathcal{SIW}_r^{q_\alpha}\left(S_\alpha + \bar{Q}_\alpha,\ \tau_{\alpha,0} + T\right).$$

2.4.5 Step 5: draw a history of volatilities $(\sigma_t^i)_{t=1}^T$

Given Q_σ^{i-1} from the previous step of the Markov chain and the just drawn history of TV-VAR coefficients and off-diagonal elements $(\beta_t^i, \alpha_t^i)_{t=1}^T$ along with the drawn covariance matrix of coefficients innovations $\left\{Q_\alpha^i, Q_\beta^i\right\}$, we want to draw from:

$$f\left(\left(\sigma_t\right)_{t=1}^T \Big| \left(\alpha_t^i, \beta_t^i,\ \sigma_t^{i-1}\right)_{t=1}^T,\ Q_\beta^i, Q_\alpha^i, Q_\sigma^{i-1}\right).$$

Also in this case we draw separately the transformed time-varying components $\Lambda_\sigma \widetilde{\sigma}_t$ and the non-TV component $M_\sigma \log \sigma_0$, since

$$\log \sigma_t = \Lambda_\sigma \widetilde{\sigma}_t + M_\sigma \log \sigma_0.$$

We can transform differently the VAR formulation used in the previous steps to obtain:

$$y_t = X_t'\beta_t + u_t = X_t'\beta_t + A_t^{-1}\Sigma_t\varepsilon_t,$$

$$\underbrace{A_t\left(y_t - X_t'\beta_t\right)}_{\widetilde{y}_t} = \Sigma_t\varepsilon_t,$$

$$\widetilde{y}_t = \Sigma_t\varepsilon_t,$$

where

$$\underbrace{\Sigma_t}_{n\times n} = \begin{bmatrix} \sigma_{1,t} & 0 & \dots & 0 \\ 0 & \sigma_{2,t} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_{n,t} \end{bmatrix} = diag\left(\begin{bmatrix} \sigma_{1,t} \\ \sigma_{2,t} \\ \vdots \\ \sigma_{n,t} \end{bmatrix}\right) = diag\left(\underbrace{\sigma_t}_{n\times 1}\right),$$

$$\Sigma_t \varepsilon_t = \Sigma_t \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \vdots \\ \varepsilon_{n,t} \end{bmatrix} = \begin{bmatrix} \sigma_{1,t}\ \varepsilon_{1,t} \\ \sigma_{2,t}\ \varepsilon_{2,t} \\ \vdots \\ \sigma_{n,t}\ \varepsilon_{n,t} \end{bmatrix} = \sigma_t \odot \varepsilon_t.$$

We then take the square element by element on both sides of our transformation of the VAR:

$$(\widetilde{y}_t)^{\cdot 2} = (\Sigma_t \varepsilon_t)^{\cdot 2} \quad \Longleftrightarrow \quad \begin{bmatrix} \widetilde{y}^2_{1,t} \\ \widetilde{y}^2_{2,t} \\ \vdots \\ \widetilde{y}^2_{n,t} \end{bmatrix} = \begin{bmatrix} \sigma^2_{1,t}\ \varepsilon^2_{1,t} \\ \sigma^2_{2,t}\ \varepsilon^2_{2,t} \\ \vdots \\ \sigma^2_{n,t}\ \varepsilon^2_{n,t} \end{bmatrix}$$

We add on the left hand-side a small constant[1] $\bar{c} = 10^{-3}$ and applying the logarithm on both sides[2], we obtain:

$$\underbrace{\begin{bmatrix} \log\left(\widetilde{y}^2_{1,t} + \bar{c}\right) \\ \log\left(\widetilde{y}^2_{2,t} + \bar{c}\right) \\ \vdots \\ \log\left(\widetilde{y}^2_{n,t} + \bar{c}\right) \end{bmatrix}}_{\widetilde{y}^*_t} = 2 \begin{bmatrix} \log \sigma_{1,t} \\ \log \sigma_{2,t} \\ \vdots \\ \log \sigma_{n,t} \end{bmatrix} + \begin{bmatrix} \log \varepsilon^2_{1,t} \\ \log \varepsilon^2_{2,t} \\ \vdots \\ \log \varepsilon^2_{n,t} \end{bmatrix},$$

$$\widetilde{y}^*_t = 2 \log \sigma_t + \log\left[(\varepsilon_t)^{\cdot 2}\right].$$

Given that $\varepsilon_t \overset{iid}{\sim} \mathcal{MN}\left(\,\mathbf{0},\ I_n\,\right)$, each $\log \varepsilon^2_{j,t}$ is a $\log \chi^2_1$ distribution, and $\log\left[(\varepsilon_t)^{\cdot 2}\right]$ is a vector of independent $\log \chi^2_1$. Kim et al. (1998) document that the $\log \chi^2_1$ can be sufficiently

[1] The addition of a small constant term has numerical stability purposes, as explained in Fuller (2009) and Primiceri (2005).

[2] Noticing that $\log \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \log a \\ \log b \end{bmatrix}$.

well approximated by a mixture of seven normal distributions:

$$f_{\log \chi_1^2}(x) \approx \sum_{j=1}^{7} m_j^p \; f_{\mathcal{N}}\left(x|\, m_j^m - 1.2704, m_j^v\right),$$

where m_j^p, m_j^m and m_j^v are respectively the weights, the mean and the variance of the normal distributions $f_{\mathcal{N}}\left(\cdot|\, \mu, \sigma^2\right)$ that compose the mixture, reported in the following table:

j	m_j^p	m_j^m	m_j^v
1	0.00730	-10.12999	5.79596
2	0.10556	-3.97281	2.61369
3	0.00002	-8.56686	5.17950
4	0.04395	2.77786	0.16735
5	0.34001	0.61942	0.64009
6	0.24566	1.79518	0.34023
7	0.25750	-1.08819	1.26261

However, Omori et al. (2007) , improving upon Kim et al. (1998), show that the $\log \chi_1^2$ is even better approximated by a mixture of 10 Normal distributions:

$$f_{\log \chi_1^2}(x) \approx \sum_{j=1}^{10} m_j^p \; f_{\mathcal{N}}\left(x|\, m_j^m, m_j^v\right),$$

where m_j^p, m_j^m and m_j^v are contained in the following table:

j	1	2	3	4	5	6	7	8	9	10
m_j^p	0.00609	0.04775	0.13057	0.20674	0.22715	0.18842	0.12047	0.05591	0.01575	0.00115
m_j^m	1.92677	1.34744	0.73504	0.02266	-0.85173	-1.97278	-3.46788	-5.55246	-8.68384	-14.65000
m_j^v	0.11265	0.17788	0.26768	0.40611	0.62699	0.98583	1.57469	2.54498	4.16591	7.33342

The latter and more precise approximation will be used in this work to draw stochastic volatilities.

Therefore, in order to have a conditionally Gaussian measurement equation, we should condition to a specific history of normal components of the mixtures. Indeed, let's define the matrix S, whose element (h, t) is such that:

$$\underbrace{S}_{n\times T} = \left[s_{h,t}\right]_{h\in\{1,\ldots,n\}}^{t\in\{1,\ldots,T\}}, \qquad s_{h,t} = j \iff \log \varepsilon_{h,t}^2 \sim \mathcal{N}\left(m_j^m, m_j^v\right)$$

We can implement the estimation of an history of stochastic volatilies by conditioning on a history of specific indexes contained in the matrix S. The procedure is then divided in three substeps[3].

2.4.5.1 SubStep: Draw a matrix S_i conditional on $\left(\sigma_t^{i-1}\right)_{t=1}^T$

Given the volatilies of the previous point of the chain, $\left(\sigma_t^{i-1}\right)_{t=1}^T$, we consider the following model form:

$$\log\left[(\varepsilon_t)^{\cdot 2}\right] = \widetilde{y}_t^* - 2\log\sigma_t,$$

where $\left(\log\left[(\varepsilon_t)^{\cdot 2}\right]\right)_{t=1}^T$ is a sequence of vectors of independent (across times and equations) and identically distributed $\log\chi_1^2$, which are approximated by the mixtures of ten normal components. Then, the element $s_{h,t}$, which indexes the specific component from which $\log\varepsilon_{h,t}^2$ is drawn, has support $J = \{1,\ldots,10\}$, and the following discrete probability distribution:

$$\forall j\in J,\qquad Pr\left[s_{h,t}=j\middle|\,\widetilde{y}_{h,t}^*,\sigma_t\right] = \frac{m_j^p\ f_{\mathcal{N}}\left(\widetilde{y}_{h,t}^* - 2\log\sigma_{h,t}\middle|\,m_j^m, m_j^v\right)}{\sum_{\iota=1}^{10} m_\iota^p\ f_{\mathcal{N}}\left(\widetilde{y}_{h,t}^* - 2\log\sigma_{h,t}\middle|\,m_\iota^m, m_\iota^v\right)}.$$

Independent draws for all variables $h\in\{1,\ldots,n\}$ at all periods $t\in\{1,\ldots,T\}$ from this distribution will form the new matrix S_i.

2.4.5.2 SubStep: draw a history of TV components $(\widetilde{\sigma}_t^i)_{t=1}^T$

Considering the transformation of the VAR form derived in the previous substep

$$\widetilde{y}_t^* = 2\log\sigma_t + \log\left[(\varepsilon_t)^{\cdot 2}\right],$$

and the reduced rank decomposition of $\log\sigma_t$:

$$\log\sigma_t = \Lambda_\sigma\widetilde{\sigma}_t + M_\sigma\log\sigma_0,$$

[3] The order follows the corrigendum work of Del Negro and Primiceri (2015).

we combine them to derive:

$$\widetilde{y}_t^{*} = 2\Lambda_\sigma \widetilde{\sigma}_t + 2M_\sigma \log \sigma_0 + \log\left[(\varepsilon_t)^{\cdot 2}\right],$$

$$\underbrace{\widetilde{y}_t^{*} - 2M_\sigma \log \sigma_0}_{\widetilde{y}_t^{**}} = 2\Lambda_\sigma \widetilde{\sigma}_t + \log\left[(\varepsilon_t)^{\cdot 2}\right].$$

Then, conditioning on the just drawn matrix S_i specifying the normal components' indexes for the volatilities, the vector $\log\left[(\varepsilon_t)^{\cdot 2}\right]$ has (approximately) the following Gaussian distribution:

$$\log\left[(\varepsilon_t)^{\cdot 2}\right] \sim \mathcal{MN}\left(\underbrace{\begin{bmatrix} m^m_{s_{1,t}} \\ m^m_{s_{2,t}} \\ \vdots \\ m^m_{s_{n,t}} \end{bmatrix}}_{\varphi_t}, \underbrace{\begin{bmatrix} m^v_{s_{1,t}} & 0 & \dots & 0 \\ 0 & m^v_{s_{2,t}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & m^v_{s_{n,t}} \end{bmatrix}}_{\Upsilon_t} \right).$$

At this point, we can define our state space form as:

$$\widetilde{y}_t^{**} = \varphi_t + 2\Lambda_\sigma \widetilde{\sigma}_t + \zeta_t, \qquad \zeta_t \overset{iid}{\sim} \mathcal{MN}\left(\ \mathbf{0},\ \Upsilon_t\ \right),$$

$$\underbrace{\widetilde{\sigma}_t}_{q_\sigma \times 1} = \widetilde{\sigma}_{t-1} + \widetilde{\nu}_{\sigma,t}, \qquad \underbrace{\widetilde{\nu}_{\sigma,t}}_{q_\sigma \times 1} \overset{iid}{\sim} \mathcal{MN}\left(\ \mathbf{0},\ I_{q_\sigma}\ \right).$$

Using the FFBS procedure, we assume the moments $\left(\bar{\sigma},\ \bar{P}_\sigma\right)$ of the prior distribution of $\log \sigma_0$, and we initialize the filter at:

$$\widetilde{\sigma}_{0|0} = R_\sigma \bar{\sigma}, \qquad P_{\sigma,0|0} = R_\sigma \bar{P}_\sigma R'_\sigma.$$

Recursively, for each $\left(\widetilde{\sigma}_{t-1|t-1},\ P_{\sigma,t-1|t-1}\right)$, we compute the filter:

$$P_{\sigma,t|t-1} = P_{\sigma,t-1|t-1} + I_{q_\sigma},$$
$$K_{\sigma,t} = 2P_{\sigma,t|t-1}\Lambda'_\sigma \left(4\Lambda_\sigma P_{\sigma,t|t-1}\Lambda'_\sigma + \Upsilon_t\right)^{-1},$$
$$\widetilde{\sigma}_{t|t} = \widetilde{\sigma}_{t-1|t-1} + K_{\sigma,t}\left(\widetilde{y}_t^{**} - 2\Lambda'_\sigma\ \widetilde{\sigma}_{t-1|t-1} - \varphi_t\right),$$
$$P_{\sigma,t|t} = P_{\sigma,t|t-1} - 2K_{\sigma,t}\Lambda_\sigma P_{\sigma,t|t-1}.$$

Having an entire set of updating and prediction steps $\left(\widetilde{\sigma}_{t|t},\ P_{\sigma,t|t},\ P_{\sigma,t|t-1}\right)_{t=1}^T$, we start to sample backward, beginning by sampling $\widetilde{\sigma}_T$ from $\mathcal{MN}\left(\ \widetilde{\sigma}_{T|T},\ P_{\sigma,T|T}\ \right)$, and then for each $t\ \in \{T-1, T-2, \ldots, 2, 1\}$ we sample recursively each $\widetilde{\sigma}_t$ from $\mathcal{MN}\left(\ \widetilde{\sigma}_{t|t+1},\ P_{\sigma,t|t+1}\ \right)$ where:

$$\widetilde{\sigma}_{t|t+1} = \widetilde{\sigma}_{t|t} + P_{\sigma,t|t}P^{-1}_{\sigma,t+1|t}\left(\widetilde{\sigma}_{t+1} - \widetilde{\sigma}_{t|t}\right),$$
$$P_{\sigma,t|t+1} = P_{\sigma,t|t} - P_{\sigma,t|t}P^{-1}_{\sigma,t+1|t}P_{\sigma,t|t}.$$

2.4.5.3 SubStep: draw a vector of time invariant component $M_\sigma \log\sigma_0$ of σ_t

At this step we can condition on $\left(\Lambda_\sigma, S, (\widetilde{\sigma}_t, \alpha_t, \beta_t,\ y_t,\ X_t)_{t=1}^T\right)$. As done similarly for $M_\beta\beta_0$ and $M_\alpha\alpha_0$, let's decompose the measurement equation in the following way:

$$\widetilde{y}_t^* = 2\Lambda_\sigma\widetilde{\sigma}_t + 2M_\sigma\log\sigma_0 + \underbrace{\log\left[(\varepsilon_t)^{\cdot 2}\right]}_{\varphi_t + \zeta_t},$$

$$\underbrace{\widetilde{y}_t^* - 2\Lambda_\sigma\widetilde{\sigma}_t - \varphi_t}_{\widetilde{y}_t^\diamond} = \underbrace{2M_\sigma\log\sigma_0}_{\mu_{\sigma,0}} + \zeta_t,$$

$$\widetilde{y}_t^\diamond = \mu_{\sigma,0} + \zeta_t, \qquad \zeta_t \overset{iid}{\sim} \mathcal{MN}\left(\ \mathbf{0},\ \Upsilon_t\ \right).$$

Also in this case we want to impose the $q_\sigma < n$ restrictions given by:

$$\underbrace{R_\sigma}_{q_\sigma\times n}\ \underbrace{\mu_{\sigma,0}}_{n\times 1} = \mathbf{0}_{q_\sigma\times 1}.$$

We can stack the restricted regression model as:

$$\underbrace{\widetilde{Y}^{\diamond}}_{nT\times 1} = \underbrace{\left(\underbrace{\begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}}_{T\times 1} \otimes I_n \right)}_{\mathcal{I}_{nT}} \mu_{\sigma,0} + \zeta^*, \qquad \zeta^* \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ \Upsilon^* \right),$$

$$\underbrace{R_\sigma}_{q_\sigma \times n} \underbrace{\mu_{\sigma,0}}_{n\times 1} = \mathbf{0}_{q_\sigma \times 1}.$$

Then, given the prior distribution for $\log \sigma_0$ and the unrestricted Bayesian regression moments:

$$\begin{aligned}
\log \sigma_0 &\sim \mathcal{MN}\left(\bar{\sigma},\ \bar{P}_\sigma \right), \\
\Psi_\sigma &= \left(\mathcal{I}'_{nT}\ \Upsilon^{*-1}\ \mathcal{I}_{nT} + \bar{P}_\sigma^{-1} \right)^{-1}, \\
\psi_\sigma &= \Psi_\sigma \left(\mathcal{I}'_{nT}\ \Upsilon^{*-1}\ \widetilde{Y}^{\diamond} + \bar{P}_\sigma^{-1} \bar{\sigma} \right),
\end{aligned}$$

the sampling posterior distribution for $\mu_{\sigma,0}$ is given by:

$$\begin{aligned}
\mu_{\sigma,0} &\sim \mathcal{MN}\left(\widetilde{\mu}_{\sigma,0}\ ,\ P_{\mu_\sigma} \right), \\
\widetilde{\mu}_{\sigma,0} &= \Psi_{\mu_\sigma} \psi_\sigma, \\
P_{\mu_\sigma} &= \Psi_{\mu_\sigma} \Psi_\sigma, \\
\Psi_{\mu_\sigma} &= I_n - \Psi_\sigma R'_\sigma \left(R_\sigma \Psi_\sigma R'_\sigma \right)^{-1} R_\sigma.
\end{aligned}$$

2.4.5.4 SubStep: sum the TV and non-TV draws of $\log \sigma_t$

We can sum the draws obtained in the two previous substeps

$$\forall t \in \{1, \ldots, T\}, \qquad \log \sigma_t = \Lambda_\sigma \widetilde{\sigma}_t + M_\sigma \log \sigma_0,$$

and apply the exponential operator element by element, to obtain a history of stochastic volatilities $(\sigma_t^i)_{t=1}^T$.

2.4.6 Step 6: draw a reduced rank covariance matrix Q_σ

Conditioning on the new draw $(\sigma_t^i)_{t=0}^T$, we can draw the reduced rank covariance matrix Q_σ. Indeed, recalling that:

$$\log \sigma_t = \log \sigma_{t-1} + \nu_{\sigma,t}, \qquad \nu_{\sigma,t} \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{Q_\sigma}_{n\times n}\right),$$

having a complete history of the sigmas is equivalent to have a complete histories of innovations $\nu_{\sigma,t}$. Stacking the $\nu_{\sigma,t}$ across time, we get:

$$\underbrace{\nu_\sigma^*}_{n\times T} = \begin{bmatrix}\nu_{\sigma,1} & \nu_{\sigma,2} & \dots & \nu_{\sigma,T}\end{bmatrix},$$

and we can easily compute the innovations sum of squares matrix:

$$\underbrace{S_\sigma}_{n\times n} = \underbrace{\nu_\sigma^*}_{n\times T}\underbrace{\nu_\sigma^{*\prime}}_{T\times n}.$$

By construction, the rank of S_σ would be $q_\sigma < n$. Then, if the prior on the matrix Q_σ is a $n \times n$ singular inverse-wishart with rank $q_\sigma < n$, scale matrix $\bar{Q}_\sigma$ and degrees of freedom $\tau_{\sigma,0}$:

$$Q_\sigma \sim \mathcal{SIW}_n^{q_\sigma}\left(\bar{Q}_\sigma,\ \tau_{\sigma,0}\right),$$

then the posterior is conjugate and given by:

$$Q_\sigma \big| \left(\sigma_t^i\right)_{t=0}^T \sim \mathcal{SIW}_n^{q_\sigma}\left(S_\sigma + \bar{Q}_\sigma,\ \tau_{\sigma,0} + T\right).$$

2.5 Other time-varying VAR model specifications

2.5.1 TV coefficients VAR with constant volatility (TV-VAR)

A simpler approach to parameter instability leaves coefficients to be time-varying but keeps the volatility constant (this is the strategy of Cogley and Sargent, 2001). A disadvantage of this approach, discussed in the literature (in particular by Sims, 2001), is that the potential shocks heteroskedasticity may be significantly exhanged for time variation of the VAR coefficients, potentially misleading any policy implications from the analysis.

As far as the implementation is concerned, we are referring to the full model:

$$y_t = X_t'\beta_t + u_t, \qquad u_t \overset{iid}{\sim} \mathcal{MN}\,(\,\mathbf{0},\ \Omega\,)\,.$$

Within the Gibbs Sampler, the time-varying coefficients can be estimated via the procedure described in the previous subsections, conditioning on the reduced form covariance matrix Ω. Given a history of coefficients $(\beta_t)_{t=1}^T$, it is possible to compute a history of reduced form residuals, and the matrix containing the sum of squared residuals $\widehat{\Omega}$:

$$\begin{aligned}
&\widehat{u}_t = y_t - X_t'\beta_t, \\
&\widehat{\Omega} = \widehat{U}\widehat{U}', \qquad\qquad \widehat{U} = \begin{bmatrix}\widehat{u}_1 & \widehat{u}_2 & \dots & \widehat{u}_T\end{bmatrix}.
\end{aligned}$$

Then, given the following conjugate prior for Ω:

$$\Omega \sim \mathcal{IW}\left(\bar{\Omega}, \tau_\Omega\right),$$

it is possible to draw from the posterior

$$\Omega \sim \mathcal{IW}\left(\bar{\Omega} + \widehat{\Omega}, \tau_\Omega + T\right).$$

2.5.2 Full rank TV-VAR-SV model

The full rank TV-VAR-SV model described in Primiceri (2005) is nested within the reduced rank model of de Wind and Gambetti (2014). To see this, referring to what showed in section 2.2, if matrix Q_z is full rank, then all eigenvalues are non-zero, and the matrix Λ_z, obtained via the eigenvector/eigenvalues decomposition is an invertible matrix.

$$Q_z = \underbrace{\Lambda_z}_{m\times m} \Lambda_z', \qquad \underbrace{\Lambda_z}_{m\times m} = \underbrace{V_z}_{m\times m}\underbrace{D_z}_{m\times m}.$$

In turn, this implies that

$$\begin{aligned} P_z &= \Lambda_z \left(\Lambda_z' \Lambda_z\right)^{-1} \Lambda_z' \\ &= \Lambda_z \left(\Lambda_z\right)^{-1} \left(\Lambda_z'\right)^{-1} \Lambda_z' \\ &= I_m, \end{aligned}$$

and

$$\begin{aligned} M_z &= I_m - P_z \\ &= I_m - I_m = \mathbf{0}, \end{aligned}$$

which means that there is no time invariant residual, and all components are time-varying. Therefore, the Gibbs Sampler for this model is not including any step regarding draws of time invariant components.

2.5.3 Constant Coefficients with Stochastic Volatility (CVAR-SV)

A VAR model with constant coefficients but with stochastic volatility (CVAR-SV) can be estimated modifying the coefficients step of the Gibbs Sampler of the TV-VAR-SV. The reduced form of the CVAR-SV is the following:

$$\begin{aligned} \underbrace{y_t}_{n\times 1} &= c + \sum_{\ell=1}^{p} \underbrace{B_\ell}_{n\times n} y_{t-\ell} + u_t, \qquad u_t \overset{i}{\sim} \mathcal{MN}\left(\mathbf{0}, \underbrace{\Omega_t}_{n\times n}\right), \\ \Omega_t &= A_t^{-1}\Sigma_t\Sigma_t\left(A_t^{-1}\right)'. \end{aligned}$$

Defining the vector β containing all coefficients:

$$\underbrace{\beta}_{k\times 1} \equiv vec\left(\begin{bmatrix} c & B_1 & \dots & B_p \end{bmatrix}'\right),$$

the stacked VAR takes the form:

$$y_t = X_t'\beta + u_t, \qquad u_t \overset{i}{\sim} \mathcal{MN}\left(\mathbf{0}, \underbrace{\Omega_t}_{n\times n}\right).$$

Conditional on the history of reduced form volatilities $(\Omega_t)_{t=1}^T$, the Gibbs Sampler step to draw the constant coefficients starts from the following transformation of the model:

$$y_t' = x_t' \cdot \begin{bmatrix} c_t & B_1 & \dots & B_p \end{bmatrix}' + u_t',$$

that allows the following stacked form:

$$\begin{bmatrix} y_1' \\ y_2' \\ \vdots \\ y_T' \end{bmatrix} = \begin{bmatrix} x_1' \\ x_2' \\ \vdots \\ x_T' \end{bmatrix} \cdot \begin{bmatrix} c_t & B_1 & \dots & B_p \end{bmatrix}' + \begin{bmatrix} u_1' \\ u_2' \\ \vdots \\ u_T' \end{bmatrix},$$

$$\underbrace{y}_{T\times n} = \underbrace{x}_{T\times\ (np+1)} \cdot B + u.$$

Exploiting the properties of vectorization and the Kronecker product, it becomes:

$$vec(y) = \underbrace{vec\left(x \cdot B \cdot I_n\right)}_{(I_n \otimes x)\cdot\beta} + vec(u), \qquad X \equiv I_n \otimes x,$$

$$\underbrace{Y}_{nT\times 1} = \underbrace{X}_{nT\times k} \cdot \beta + U, \qquad U \sim \mathcal{MN}\left(\mathbf{0}, \underbrace{V_u}_{n\times n}\right),$$

where

$$V_u \equiv \begin{bmatrix} \Omega_1^{(1,1)} & 0 & \cdots & 0 & \cdots & \cdots & \Omega_1^{(1,n)} & 0 & \cdots & 0 \\ 0 & \Omega_2^{(1,1)} & \ddots & \vdots & \cdots & \cdots & 0 & \Omega_2^{(1,n)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & \cdots & \cdots & \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \Omega_T^{(1,1)} & \cdots & \cdots & 0 & \cdots & 0 & \Omega_T^{(1,n)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \Omega_1^{(n,1)} & 0 & \cdots & 0 & \cdots & \cdots & \Omega_1^{(n,n)} & 0 & \cdots & 0 \\ 0 & \Omega_2^{(n,1)} & \ddots & \vdots & \cdots & \cdots & 0 & \Omega_2^{(n,n)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & \cdots & \cdots & \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \Omega_T^{(n,1)} & \cdots & \cdots & 0 & \cdots & 0 & \Omega_T^{(nn)} \end{bmatrix}.$$

To use an informative prior on β, the approach followed by Gelman et al. (2014) is followed. The strategy incorporates the prior information as observations. Considering a multivariate normal prior with the following moments:

$$\beta \sim \mathcal{MN}\left(\, b,\ V_\beta \,\right),$$

it is possible to augment the model with k observations that express such prior information:

$$\begin{aligned}
\begin{bmatrix} Y \\ b \end{bmatrix} &= \begin{bmatrix} X \\ I_k \end{bmatrix} \beta + \begin{bmatrix} U \\ U_\beta \end{bmatrix}, \\
Y^\diamond &= X^\diamond \beta + U^\diamond, \quad U^\diamond \sim \mathcal{MN}\left(\, \mathbf{0}_{nT+k},\ V^\diamond \,\right), \\
V^\diamond &= \begin{bmatrix} V_u & \mathbf{0}_{nT\times k} \\ \mathbf{0}_{k\times nT} & V_\beta \end{bmatrix}.
\end{aligned}$$

A draw for β can be finally obtained from the following posterior:

$$\begin{aligned}
\beta &\sim \mathcal{MN}\left(\, \widetilde{\beta},\ \left(X^{\diamond\prime} V^{\diamond-1} X^\diamond\right)^{-1} \,\right), \\
\widetilde{\beta} &= \left(X^{\diamond\prime} V^{\diamond-1} X^\diamond\right)^{-1} X^{\diamond\prime} V^{\diamond-1} Y^\diamond.
\end{aligned}$$

2.5.4 Full rank TV-VAR-SV model with autoregressive states

Instead of assuming a multivariate random walk process for the time-varying coefficients, the TV-VAR-SV model can be modified so to have each TV coefficient as a zero-mean AR(1) process[4], that is:

$$\beta_t = \Gamma_\beta\ \beta_{t-1} + \nu_{\beta,t}, \qquad \nu_{\beta,t} \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{Q_\beta}_{k\times k}\right).$$

[4]It would be also possible to specify a non-zero mean AR(1) process.

where Γ_β is a diagonal matrix:

$$\underbrace{\Gamma_\beta}_{k\times k} = \begin{bmatrix} \gamma_{\beta,1} & 0 & \dots & 0 \\ 0 & \gamma_{\beta,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \gamma_{\beta,k} \end{bmatrix}.$$

This requires some modifications of the drawing step of TV coefficients $(\beta_t)_{t=1}^T$. The Gaussian state space representation in this case is:

$$\begin{aligned} y_t &= X_t'\beta_t + u_t, & u_t &\sim \mathcal{MN}\left(\,\mathbf{0},\ \Omega_t\,\right), \\ \beta_t &= \Gamma_\beta\ \beta_{t-1} + \nu_{\beta,t}, & \nu_{\beta,t} &\overset{iid}{\sim} \mathcal{MN}\left(\,\mathbf{0},\ Q_\beta\,\right). \end{aligned}$$

Borrowing from the general implementation described in Carter and Kohn (1994), given Γ_β, Q_β and the data, we can apply the following FFBS procedure. We can initialize the Kalman filter borrowing from the prior of β_0, so setting:

$$\beta_{0|0} = \bar{\beta}, \qquad P_{\beta,0|0} = \bar{P}_\beta.$$

Afterwards, given $\left(\beta_{t-1|t-1}, P_{\beta,t-1|t-1}\right)$, we can compute an entire set of updating and prediction steps $\left(\beta_{t|t-1},\ \beta_{t|t},\ P_{\beta,t|t},\ P_{\beta,t|t-1}\right)_{t=1}^T$ by iterating the following:

$$\begin{aligned} P_{\beta,t|t-1} &= \Gamma_\beta P_{\beta,t-1|t-1}\Gamma_\beta + Q_\beta, \\ K_{\beta,t} &= P_{\beta,t|t-1}X_t\left(X_t'P_{\beta,t|t-1}X_t + \Omega_t\right)^{-1}, \\ \beta_{t|t} &= \Gamma_\beta\beta_{t-1|t-1} + K_{\beta,t}\left(y_t - X_t'\Gamma_\beta\beta_{t-1|t-1}\right), \\ P_{\beta,t|t} &= P_{\beta,t|t-1} - K_{\beta,t}X_t'P_{\beta,t|t-1}. \end{aligned}$$

Using the updating and prediction steps, we can finally sample backward starting from β_T, which is distributed as a $\mathcal{MN}\left(\,\beta_{T|T},\ P_{\beta,T|T}\,\right)$, and then sampling recursively, for all $t\ \in \{T-1, T-2, \dots, 2, 1\}$, each β_t from $\mathcal{MN}\left(\,\beta_{t|t+1},\ P_{\beta,t|t+1}\,\right)$ where:

$$\begin{aligned} \beta_{t|t+1} &= \beta_{t|t} + P_{\beta,t|t}\Gamma_\beta P_{\beta,t+1|t}^{-1}\left(\beta_{t+1} - \Gamma_\beta\beta_{t|t}\right), \\ P_{\beta,t|t+1} &= P_{\beta,t|t} - P_{\beta,t|t}\Gamma_\beta P_{\beta,t+1|t}^{-1}\Gamma_\beta P_{\beta,t|t}. \end{aligned}$$

The k autoregressive coefficients contained in Γ_β can be estimated by augmenting the Gibbs Sampler with a Bayesian regression step. Indeed, let's simply rewrite the state equation as:

$$\beta_t = \underbrace{\begin{bmatrix} \beta_{t-1,1} & 0 & \dots & 0 \\ 0 & \beta_{t-1,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \beta_{t-1,k} \end{bmatrix}}_{\mathcal{B}_t} \underbrace{\begin{bmatrix} \gamma_{\beta,1} \\ \gamma_{\beta,2} \\ \vdots \\ \gamma_{\beta,k} \end{bmatrix}}_{\gamma_\beta} + \nu_{\beta,t},$$

that can be stacked, becoming

$$\underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_T \end{bmatrix}}_{\beta} = \underbrace{\begin{bmatrix} \mathcal{B}_1 \\ \mathcal{B}_2 \\ \vdots \\ \mathcal{B}_T \end{bmatrix}}_{\mathcal{B}} \gamma_\beta + \underbrace{\begin{bmatrix} \nu_{\beta,1} \\ \nu_{\beta,2} \\ \vdots \\ \nu_{\beta,T} \end{bmatrix}}_{\nu_\beta}, \qquad \nu_\beta \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{I_T \otimes Q_\beta}_{kT \times kT} \right).$$

Within the sampler, given a set of coefficients $(\beta_t)_{t=1}^T$, and variance matrix Q_β, and assuming a conjugate prior for γ_β :

$$\gamma_\beta \sim \mathcal{MN}\left(\bar{\gamma}_\beta\ \mathbf{1}_{k\times 1},\ \bar{\sigma}^2_{\gamma,\beta}\ I_k \right),$$

we can draw from the following posterior:

$$\gamma_\beta \sim \mathcal{MN}\left(\widetilde{\gamma}_\beta\ ,\ \Psi_{\gamma,\beta} \right),$$

where

$$\Psi_{\gamma,\beta} = \left[\mathcal{B}' \left(I_T \otimes Q_\beta \right)^{-1} \mathcal{B} + \bar{\sigma}^{-2}_{\gamma,\beta}\ I_k \right]^{-1},$$

$$\widetilde{\gamma}_\beta = \Psi_{\gamma,\beta} \left[\mathcal{B}' \left(I_T \otimes Q_\beta \right)^{-1} \beta + \bar{\sigma}^{-2}_{\gamma,\beta}\ \bar{\gamma}_\beta\ \mathbf{1}_{k\times 1} \right].$$

3 The Algorithm

The estimation of all models showed in this paper is implemented in MATLAB®, via a package of several functions[5]. Almost all the functions have been programmed from scratch. The code's efficiency has been thoroughly worked out using the Matlab time profiler. Loops have been used only when unavoidable (for instance because of chain dependence), and this was accomplished through the extensive usage of the multidimensional matrix multiplication package *mtimesx*, compiled in C by *James Tursa* (shared via the Matlab Central File Exchange). Indeed, when estimating time-varying models, avoiding repeated loops over the time dimension $\{1, \ldots, T\}$ is extremely beneficial in terms of efficiency. Most matrix reshaping, extensively needed throughout the algorithm, has been performed by exploiting the linear indexing of sparse and full matrices written in the Matlab language. A general routine for the Forward Filtering Backward Sampling implementation has been programmed, and it allows for VAR(1) states and heteroskedastic innovations. The algorithm of drawing from the Singular Inverse Wishart distribution is programmed to reduce the computational burden due to large singular matrix inversion and eigenvalue sorting (see Sætrom, 2010 and Bodnar and Okhrin, 2008). Minor numerical issues, like slightly imprecise matrix semi-definitive positiveness or imperfect symmetry (with magnitude 10^{-6} or smaller), arising because of large dimensional matrices calculations and floating point numerical system, have been tackled. For each set of draws from the Markov Chain, a 30% of initial burn-in draws are automatically discarded by the algorithms.

4 Concluding remarks

This paper reviews several alternative models that introduce time variation in VAR coefficients and volatilities of residuals. In particular, a rank reduction strategy proposed by de Wind and Gambetti (2014) is analyzed in depth and finely reshaped in some steps. Moreover, the Gibbs Sampler algorithms presented have been extensively tested and always converged well enough to stationary posterior distributions. The full rank estimation is more costly in terms of computation time than its reduced rank counterpart, especially when models get large. However, using macro-economic data at low frequencies, VAR

[5]The packages are available upon request.

coefficients typically display a smaller degree of time variation and the reduced rank inference is qualitatively and quantitavely very close to the full rank counterpart.

A possible path is to produce hybrid strategies when modeling time variation of the VAR ingredients. Indeed, a model with reduced rank time-varying coefficients but full rank stochastic volatilities might be very useful given the typically large discrepancy between the number of objects involved: stochastic volatilities amount to a much smaller number of objects that grow less quickly with the number of variables with respect to VAR coefficients, making a full rank strategy relatively less costly for SV than for TVC.

These types of models and algorithms, along with some novel ones, will be studied in the following chapters, also by means of some empirical applications.

There is an important number of questions and issues that future research on this field should focus on. Given the good behavior showed by the algorithm, further research should assess if MCMC convergence for these models is reached with a smaller number of draws and, more importantly, whether merging several smaller chains produced in parallel computing may alter the inference in terms of posterior distribution shape and moments[6]. In case of partially or completely affirmative answer, there are huge potential gains in terms of computational costs and feasibility of a larger number of experiments. In-sample evaluation strategies as computing the marginal data densities should also be attempted, even though the substantial non-linearity of these models may enlarge the role of numerical approximation. As a general remark, given the crucial role of the prior on innovation variances, a prior sensitivity analysis should be conducted.

[6]This is the approach suggested by the work of Gelman and Rubin (1992).

Bibliography

Ataei, A. (2014). Improved qrginv algorithm for computing moore-penrose inverse matrices. *ISRN Applied Mathematics*.

Bodnar, T. and Okhrin, Y. (2008). Properties of the singular, inverse and generalized inverse partitioned wishart distributions. *Journal of Multivariate Analysis*, 99(10):2389–2405.

Carter, C. K. and Kohn, R. (1994). On gibbs sampling for state space models. *Biometrika*, 81(3):541–553.

Cogley, T. and Sargent, T. J. (2001). Evolving post-world war ii us inflation dynamics. *NBER macroeconomics annual*, 16:331–373.

Cogley, T. and Sargent, T. J. (2005). Drifts and volatilities: monetary policies and outcomes in the post wwii us. *Review of Economic dynamics*, 8(2):262–302.

de Wind, J. and Gambetti, L. (2014). Reduced-rank time-varying vector autoregressions. *CPB Netherlands Bureau for Economic Policy Analysis, No. 270.*

Del Negro, M. and Primiceri, G. E. (2015). Time varying structural vector autoregressions and monetary policy: a corrigendum. *The review of economic studies*, 82(4):1342–1345.

Dıaz-Garcıa, J. A., Jáimez, R. G., and Mardia, K. V. (1997). Wishart and pseudo-wishart distributions and some applications to shape theory. *Journal of Multivariate Analysis*, 63(1):73–87.

Fuller, W. A. (2009). *Introduction to statistical time series*, volume 428. John Wiley & Sons.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2014). *Bayesian data analysis*, volume 2. CRC press Boca Raton, FL.

Gelman, A. and Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, pages 457–472.

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with arch models. *The review of economic studies*, 65(3):361–393.

Omori, Y., Chib, S., Shephard, N., and Nakajima, J. (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics*, 140(2).

Primiceri, G. E. (2005). Time Varying Structural Vector Autoregressions. *The Review of*

Economic Studies, 72(3):821–852.

Sætrom, J. (2010). Reduction of dimensionality in spatiotemporal models. *Doctoral Thesis, Department of Mathematical Sciences, Norwegian University of Science and Technology.*

Sims, C. A. (2001). [evolving post-world war ii us inflation dynamics]: Comment. *NBER macroeconomics annual*, 16:373–379.

Chapter 2

Labor Market and Financial Shocks: a Time-Varying Analysis

co-authored with Valerio Nispi Landi

1 Introduction

The Great Recession has been particularly detrimental for the US labor market, in terms of both magnitude and persistence of the effects (Figure 1). The unemployment rate reached a peak of about 10% at the end of 2009 and it fully recovered only in 2016; this employment collapse was mainly driven by an impressive layoffs growth in the private sector; in addition, hours worked per employee have heavily fallen down throughout the crisis and in 2017 they are still not back to their pre-crisis level. Labor force participation was also strongly affected ans has accelerated the downward trend started in the early 2000s. Finally, not surprisingly, the vacancy rate experienced a significant drop.

Given that the last recession was generated by a large turmoil in the financial sector (Figure 2), studying the linkages between financial and labor markets seems crucial to gain a better comprehension of business cycle fluctuations. Moreover, the last crisis has fostered a large debate about effectiveness of macroprudential policies aiming to maintain financial stability and to dampen financial shock impact: therefore, analyzing the effects of a financial tightening on the real economy, and especially on the labor market, can provide some useful guidance to policy makers.

In the empirical economic literature there are several works analyzing the macroeconomic effects of a financial shock, defined as a tightening of financial conditions (Gilchrist et al., 2009, and Gilchrist and Zakrajšek, 2012 are eminent examples). These works typically use a VAR model with constant standard deviations (CVAR henceforth) to identify the impact of a financial shock, using sign or short-run restrictions. However, the assumption of constant shock's standard deviation seems quite restrictive when the object of interest is a financial shock: financial time series are typically asymmetric and exhibit changes in volatility over time.[1] Figure 3 shows the time series of the Chicago Fed National Financial Condition Index (NFCI henceforth) and the Gilchrist and Zakrajsek Spread (GZ from now on). These indicators measure the tightness of financial conditions in the US, with higher values denoting periods of tighter conditions: both measures feature small fluctuations on average (associated with periods of financial calm) and few big and volatile peaks (associated with periods of financial distress); a VAR with stochastic volatility is able to distinguish between periods of high and low volatility: as a result the estimated standard deviation of financial shocks is likely to be high in periods of financial turbulence, as the data seem to suggest. Furthermore, the bias resulting from a CVAR estimation is

[1] See Kim et al. (1998) for an overview of stochastic volatility models.

magnified if the sample period includes the recent financial crisis: the estimated impact of a shock in a CVAR tries to summarize the effects of such a shock over time. Given that the Great Recession faced a huge drop in the real activity due to massive distress in financial markets, a CVAR may overestimate the effect of a financial innovation across the sample. In Section 4 we argue that this overestimation is particularly strong for the unemployment rate.

Nevertheless, time-varying volatility may not be enough to correctly capture the effect of a financial shock. Indeed, not only the size of a financial shock can be time-varying, but also its transmission mechanism. As shown by Gaiotti (2013), Silvestrini and Zaghini (2015) and Prieto et al. (2016), a change in financial conditions has a higher impact on the real economy during periods of financial distress: this can be explained theoretically by the presence of financial constraints that can be binding during recessions, as pointed out by Mendoza (2010), Bianchi (2011) and Guerrieri and Iacoviello (2017). Moreover, several recent papers show that labor market time series feature a significant degree of skewness.[2]

Motivated by all these considerations, we assess the interactions between labor market and financial shocks in a time-varying framework. In particular, the following research questions are tackled:

- What is the effect of a financial tightening on the labor market?
- How relevant are financial shocks in explaining historical labor market fluctuations?
- Is the response of labor market time-varying?
- Is the size of financial shocks changing over time?
- What are the main transmission mechanisms and how are they captured by VAR models?

To the best of our knowledge, there are no empirical studies in the literature focusing on the effects of financial shocks in the labor market and this is particularly surprising, given what happened during the Great Recession. Therefore, following the work of Primiceri (2005), we estimate a time-varying VAR with stochastic volatility (TV-VAR-SV henceforth), over the sample period 1973-2016. We include in the model eight US quarterly time series: five labor-market variables (unemployment rate, participation rate, hours worked

[2]See, for instance, McKay and Reis (2008), Ferraro (2017) and Pizzinelli and Zanetti (2017).

per employee, nominal wage and vacancy rate), two macroeconomic variables (inflation rate and real GDP) and the National Financial Condition Index (the NFCI henceforth). The NFCI series and its three components (risk, credit and leverage) are published by the Federal Reserve of Chicago at weekly frequency. This index is a dynamic factor built from a balanced panel of 100 mixed-frequency indicators of financial activity. The financial shock that we aim to identify is a tightening of financial conditions, modeled as an unanticipated exogenous increase in the NFCI: as such, this is a credit supply shock, deeply analyzed by theoretical studies.[3] The shock is identified through a short-run restriction: the labor market and the macroeconomy cannot simultaneously react to innovations in the NFCI, a standard assumption in the structural VAR literature. This restriction is even more reasonable in our framework, since the labor market is by its own nature sluggish to respond to economic shocks.

Our contribution to the literature is threefold. First we show that a tightening of financial conditions depresses the labor market and the GDP while the effect on inflation is positive and borderline significant.[4] Second, we find that financial shocks have hit the unemployment rate asymmetrically over the last three decades: while negative financial shocks have been responsible for the high unemployment rates in the early 1980s and during the Great Recession, our model does not find relevant contribution of financial shocks throughout expansion periods. We show that a CVAR is not able to capture this asymmetry. Third, we observe strong evidence of time variation in the financial shock volatility, with peaks during economic downturns, consistently with the works of Justiniano and Primiceri (2008), Stock and Watson (2012) and Prieto et al. (2016): accordingly, we find larger negative than positive shocks. Since the systematic transmission mechanism does not display time variation, we can conclude that the asymmetric response of the unemployment rate is entirely due to financial shock's drifting volatility.

The remainder of the paper is organized as follows: in Section 2 we provide a brief review of theoretical and empirical literatures on the transmission of financial shocks to real activity; Section 3 presents the time series used in the analysis; Section 4 introduces the estimation methodology; Section 5 describes the results; in Section 6 some robustness checks are reported; Section 7 concludes.

[3]See, for instance, the DSGE model by Jermann and Quadrini (2012).

[4]Notice that the DSGE literature is not unanimous about the effects of financial shocks on the inflation rate.

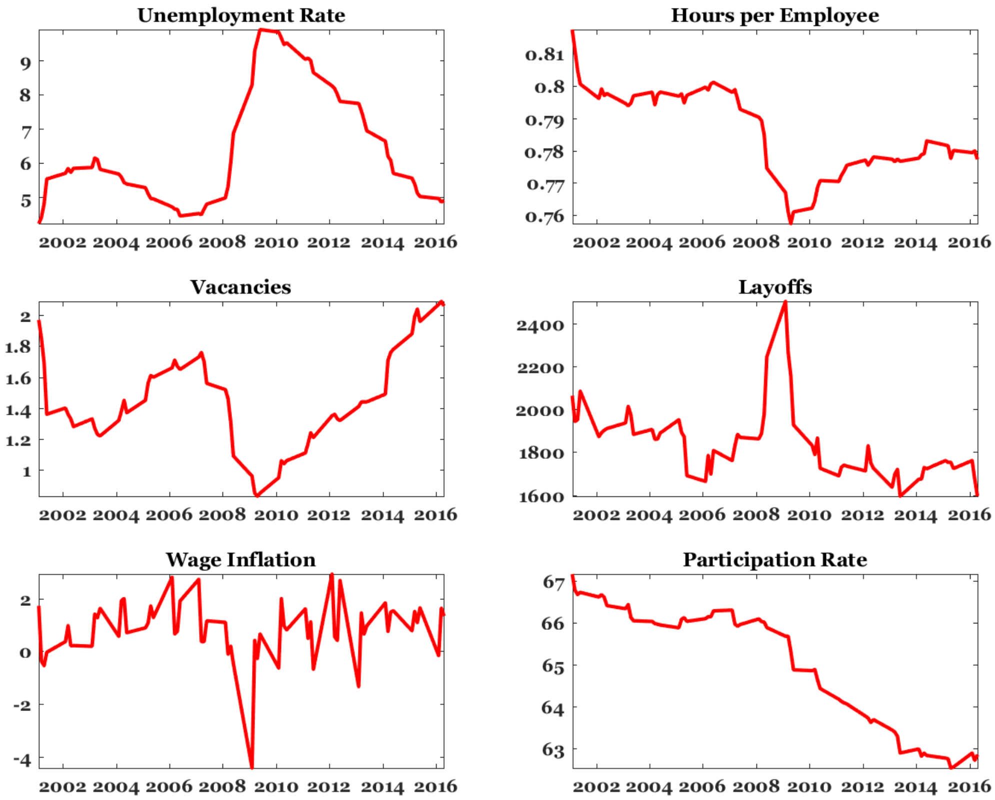


Figure 1: US labor market during the Great Recession. Unemployment rate, participation rate and wage inflation are expressed in percentage points. Hours per employee is the ratio between hours in nonfarm business sector (an index with base year 2010), and the number of employees in that sector. Layoffs in nonfarm business sector are expressed in thousands of workers. Vacancies are measured by the Help-Wanted Index developed by Barnichon (2010). All these variables but the vacancy index come from the BLS database.

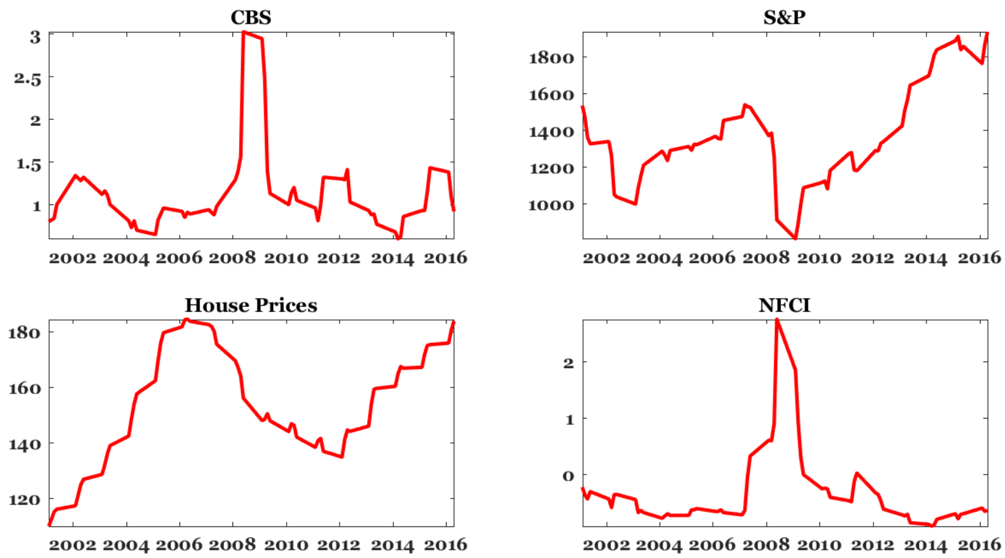


Figure 2: US financial markets during the Great Recession. CBS is the Moody's BAA-AAA corporate bond spread expressed in percentage points. S&P is the Standard & Poor stock market index. House prices are measured by the Standard & Poor's Case–Shiller Home Price Index (base year 2000). NFCI is the National Financial Condition Index developed by the Chicago Fed. Source: St. Louis Fed.

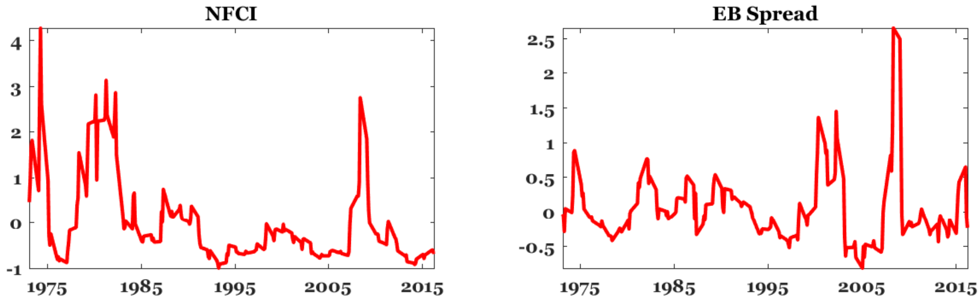


Figure 3: EB spread is the excess bond premium component of the GZ spread built by Gilchrist and Zakrajšek (2012).

2 Related Literature

2.1 Theoretical Literature

The macroeconomic theoretical literature has deeply analyzed the transmission mechanism linking real activity with the financial sector. The seminal paper by Bernanke and

Gertler (1989) introduces the concept of financial accelerator: in presence of asymmetric information, lenders are more willing to provide funds if borrowers are endowed with high levels of net worth; since the latter is likely to be procyclical, following a recessionary shock firms find it harder to obtain outside funds; as a result, investment falls, reinforcing the initial drop in net worth, eventually magnifying the negative consequences on the real economy. Bernanke and Gertler (1995) argue that an unanticipated monetary policy tightening reduces firms cash flows and so their net worth, inducing smaller investments, reducing further the net worth and exacerbating the shock's negative effects. Kiyotaki and Moore (1997) build a model following the same intuition: if borrowers secure loans with capital and they face binding collateral constraints, any negative shock to the collateral's value force them to deleverage. As a result, they reduce investment in new capital, amplifying the deleveraging process and depressing future investment as well.

The aforementioned studies have formed the basis for business cycle models that incorporate the financial accelerator mechanism in a general equilibrium framework.[5] In particular, Jermann and Quadrini (2012) develop a model in which the financial sector is not only a source of amplification but also a driver of business cycle fluctuations, as witnessed by the Great Recession. In this model, a tightening of financial conditions reduces firms' borrowing ability and forces them to cut hours of work and production. Also in Gertler and Karadi (2011), the main driver of fluctuations originates in the financial sector: a deterioration in the quality of financial assets reduces the net worth of banks. In turn, banks raise the lending cost to remain profitable and avoid deposit withdrawals. The shock that we aim to identify in our model resembles those in Jermann and Quadrini (2012) and Gertler and Karadi (2011): it is a credit supply shock that makes it harder for firms to borrow funds; using the NFCI goes precisely in this direction, given that this indicator captures the tightness of financial conditions.

A recent strand of the macro-financial literature has departed from the assumption - made for computational reasons - that financial constraints are always binding: scholars have been starting to set up models characterized by occasionally binding collateral constraints, in order to study the (potentially) non-linear effects of economic shocks.[6] During recessions, such constraints bind and firms deleverage; on the other hand, in expansions, a loosening of financial conditions does not have a direct impact on collateral constraints,

[5] See also Bernanke et al. (1999) and Christiano et al. (2003).

[6] For example, occasionally binding collateral constraints characterize the model of Mendoza (2010) and Guerrieri and Iacoviello (2017).

which continue to be slack.

A common denominator of these papers is the abstraction from unemployment issues: the labor input is modeled as an intensive margin and so there is no room for unemployment. The Great Recession has provided motivation to fill this gap: frictional financial markets have been introduced in models featuring equilibrium unemployment à la Mortensen and Pissarides (1994): one contribution of our work is to qualitatively capture the dynamics highlighted in these models. A not comprehensive list includes Christiano et al. (2011), Petrosky-Nadeau and Wasmer (2013), Petrosky-Nadeau (2013), Iliopulos et al. (2014) and Zanetti (2017). In these works, financial and labor-matching frictions coexist: in particular, firms need external funding to finance the cost of posting vacancies and attracting unemployed workers. During crises, firms have limited access to external credit and thus reduce the number of posted vacancies, so that financial frictions magnify the role played by search and matching frictions in creating unemployment. Notably, Petrosky-Nadeau (2013) examines the impact of a credit shock on the labor market. In this model, a worsening of financial conditions induces firms to reduce the number of vacancies: the probability to find a job for an unemployed falls and the unemployment rate rises; as a consequence, labor market tightness (defined as the vacancy-unemployment ratio) collapses, making easier for firms to fill a vacancy. The obvious by-product of such a negative shock is a drop of aggregate production. On the other hand, Monacelli et al. (2011) propose a different mechanism. The authors argue that external debt allows firms to get a better position when bargaining wage with workers. In such a case, a negative financial shock forces firms to reduce borrowing and places them in a less favorable bargaining position with workers: as a result, firms will create less jobs.

2.2 Empirical Literature

There is a growing empirical literature analyzing the effects of financial shocks. One seminal paper of this literature is Gilchrist and Zakrajšek (2012): they construct the GZ credit spread index, an indicator with a significant predictive power for economic activity. The GZ spread is decomposed in two components: the expected default and the excess bond premium; the latter is used as financial variable in a macroeconomic VAR: an innovation in the excess bond premium leads to economically and statistically significant declines in the real activity. Fornari and Stracca (2012) estimate a panel VAR for a large set of advanced economies, in order to evaluate the effects of a credit supply shock, identified with sign restrictions. Their findings show that financial shocks exert a relevant influence on the real economy, both in terms of shock's size and variance explained. These results are confirmed by Caldara et al. (2016): they distinguish between financial and uncertainty shocks, finding that the former ones significantly affect economic conditions and are an important source of business cycle fluctuations; moreover, they amplify the negative effects of adverse uncertainty shocks. One of our goals is contributing to this literature, by augmenting the set of endogenous variables typically included in VAR with some labor-market variables.

The previous works all consider linear frameworks, in which estimated parameters are constant over time. Our work goes one step ahead by allowing time variation in parameters and shock standard deviations. Indeed, some papers point out that including time variation in a financial shock analysis yields some important insights. In particular, many studies find that standard deviation of financial shocks is changing over time and it is relatively higher in periods of financial distress. Using Italian firm-level, Gaiotti (2013) shows that quantitative credit constraints have a larger effect during economic slack. Silvestrini and Zaghini (2015) study the linkages between financial shocks and the macroeconomy in a time-varying framework. They estimate a time-varying VAR with euro-area variables, finding evidence of time variation in the transmission mechanism of a shock to the Composite Indicator of Systemic Stress, identified with short-run restrictions: notably, this financial shock has a more significant adverse impact during financial crises. Prieto et al. (2016) estimate a time-varying VAR for the US economy with three financial market variables (house prices, stock prices and credit spread). They find that during financial crisis, contribution of financial shocks in explaining fluctuations in GDP more than doubles, and this is mainly explained by a higher shock volatility. In a reduced-form model, Alessandri et al. (2017) find that changes in the GZ Excess Bond Premium have stronger predictive power for crisis than for expansions.

3 Data

The estimation is performed over the sample period 1973Q1-2016Q3, using eight US time series; the observations going from 1973Q1 to 1980Q4 are used as training sample to calibrate prior distributions. The sample starting point is dictated by availability of the NFCI.

In order to capture financial shocks' effects both on the extensive and intensive margin, the vector of endogenous variables comprises the unemployment rate, the log-first difference of the participation rate (defined as labor force over working age population) and the log-first difference of hours of work per capita in the non-farm business sector; in addition, we include the log-first difference of nominal wage, the vacancy index constructed by Barnichon (2010), GDP deflator and real GDP. Our financial measure is the NFCI.[7]

It deserves to spend some words about the NFCI, given the crucial role played in our analysis. The NFCI, published at weekly frequency by the Federal Reserve Bank of Chicago, is a weighted average of a set of financial indicators: the weight of each indicator is estimated by means of the principal component technique. The index captures measures of risk, liquidity and leverage. The risk component summarizes the premium yielded by risky assets and their volatility; the liquidity component provides an index of willingness to borrow and lend at prevailing prices; the leverage component assesses the economy-wide level of financial debt relative to equity. The index and its components are rescaled to have a zero mean and a unitary standard deviation; positive (negative) values denote financial conditions that are tighter (looser) than average. The main advantage of this indicator comes from an overall description of US financial conditions: this is particularly important due to the strong interconnectivity of US financial markets, in which shocks tend to have an impact on aggregate financial conditions rather than on one specific segment.[8]

In Figure 4 we plot the seven non-financial series we use in estimation, excluding training sample observations. The unemployment rate is above 5% in several periods of our sample:[9] it is particularly high during the 1980s and throughout the Great Recession. During the recent financial crisis we also observe unusual negative values for price and wage inflation along with a large drop in hours of work, vacancies and GDP. Figure 5 plots the NFCI and its subcomponents. The index displays positive values at the beginning of the sample (in correspondence with the less-developed-countries debt crisis occurred in the early 1980s), and in the late 1980s, in addition to the big spike in 2009; the years between the 1990's and the boom preceding the Great Recession were characterized by

[7]We are not the first ones to use the NFCI in a structural VAR: Fink and Schüler (2015) identify a financial shock as an innovation in the NFCI; Metiu et al. (2015) include the NFCI in a VAR in order to capture the transmission mechanism of a financial shock defined as an innovation to the GZ spread.

[8]See Brave and Butters (2011) and Brave and Butters (2012) for more details.

[9]According to FED estimates, the long-run normal rate of unemployment ranges from 4.5 to 5.0 percent.

relatively looser financial conditions. The three NFCI components exhibit similar patterns: interestingly, the leverage component is much more volatile, while the credit index is more persistent.

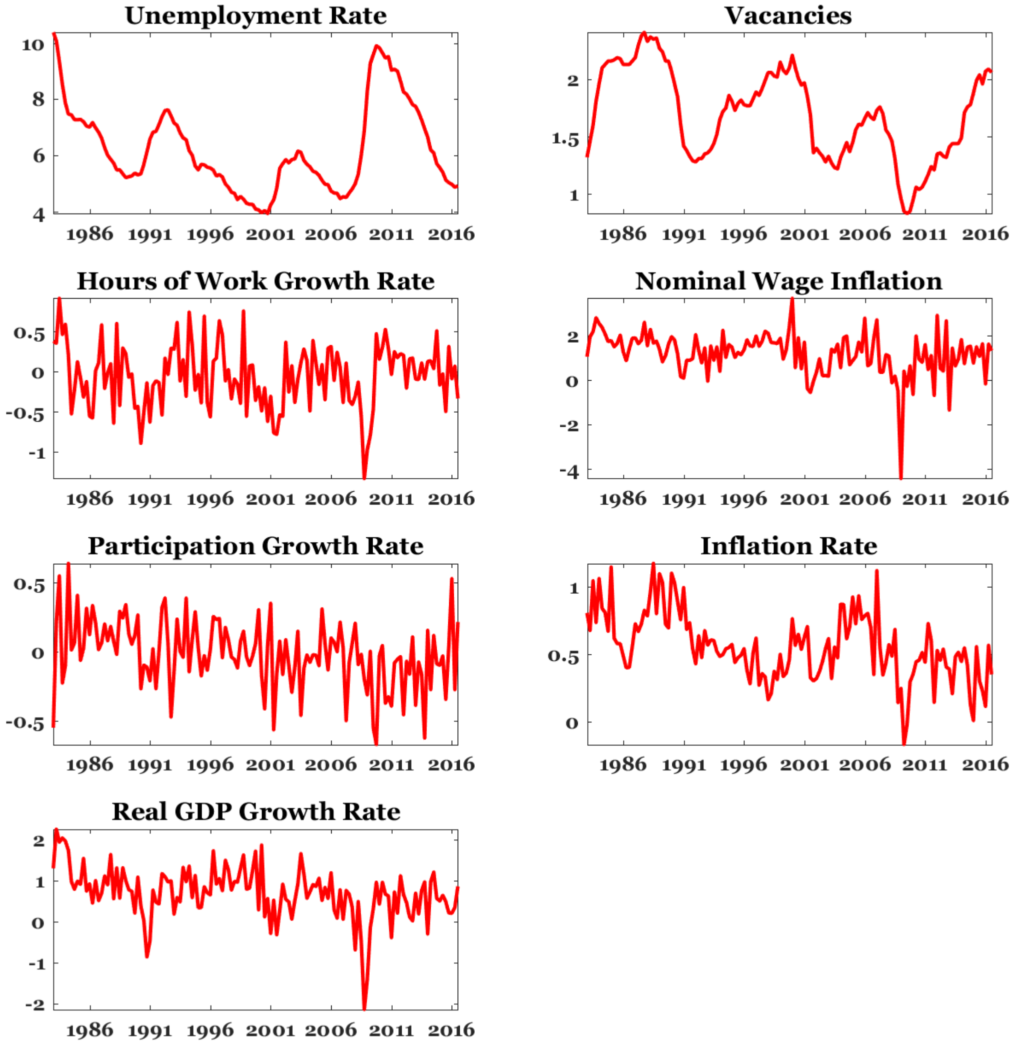


Figure 4: Labor and macroeconomic variables. Variables are expressed in percentage points except vacancies.

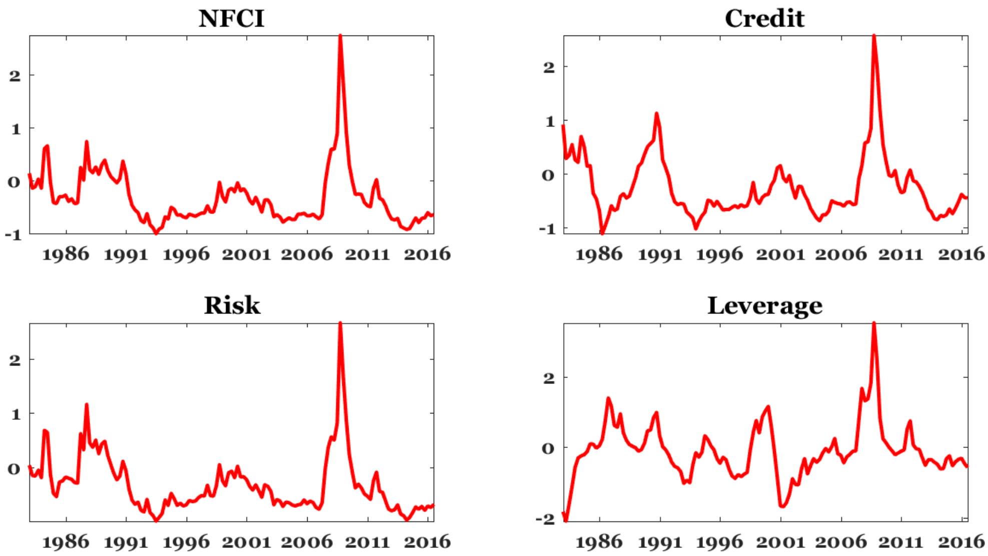


Figure 5: NFCI and its subcomponents.

4 Econometric Methodology

4.1 TV-VAR-SV and rank reduction

The n variables of interest are modeled through a time-varying coefficient VAR with stochastic volatility, whose reduced form follows the seminal work of Primiceri (2005):

$$\underbrace{y_t}_{n\times 1} = c_t + \sum_{\ell=1}^{p} \underbrace{B_{\ell,t}}_{n\times n} y_{t-\ell} + u_t, \qquad u_t \overset{i}{\sim} \mathcal{N}\left(\mathbf{0},\ \underbrace{\Omega_t}_{n\times n} \right).$$

Following the literature, for estimation purposes the model is transformed as:

$$y_t = X_t' \beta_t + A_t^{-1} \Sigma_t \varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} \mathcal{N}\left(\mathbf{0},\ I_n \right),$$

$$\underbrace{X_t'}_{n\times k} \equiv I_n \otimes \begin{bmatrix} 1 & y_{t-1}' & y_{t-2}' & \dots & y_{t-p}' \end{bmatrix}$$

where the time-varying coefficients are stacked in a set of k-dimensional[10] vectors $(\beta_t)_{t=1}^T$ and the $n \times n$ reduced form covariances $(\Omega_t)_{t=1}^T$ face a triangular reduction, producing the lower triangular matrices $(A_t)_{t=1}^T$ and diagonal matrices $(\Sigma_t)_{t=1}^T$:

$$\underbrace{\beta_t}_{k\times 1} \equiv vec\left(\begin{bmatrix} c_t & B_{1,t} & \dots & B_{p,t} \end{bmatrix}'\right)$$

$$\Omega_t = A_t^{-1}\Sigma_t\Sigma_t\left(A_t^{-1}\right)'.$$

Each group of time-varying elements follows a multivariate random walk with Gaussian innovations, whose covariances are parametrized by matrices $\{Q_\alpha, Q_\beta, Q_\sigma\}$. Indeed, defining the r-dimensional vectors[11] $(\alpha_t)_{t=1}^T$ and the n-dimensional vectors $(\sigma_t)_{t=1}^T$ as the sets of vectors containing, respectively, the off-diagonal elements in $(A_t)_{t=1}^T$ and the diagonal elements in $(\Sigma_t)_{t=1}^T$, we can write the assumed law of motions:

$$\begin{aligned} \beta_t &= \beta_{t-1} + \nu_{\beta,t}, & \nu_{\beta,t} &\overset{iid}{\sim} \mathcal{N}\left(\ \mathbf{0},\ Q_\beta\ \right) \\ \alpha_t &= \alpha_{t-1} + \nu_{\alpha,t}, & \nu_{\alpha,t} &\overset{iid}{\sim} \mathcal{N}\left(\ \mathbf{0},\ Q_\alpha\ \right) \\ \log\sigma_t &= \log\sigma_{t-1} + \nu_{\sigma,t}, & \nu_{\sigma,t} &\overset{iid}{\sim} \mathcal{N}\left(\ \mathbf{0},\ Q_\sigma\ \right). \end{aligned}$$

As stressed in Primiceri (2005), the random walk assumption is not harmful for estimation purposes, especially since the process is operating for a finite number of periods. Moreover, the random walk specification is parsimonious and does not restrict much the elements' dynamics.

However, a rank reduction strategy proposed by de Wind and Gambetti (2014, WG henceforth) is implemented for the time-varying coefficients. The reason is twofold: first, the rank reduction methodology makes possible to reduce the computational burden of non-small TV-VAR-SV models (say, for larger models than the small one estimated in Primiceri, 2005); second, in line with the findings of Cogley and Sargent (2005) and WG, these models typically do not show large time variation of VAR coefficients when estimated with macro-financial variables. Therefore, a smaller number of components are able to capture coefficients time variation, especially considering that the number of VAR coefficients increases very fast with the number of variables and lags, and there is no

[10] $k = n\,(np + 1)$.
[11] $r = n\,(n-1)\,/2$.

reason to conceive such a large number of economic forces able to alter the transmission mechanism at every period.

In particular, Q_β is assumed to be a reduced-rank matrix with rank $q_\beta < k$ and it is eigendecomposed in the following way:

$$Q_\beta = \underbrace{\Lambda_\beta}_{k\times q_\beta} \Lambda'_\beta = \underbrace{V_\beta}_{k\times q_\beta} \underbrace{D_\beta}_{q_\beta\times q_\beta} V'_\beta, \qquad \Lambda_\beta \equiv V_\beta D_\beta^{1/2}$$

where D_β is a diagonal matrix containing the non-zero eigenvalues of Q_β and V_β is the matrix where the associated eigenvectors are stacked in columns. Following the procedure of WG, the time-varying coefficient law of motion can be transformed (projecting onto the column space of Λ_β) in order to reduce the number of time-varying components to q_β. WG highlight that this assumption is reasonable with macroeconomic variables, especially when k gets large; in these circumstances the time variation can be spanned using a smaller space without significant changes in the inference. To sum up, the time-varying coefficients are eventually linear combinations of time-varying components with the following associated laws of motion and time-invariant residual:[12]

$$\underbrace{\beta_t}_{k\times 1} = \Lambda_\beta \widetilde{\beta}_t + M_\beta \beta_0, \qquad \underbrace{\widetilde{\beta}_t}_{q_\beta\times 1} = \widetilde{\beta}_{t-1} + \widetilde{\nu}_{\beta,t}, \qquad \underbrace{\widetilde{\nu}_{\beta,t}}_{q_\beta\times 1} \overset{iid}{\sim} \mathcal{N}\left(\mathbf{0},\ I_{q_\beta} \right).$$

Inference on the covariance matrices of states' innovations and on the unobservable states (time-varying coefficients, off-diagonal elements and stochastic volatilities) is produced by means of a Gibbs Sampler that draws from conditional posteriors. The Gibbs Sampler steps are listed below.

Gibbs Sampler Steps[13]

1. Initialize the Gibbs Sampler at some $(\beta_t^0,\ \alpha_t^0,\ \sigma_t^0)_{t=0}^T,\ Q_\beta^0, Q_\alpha^0, Q_\sigma^0$, set $i = 1$.
2. Draw a history of time-varying coefficients $(\beta_t^i)_{t=1}^T$:
   (a) Draw a history of time-varying components $\left(\widetilde{\beta}_t^i\right)_{t=1}^T$;

[12] The time invariant residual is obtained through the matrix $M_\beta \equiv I_m - \Lambda_\beta \left(\Lambda'_\beta \Lambda_\beta\right)^{-1} \Lambda'_\beta$. As shown in Corsello (2018), the residual only depends on the initial condition β_0.

[13] See Corsello (2018) for analytical details on each step.

(b) Draw a vector of time invariant component $M_\beta \beta_0$ of $(\beta_t^i)_{t=1}^T$.

3. Draw a reduced-rank covariance matrix Q_β^i.

4. Draw a history of off-diagonal elements $(\alpha_t^i)_{t=1}^T$.

5. Draw a full rank covariance matrix Q_α^i.

6. Draw a history of volatilities $(\sigma_t^i)_{t=1}^T$.

7. Draw a full-rank covariance matrix Q_σ^i, set $i = i + 1$ and restart at [2].

Draws from the posterior of unobservable states are obtained by means of the Carter and Kohn (1994) algorithm. Draws from the Singular Inverse Wishart are implemented differently from what proposed by WG, following the approach of Corsello (2018). The stochastic volatilities draws are implemented as in Primiceri (2005), but respecting the correct order described in Del Negro and Primiceri (2015). The $\log \chi_1^2$ distribution is approximated using the mixture of normal components proposed by Omori et al. (2007), which improves upon the previous standard set by Kim et al. (1998).

The prior distribution is set following Primiceri (2005) for $\{Q_\alpha, Q_\sigma\}$ and for the initial conditions α_0 and σ_0, while WG is used as benchmark for calibrating the prior distributions of Q_β and β_0. In particular, the training sample used to calibrate prior distributions runs from 1973Q1 to 1980Q4. Dealing with quarterly data, we set $p = 2$ lags (as in Primiceri, 2005), so having a total of $k = 78$ coefficients in the VAR. The reduced rank of the time-varying coefficients is chosen to be $q_\beta = 40$, which constitutes a good compromise between dimensionality and span time variation.[14]

4.2 Identification Strategy

Our identification scheme relies on a short-run restriction: we assume that both the labor market (unemployment, vacancies, participation, hours and wage) and the macroeconomy (GDP and inflation) do not contemporaneously respond to innovations in the NFCI equation. Accordingly, we order NFCI last in the endogenous vector and impose a Cholesky structure to the matrix of contemporaneous relations.

[14]Increasing the rank does not produce any variation on time-varying coefficient inference, showing that the chosen dimension is sufficient to span the measured time-varying coefficient variations over time.

This identification strategy is quite standard in the economic literature dealing with structural shocks identification. Christiano et al. (1999) and Bernanke et al. (2005) are pioneering examples for the identification of monetary shocks using a Cholesky approach. Gilchrist and Zakrajšek (2012), Prieto et al. (2016) and Silvestrini and Zaghini (2015) use short-run restrictions to identify a financial shock.

The rationale behind this assumption is that the private sector typically tends to react with a delay to economic shocks, due to real and nominal rigidities, such as, price and wage adjustment costs or habits persistence in consumption. A Cholesky identification scheme seems even more reasonable when the focus is on labor market, which is sluggish by its own nature. In Section 6 we estimate the model using monthly data, making the short-run assumption more realistic, and we show that results are barely affected.

5 Results

5.1 Structural Analysis

5.1.1 Impulse responses and stochastic volatilities

Following Lopez-Salido and Nelson (2010) and Prieto et al. (2016), we group quarterly observations available in the sample into financial crisis and non-financial crisis periods. In particular, we consider five periods of US financial distress: the Less-Developed-Countries Debt Crisis (1982-1984), the stock market crash of 1987, the Savings and Loan Crisis (1988-1991), the stock market crash of 2001 and the Global Financial Crisis (2008-2009). For these financially distressed periods (and for the non-crisis observations), we compute the averaged response of each variable to a standardized impulse (unitary increase) to the NFCI.[15] Standardization across periods allows to isolate the transmission mechanism from changes in shock volatilities.

All figures report the 16th and 84th percentiles as bounds for the credible region. We do not detect noticeable time variation in the transmission mechanism, between crisis and non-crisis period, since the response to a standardized impulse is almost constant over time. We find that an impulse of one standard deviation to the NFCI significantly in-

[15]Remind that NFCI is constructed to have zero mean and unitary standard deviation.

creases the unemployment rate by about 0.8% within 10 quarters and depresses hours of work per employee. Unemployed workers get discouraged and some of them exit the labor force, given the negative and persistent response of the participation rate. Moreover, wage inflation response shows a significant reduction by about 0.4% at the peak, because of lower labor demand, captured by a significant drop in the vacancy rate. The simultaneous reduction in employment and vacancy results in a fall in labor market tightness:[16] according to search and matching models, when labor market is less tight, the job finding rate decreases, while the vacancy filling rate rises. The labor market slack drives a reduction in GDP (by more than 1% in levels at the peak). Interestingly, price inflation is slightly positive and borderline significant: this result is consistent with the VAR estimates by Abbate et al. (2016), who, however, use a different identification strategy.[17]

As shown in Figure 7, we find evidence of time-varying volatility of the identified financial shock. In particular, the estimated standard deviation is very high in the beginning of the sample (in correspondence to the Less-Developed-Countries Debt Crisis) and during the Global Financial Crisis. On the other hand, consistently with a large strand of the literature, during the Great Moderation the identified financial shock displays a relatively small standard deviation. In particular, Justiniano and Primiceri (2008) report that the key factor to understand the Great Moderation is the reduction in volatility of the investment shock, defined as innovation to the technology that transforms investments in capital goods. They mention that this innovation may be tied with the cost of external financing for firms, which has shown an important reduction in those years, due to massive financial liberalization.

To better assess the role of drifting volatilities, Figure 8 reports the non-standardized responses to a one-standard-deviation impulse. Time variation is driven by the stochastic volatility that determines the size of the impulse.

[16] Labor market tightness is defined as the vacancy/unemployed ratio. The numerator decreases after the shock; the denominator is affected by two countervailing forces: unemployment rate goes up but labor participation is lower. However, given that the reduction in the participation rate is small in magnitude, we conclude that labor market is less tight following a financial shock.

[17] Theoretical models are not unanimous on the effects of a financial tightening on inflation; on the one hand, tighter financial conditions reduce households' expenditure via negative wealth effects, generating a reduction in aggregate demand and so a lower price level (this is the mechanism at work in Gertler and Karadi (2011)); on the other hand, if firms borrow in order to buy intermediate inputs, an increasing risk-premium boosts the marginal cost of firms, which respond by raising the price of their products (as in Atta-Mensah and Dib (2008)). On top of that, Gilchrist et al. (2016) build a model in which firms increase prices when an adverse shock weakens their balance sheets, in order to maintain profitability; this paper provides also empirical evidence showing that, during the Great Recession, firms with a limited internal liquidity decided to raise their prices.

As stated previously, there is a growing theoretical literature analyzing the linkages between financial frictions and the labor market. Some of these works focus on the effects of a financial shock on labor market variables: our estimated impulse responses are consistent with these models, at least qualitatively. For instance, in Monacelli et al. (2011) a negative credit shock (by one standard deviation) decreases the employment rate by about 0.35% of the steady state value (which is close to one) at the response peak: this value is close to our median estimate for the crisis periods; GDP and wage responses are close to our median values as well. In Petrosky-Nadeau (2013) a two standard deviations negative credit shock raises the unemployment rate by 2% points at the peak, a value higher than our median estimates during crisis periods; however, both shape and persistence of the reaction are very similar.

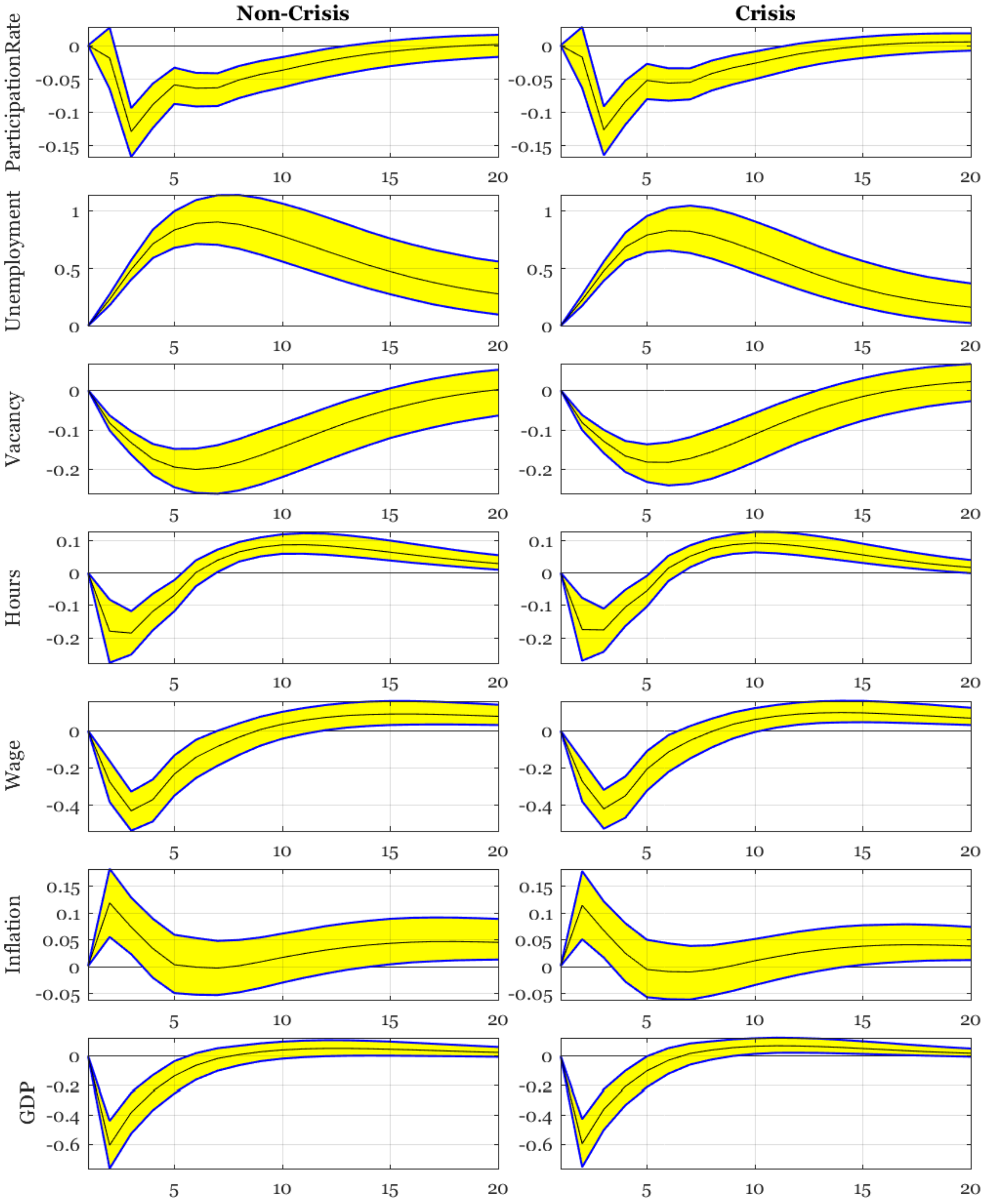


Figure 6: Time-varying impulse response functions to a standardized shock NFCI shock.

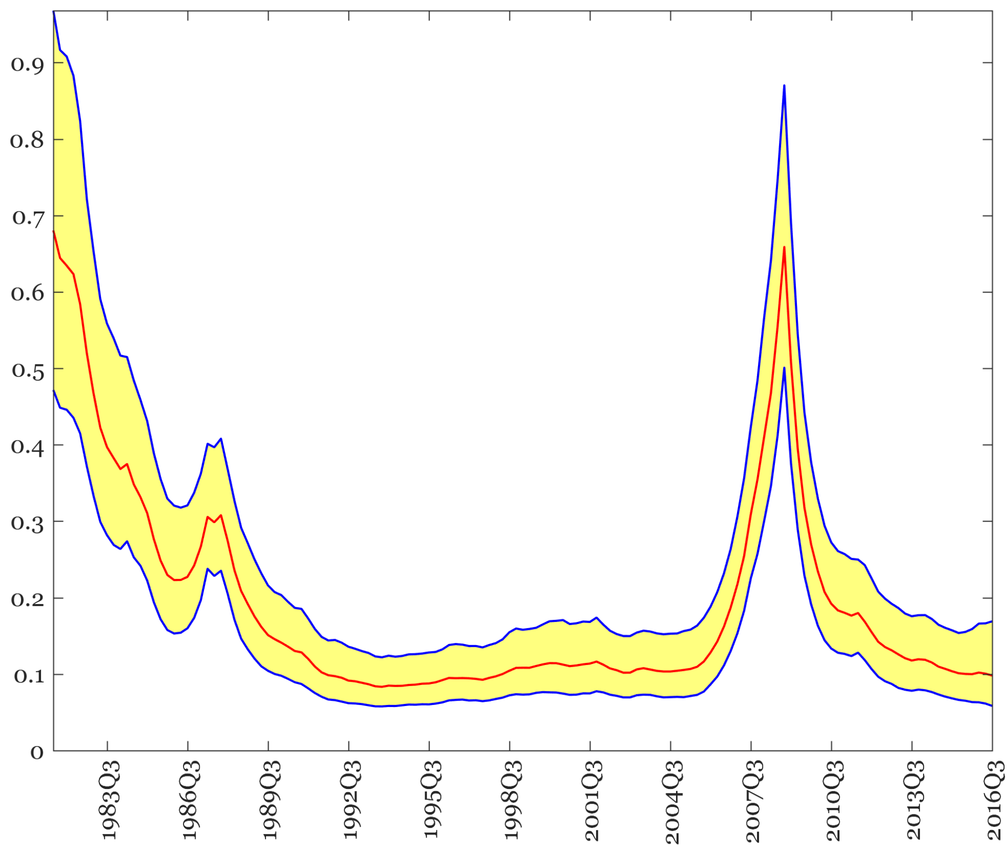


Figure 7: Stochastic volatility of the identified shock.

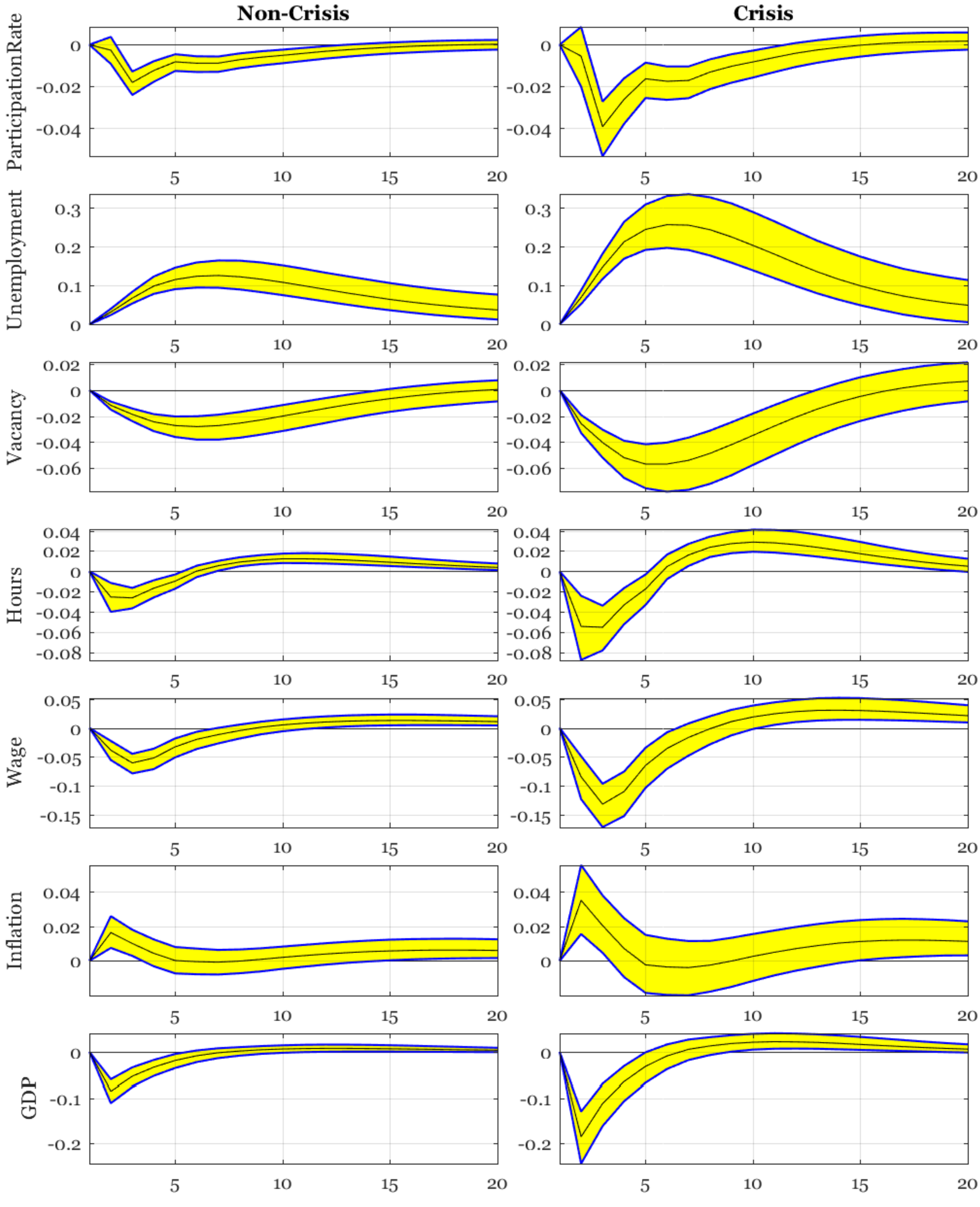


Figure 8: Time-varying impulse response functions to a (non-standardized shock) NFCI shock
.

5.1.2 Historical Decomposition: comparing CVAR and TV-VAR-SV

Having observed significant time variation in the identified financial shock volatility, where larger levels of volatility tend to be associated with financial distress periods, we perform a historical decomposition analysis for the variables of interest.

A CVAR assumes constant volatility of the identified shock. When estimated using financial crisis observations, i.e. when large economic reactions follow financial distress, a CVAR framework may overstate the role of financial shocks in explaining unemployment over the sample. The TV-VAR-SV, instead, is able to discriminate between periods of large and small volatility, complying with the asymmetry of observed financial shocks: large and negative or small and positive.

If we estimate a CVAR using the entire sample (Figure 10), we find that the financial shock's role in explaining unemployment rate fluctuations is roughly as important as for other non-identified shocks. This finding is consistent with Caldara et al. (2016) that use a CVAR to estimate the effects of financial and uncertainty shocks on macroeconomic variables.[18] However, when the CVAR estimation sample ends in 2006Q4 (Figure 9), the picture changes: the financial shock contribution is much smaller across the entire sample period. In these experiments we observe that Great Recession observations are driving the importance of financial shocks across the sample.

Figure 11 displays the historical decomposition of unemployment in a TV-VAR-SV framework. The financial shock hits the unemployment rate asymmetrically: contributions are large and negative in periods of severe financial distress like the Great Recession and the 1982-1984 crisis, while there is almost no positive contributions in the 1990s and early 2000s. This result is in contrast with the previous CVAR analysis and provides solid ground for more general time-varying models when performing structural analysis. This asymmetric behavior is explained by the time-varying volatility of the financial shock, which is estimated to be higher in period of financial distress: indeed, we find that negative shocks tend to be larger and less frequent than positive shocks.

[18]See figure B-1 in the working paper version of Caldara et al. (2016).

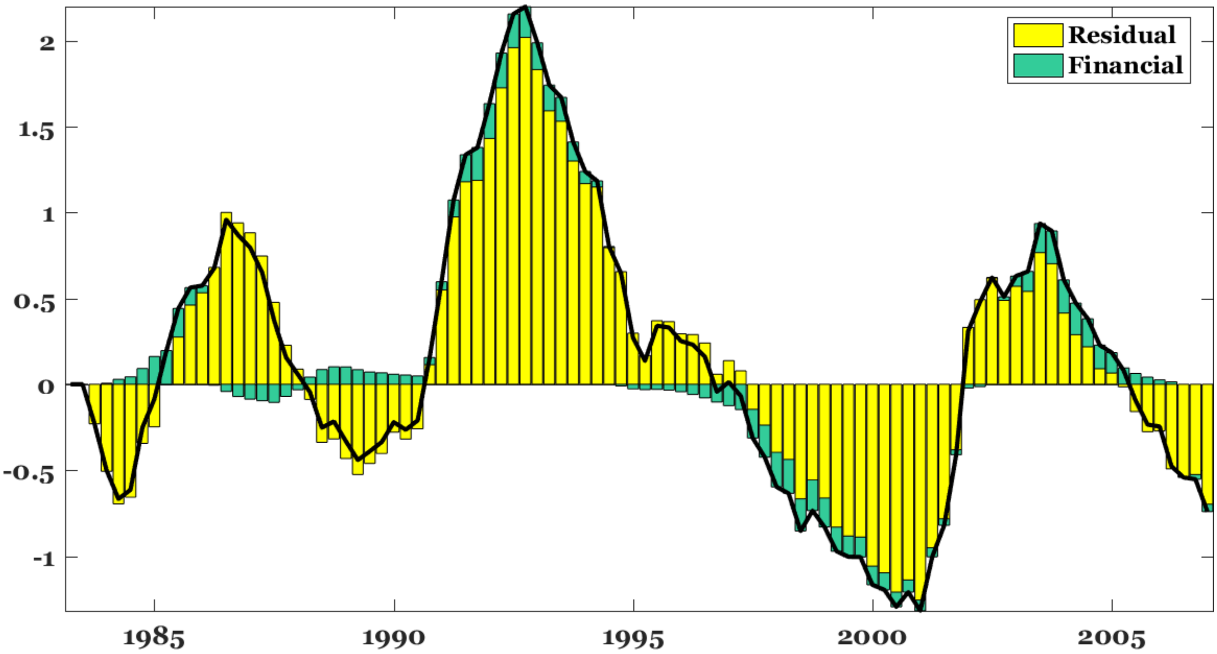


Figure 9: CVAR historical decomposition of unemployment, sample 1983Q1-2006Q4.

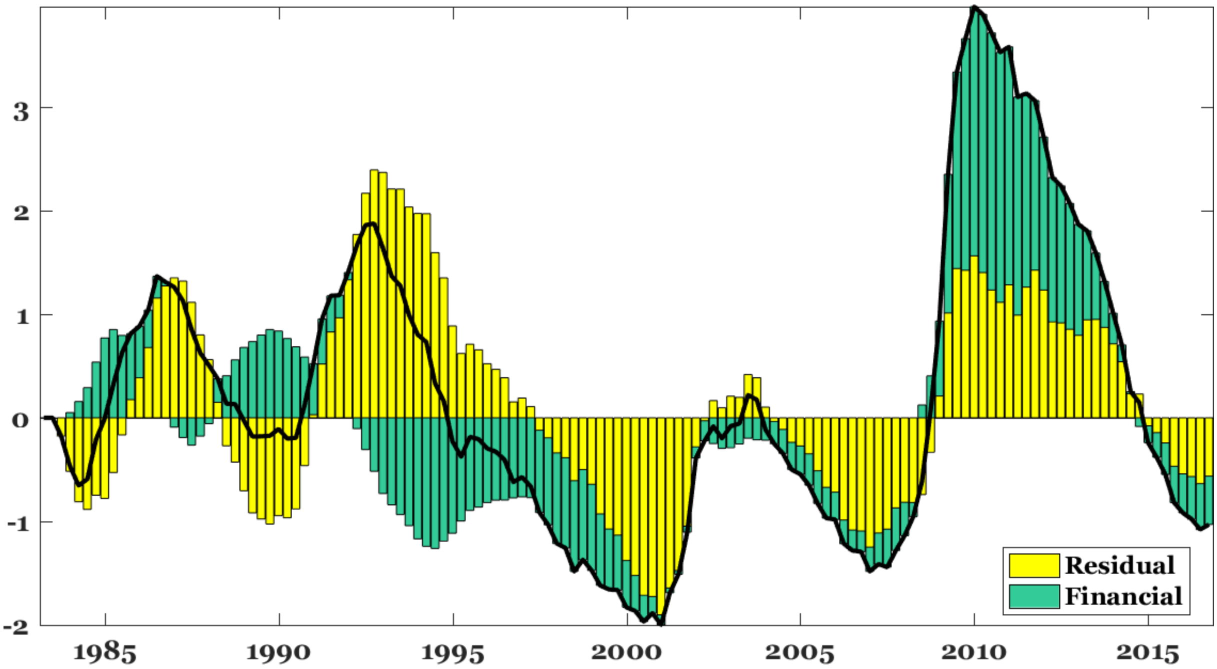


Figure 10: CVAR historical decomposition of unemployment, sample 1983Q1-2016Q3.

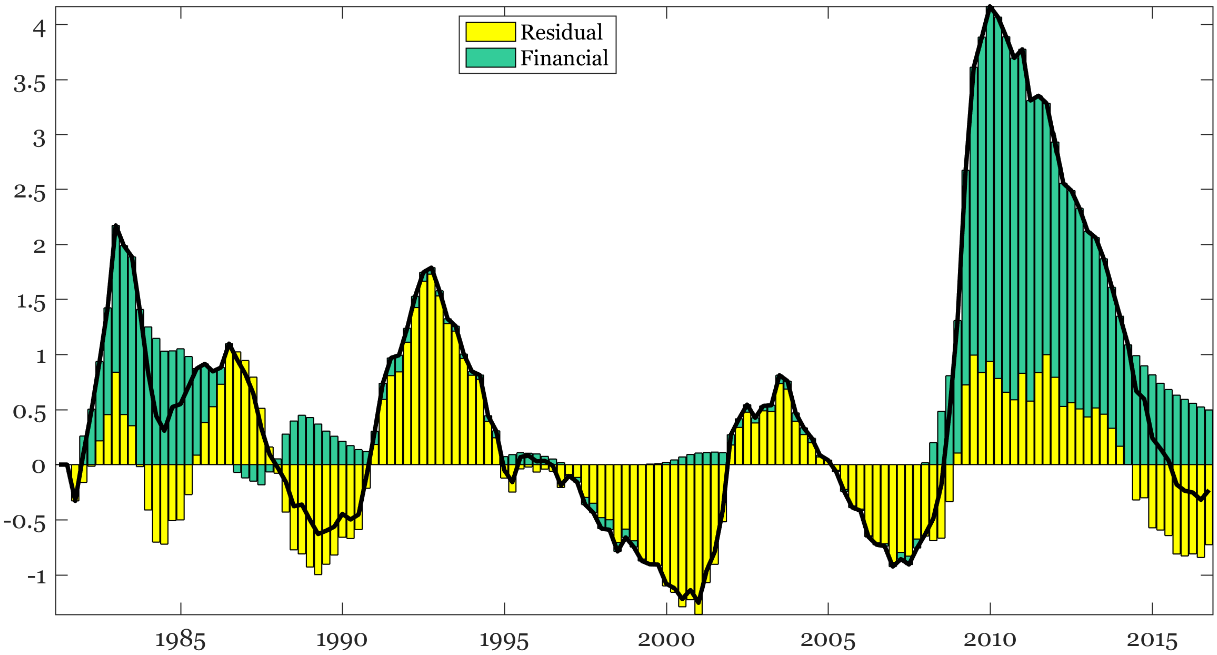


Figure 11: Historical decomposition of unemployment (TV-VAR-SV).

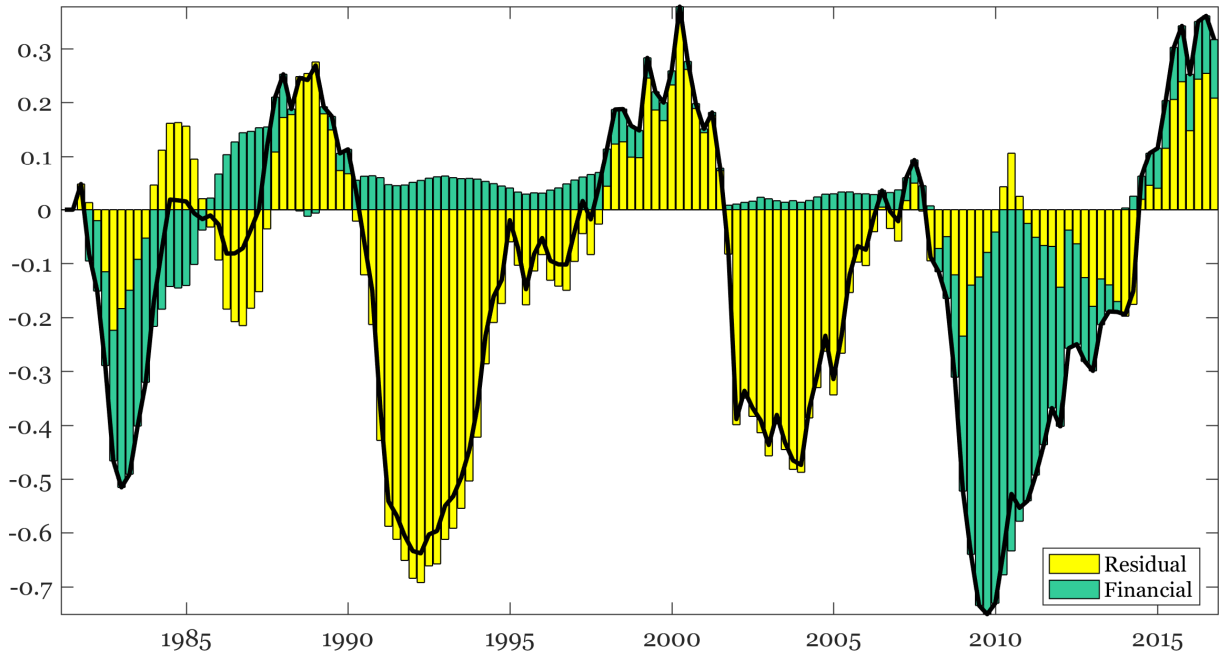


Figure 12: Historical decomposition of vacancies (TV-VAR-SV).

Also for the vacancy rate, negative financial shocks are the main responsible for the fall in 1982-1984 crisis and during the Great Recession, while positive financial shocks cause only mild response in other periods. This difference between the CVAR and TV-VAR historical decomposition is less evident for hours and GDP, as shown in the Appendix. Hours of work enter the VAR in (log) first differences: while the level of hours worked per employee is not fully back to the pre-crisis level, the growth rate turned to be positive in 2009 (most recent observations are above the pre-crisis mean).

Accordingly, our identified financial shock was less severe on hours worked in terms of persistence: it is the prolonged negative effect on the unemployment rate that generates this sizable differences between the CVAR and TV-VAR historical decomposition. The same reasoning holds for real GDP, that, like hours worked, enters our model in (log) first differences.

6 Robustness

In this section we perform two experiments to challenge the robustness of our results. First, we repeat the analysis using the credit spread index built by Gilchrist and Zakrajšek (2012). Second, we estimate the model by using monthly data. Most of the resulting figures for these robustness exercises are left in the Appendix.[19]

6.1 Excess Bond Premium

Gilchrist and Zakrajšek (2012) decompose their credit spread index in two components: the first one captures movements in expected defaults, the second one represents the excess bond premium (EB, henceforth) which measures cyclical changes in the relationship between default risk and credit spreads. In this subsection we verify the robustness of our results by using EB as financial variable, as in the VAR analysis of Gilchrist and Zakrajšek (2012); since the correlation between EB and NFCI is not so high (0.40), this represents a good test bench for our findings.[20] According to the CVAR historical decomposition, the 1990s were a period of positive financial shocks contribution. On the other hand, using a time-varying model, this positive contribution almost disappears, as in our benchmark specification.

[19] Additional figures are available upon request.

[20] An empirical analysis in a time-varying framework using the GZ index has been conducted by Abbate and Marcellino (2016), who, however, do not consider the labor market.

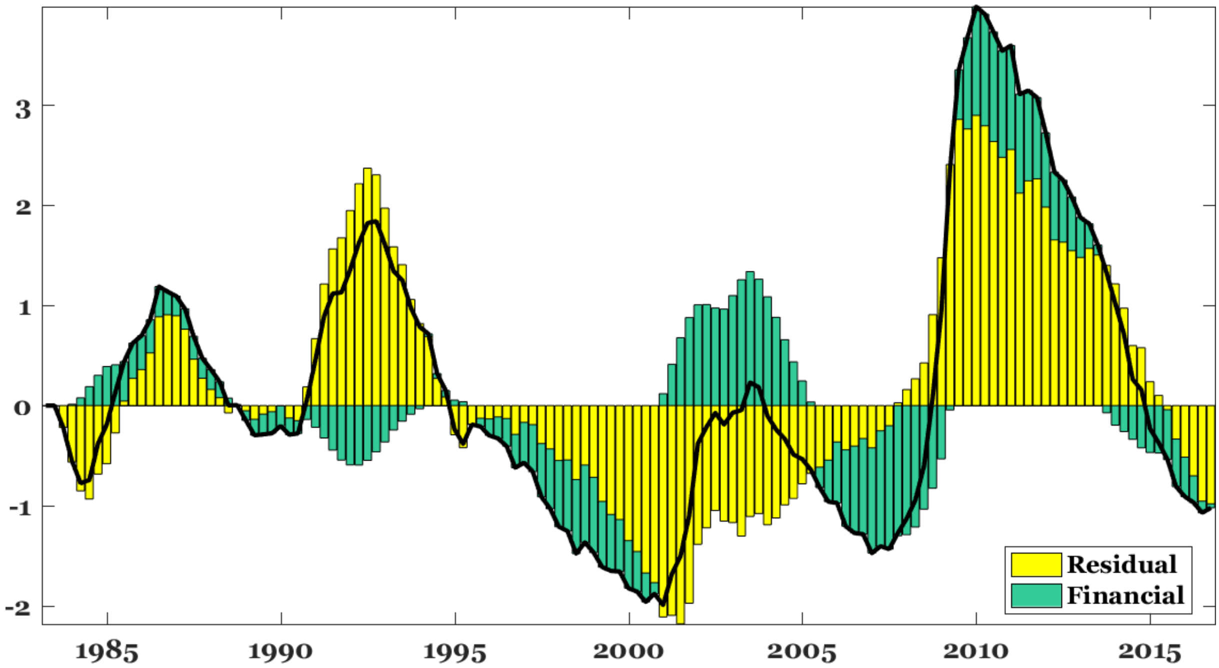


Figure 13: Historical decomposition of unemployment (EB specification, CVAR).

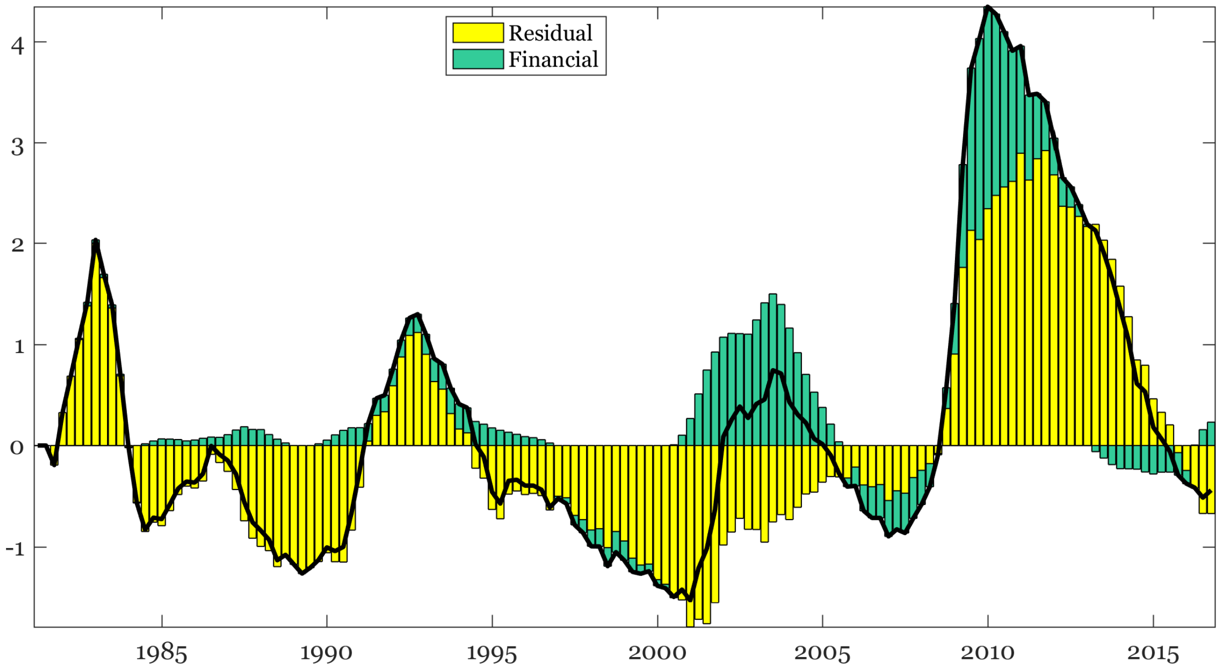


Figure 14: Historical decomposition of unemployment (EB specification, TV-SV-VAR).

Financial shock standard deviation is very low and pretty constant until the late 1990s, when it starts to increase, reaching a peak during the Great Recession. Impulse responses seem constant over time, when the impulse is standardized. Signs and sizes of responses are in line with the benchmark specification, except for the inflation response, which is negative but not statistically different from zero.

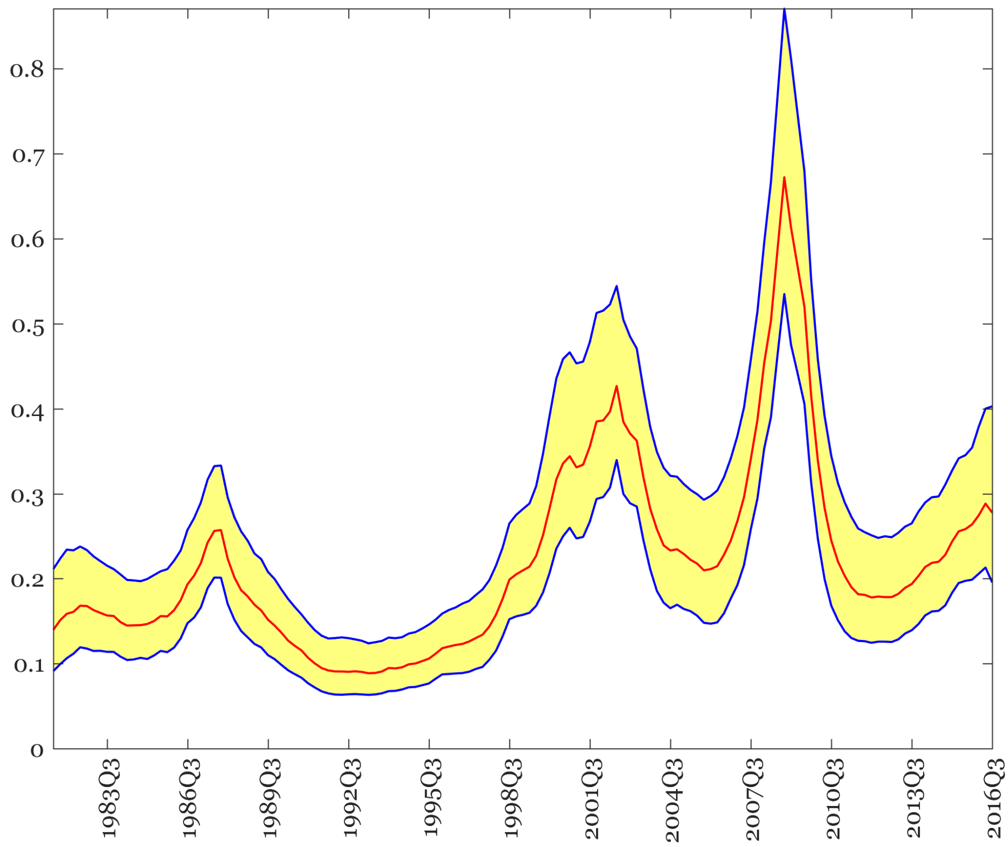


Figure 15: Stochastic volatility of the identified financial shock (EB specification).

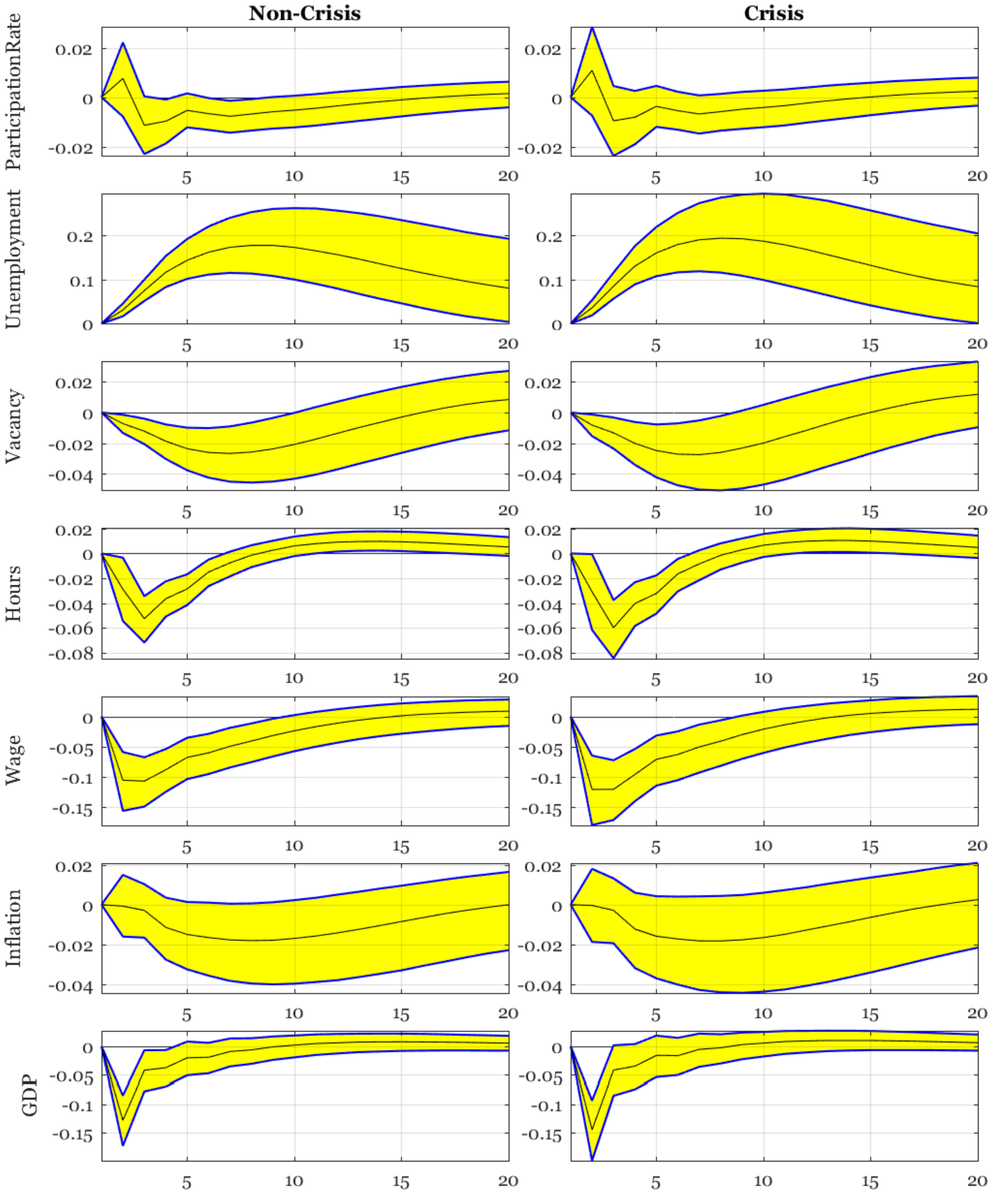


Figure 16: Time-varying impulse response functions to an EB (non-standardized) shock .

6.2 Fixed VAR Coefficients and Monthly Data

Our identification strategy is based on a short-run restriction: variables cannot contemporaneously react to an innovation in NFCI equation residuals. This assumption may seem questionable, if applied to quarterly data. In order to address this issue, we estimate the model using monthly data, keeping fixed the VAR coefficients, thus allowing time variation only in shocks' standard deviations. Indeed, when the model is estimated at monthly frequency, the number of coefficients to estimate gets huge. This is because, when the frequency switches from quarterly to monthly, the number of observations is multiplied by three, causing an increase in the number of lags p, as well as an increase of periods in which $k = n\,(np + 1)$ coefficients have to be estimated. However, since the benchmark specification does not find much time variation in VAR coefficients, a CVAR with stochastic volatility seems a good compromise.

The Gibbs Sampler to estimate a CVAR with stochastic volatility is a modification of the counterpart for the TV-VAR-SV described above, in which the coefficients' draw step is simply a Bayesian regression with informative prior, obtained by stacking the model in the following form:

$$\underbrace{Y}_{nT\times 1} = \underbrace{X}_{nT\times k} \cdot \beta + U, \qquad U \sim \mathcal{MN}\left(\mathbf{0},\, \underbrace{V_u}_{n\times n}\right)$$

where V_u is conditional on the history of stochastic volatilities drawn in the previous step. Prior information is incorporated in the posterior as suggested by Gelman et al. (2013).

Some of the series used in the baseline specification are not available at monthly frequency. The new series that we use are the following:

- Average hours of production and nonsupervisory employees in the private sector.
- Average hourly earnings of production and nonsupervisory employees in the private sector.
- The consumption price index.
- The industrial production.

The unemployment rate, the participation rate and the NFCI are available at monthly frequency, while the vacancy index is not and it is dropped from the estimation sample.

All the transformations are the same of the baseline specification, except for hours worked that now are not divided for number of employees, being an average measure.

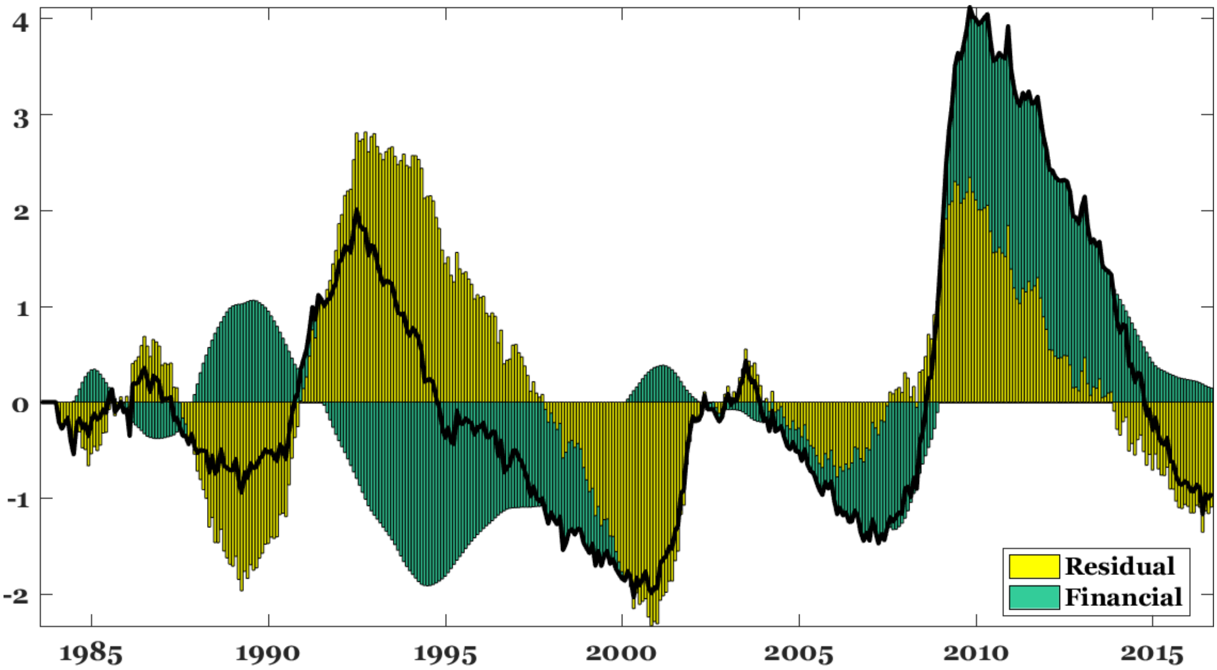


Figure 17: Historical decomposition of unemployment (CVAR, monthly specification).

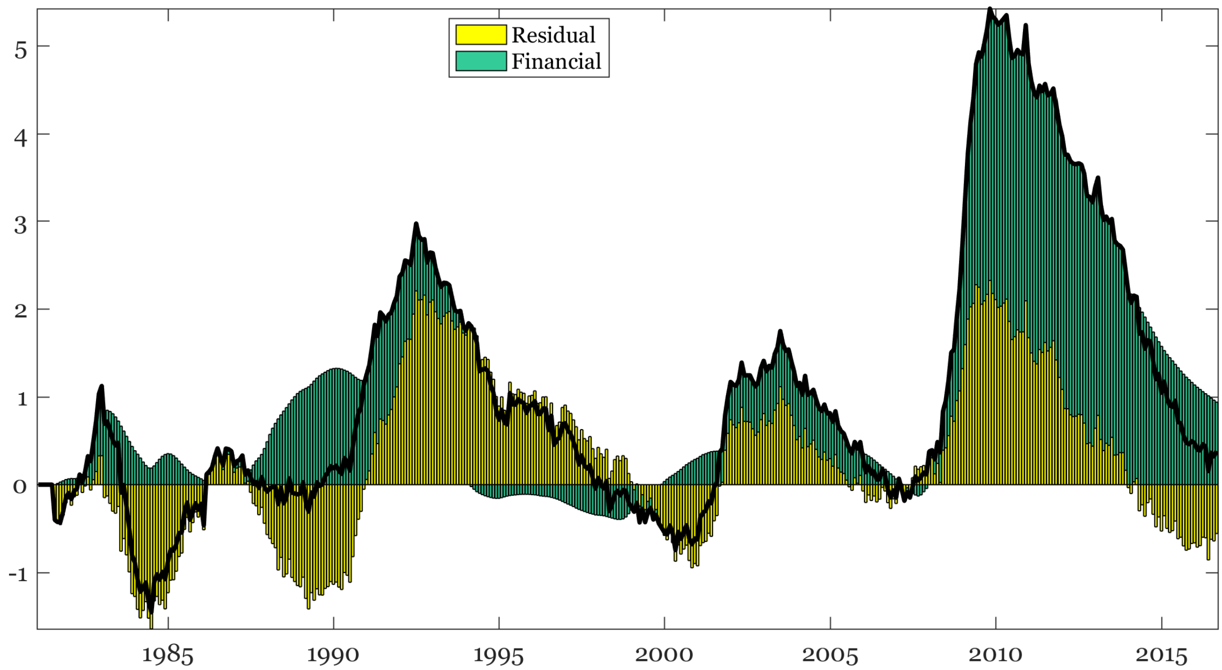


Figure 18: Unemployment historical decomposition (CVAR-SV, monthly specification).

Results from this new specification do not change the picture: positive financial shocks have a mild impact on unemployment fluctuations and this result cannot be captured by a VAR with constant shocks' variance. The estimated stochastic volatility has a pattern similar to the correspondent graph of the quarterly baseline specification. The impulse response functions of unemployment, hours, GDP and inflation do not qualitatively differ from the baseline specification. Now nominal wage inflation grows on impact; nevertheless, given inflation response, the real wage decreases and this is consistent with a fall in labor demand by firms.

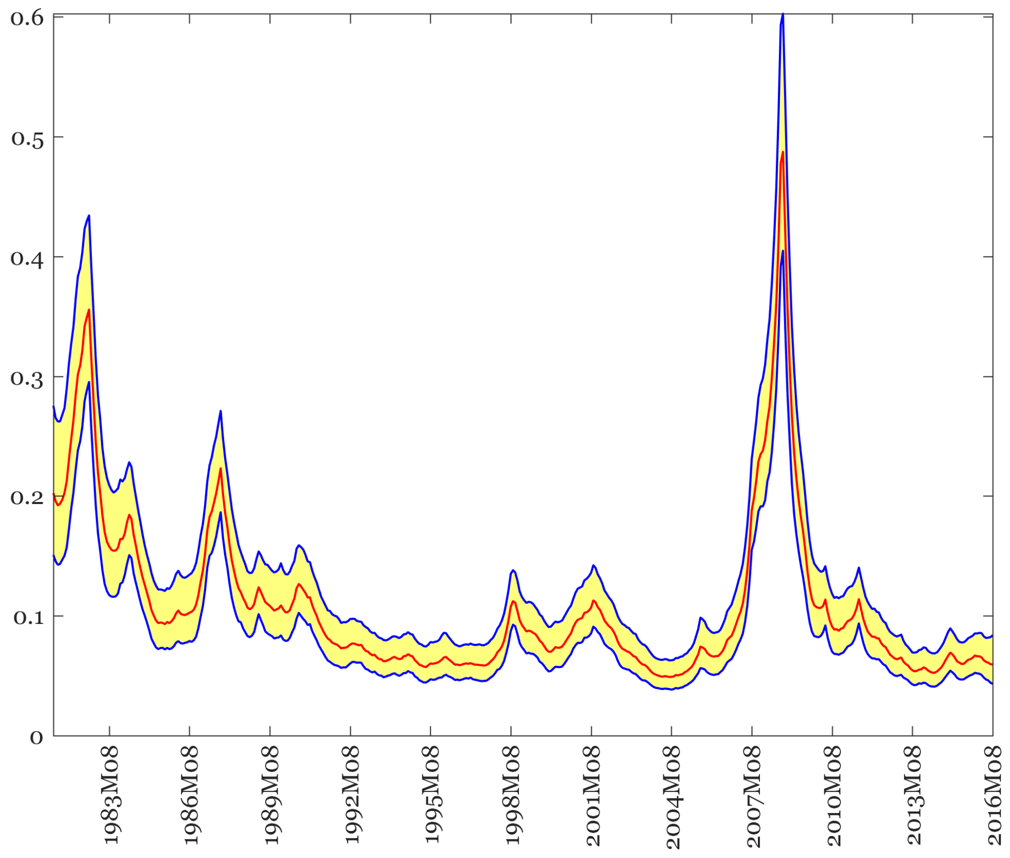


Figure 19: Stochastic volatility of the identified shock (monthly specification).

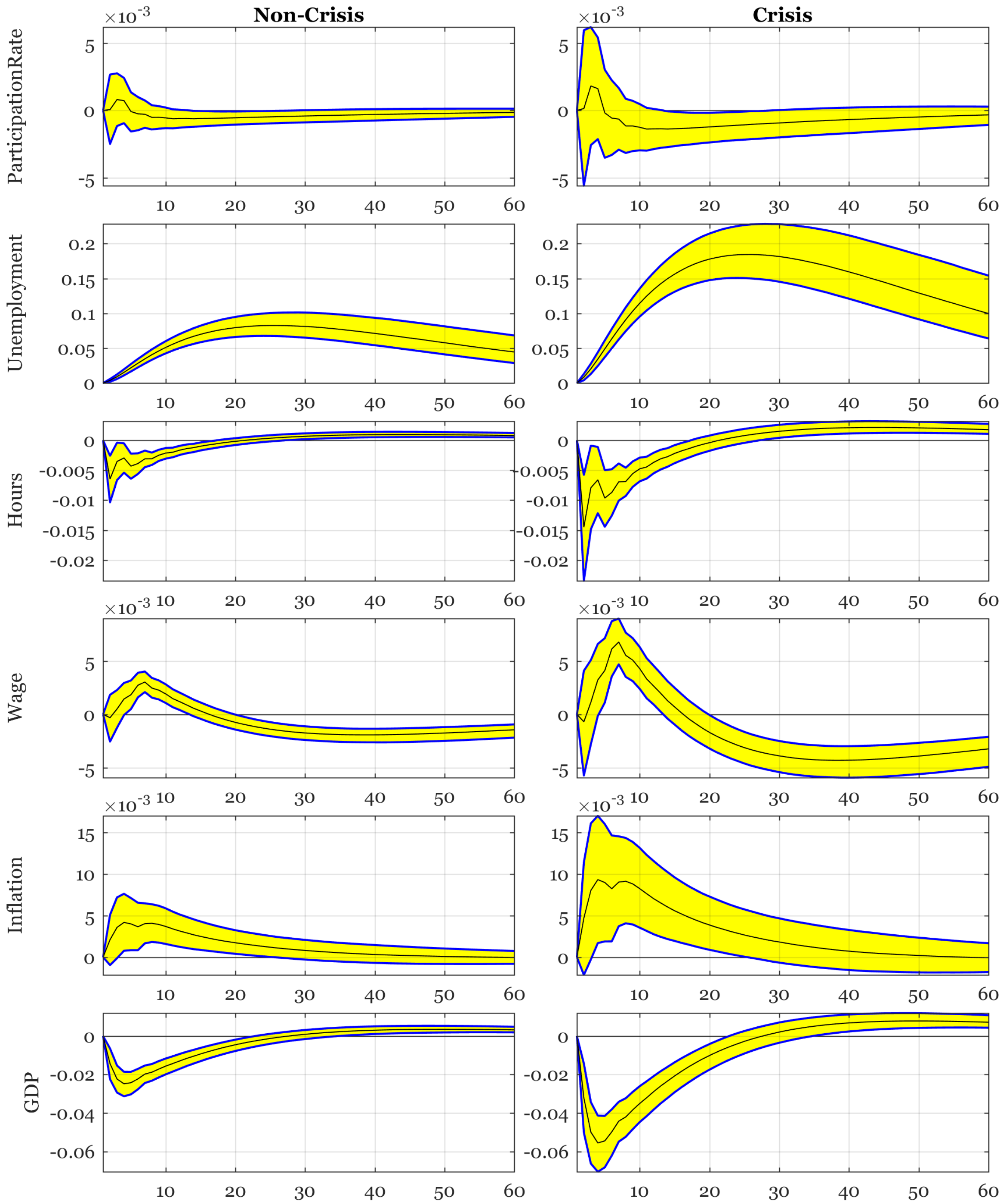


Figure 20: Impulse response functions (non-standardized, monthly specification).

7 Conclusions

In this paper we estimate a TV-VAR-SV in order to detect the effects of financial shocks on labor market variables, using the NFCI as financial indicator.

Differently from a CVAR framework, all the elements of a TV-VAR-SV are potentially time-varying: our findings report a relevant degree of time variation in the estimated financial shock's standard deviation, which is high in correspondence of periods of financial distress and low during the Great Moderation. However, our model does not find significant time variation in the transmission mechanism, since the estimated autoregressive coefficients are highly stable over time.

We show that positive financial shocks have been almost irrelevant in explaining unemployment fluctuations in the last three decades, while negative shocks have given a remarkable contribution. This result relies on the changing size of financial shocks, which is high during periods of financial turbulence, such as the Great Recession. This feature cannot be captured by a CVAR, which indeed finds both positive and negative contributions of financial shocks on the unemployment rate, in our estimation horizon.

In addition, we show that estimated response signs to an adverse financial impulse, identified with a short-run restriction, are reasonable and in line with the theoretical literature. A financial tightening generates a crisis in the labor market: firms reduce unemployment and per-capita hours worked to face the reduction in borrowing ability; they post less vacancies, so that resulting lower labor demand produces a fall in nominal wages; some unemployed workers get discouraged and the labor participation rate decreases; the response of real GDP is negative as well, due to a drop in labor input and a smaller demand of households, who become poorer. The response of inflation is positive but only borderline significant, reflecting both supply and demand effects at work.

From a policy perspective, our findings point out that macroprudential measures limiting the probability of a financial crisis may yield significant benefits to the economic activity via the labor market channel. Nonetheless, our model can not say anything about the effectiveness of these policies in reducing the standard deviation of financial shocks: thus, including a macroprudential indicator in macro-financial VARs with stochastic volatility seems an additional promising avenue to extend this line of research.

Bibliography

Abbate, A., Eickmeier, S., and Prieto, E. (2016). Financial Shocks and Inflation Dynamics. *CAMA Working Paper*, (53/2016).

Abbate, A. and Marcellino, M. (2016). Macroeconomic Activity and Risk Indicators: An Unstable Relationship. *Mimeo*.

Alessandri, P., Conti, A. M., and Venditti, F. (2017). The Financial Stability Dark Side of Monetary Policy. *Temi di discussione (Economic working papers), Bank of Italy*, (1121).

Atta-Mensah, J. and Dib, A. (2008). Bank Lending, Credit Shocks, and the Transmission of Canadian Monetary Policy. *International Review of Economics & Finance*, 17(1):159–176.

Barnichon, R. (2010). Building a Composite Help-Wanted Index. *Economics Letters*, 109(3):175–178.

Bernanke, B. S., Boivin, J., and Eliasz, P. (2005). Measuring the Effects of Monetary Policy: a Factor Augmented Vector Autoregressive (FAVAR) Approach. *Quarterly Journal of Economics*, 120(1):387–422.

Bernanke, B. S. and Gertler, M. (1989). Agency Costs, Net Worth, and Business Fluctuation. *American Economic Review*, 79(1):14–31.

Bernanke, B. S. and Gertler, M. (1995). Inside the Black Box: The Credit Channel of Monetary Policy Transmission. *Journal of Economic Perspectives*, 9(4):27–48.

Bernanke, B. S., Gertler, M., and Gilchrist, S. (1999). The Financial Accelerator in a Quantitative Business Cycle Framework. *Handbook of Macroeconomics*, 1:1341–1393.

Bianchi, J. (2011). Overborrowing and Systemic Externalities in the Business Cycle. *American Economic Review*, 101(7):3400–3426.

Brave, S. and Butters, R. A. (2011). Monitoring Financial Stability : a Financial Conditions Index Approach. *Economic Perspectives*, 35(1):22–43.

Brave, S. and Butters, R. A. (2012). Diagnosing the Financial System: Financial Conditions and Financial Stress. *International Journal of Central Banking*, 8(2):191–239.

Caldara, D., Fuentes-Albero, C., Gilchrist, S., and Zakrajšek, E. (2016). Themacroeconomic impact of financial and uncertainty shocks. *European Economic Review*, 88:185–

207.

Carter, C. K. and Kohn, R. (1994). On Gibbs Sampling for State Space Models. *Biometrika*, 81(3):541–553.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (1999). Monetary Policy Shocks: What Have We Learned and to What End? *Handbook of macroeconomics*, 1:65–148.

Christiano, L. J., Motto, R., and Rostagno, M. (2003). The Great Depression and the Friedman-Schwartz Hypothesis. *Journal of Money, Credit, and Banking*, 35(6):1119–1198.

Christiano, L. J., Trabandt, M., and Walentin, K. (2011). Introducing Financial Frictions and Unemployment into a Small Open Economy Model. *Journal of Economic Dynamics and Control*, 35(12):1999 – 2041.

Cogley, T. and Sargent, T. J. (2005). Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US. *Review of Economic dynamics*, 8(2):262–302.

Corsello, F. (2018). Rank Reduction in BVARs with Time Varying Coefficients and Stochastic Volatility: Specification and Estimation. *Bocconi University, mimeo.*

de Wind, J. and Gambetti, L. (2014). Reduced-Rank Time-Varying Vector Autoregressions. *CPB Discussion Paper*, (270).

Del Negro, M. and Primiceri, G. E. (2015). Time Varying Structural Vector Autoregressions. *Review of Economic Studies, Forthcoming.*

Ferraro, D. (2017). The Asymmetric Cyclical Behavior of the U .S . Labor Market. *Arizona State University Working Paper.*

Fink, F. and Schüler, Y. S. (2015). The Transmission of US Systemic Financial Stress: Evidence for Emerging Market Economies. *Journal of International Money and Finance*, 55:6–26.

Fornari, F. and Stracca, L. (2012). What Does a Financial Shock Do? First International Evidence. *Economic Policy*, 71(27):407–445.

Gaiotti, E. (2013). Credit Availability and Investment: Lessons from the Great Recession. *European Economic Review*, 59:212–227.

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., and Rubin, D. (2013). *Bayesian Data Analysis, Third Edition*. CRC Press.

Gertler, M. and Karadi, P. (2011). A Model of Unconventional Monetary Policy. *Journal* *of Monetary Economics*, 58(1):17–34.

Gilchrist, S., Schoenle, R., Sim, J., and Zakrajšek, E. (2016). Inflation Dynamics During the Financial Crisis. *NBER Working Paper*, (No. w22827).

Gilchrist, S., Yankov, V., and Zakrajšek, E. (2009). Credit Market Shocks and Economic Fluctuations: Evidence from Corporate Bond and Stock Markets. *Journal of Monetary Economics*, 56(4):471–493.

Gilchrist, S. and Zakrajšek, E. (2012). Credit Spreads and Business Cycle Fluctuations. *American Economic Review*, 102(4):1692–1720.

Guerrieri, L. and Iacoviello, M. (2017). Collateral Constraints and Macroeconomic Asymmetries. *Journal of Monetary Economics*, 90:28–49.

Iliopulos, E., Langot, F., and Sopraseuth, T. (2014). Welfare Cost of Fluctuations: when Labor Market Search Interacts with Financial Frictions. *Documents de travail du Centre d'Economie de la Sorbonne 14042*.

Jermann, U. and Quadrini, V. (2012). Macroeconomic Effects of Financial Shocks. *American Economic Review*, 102(1):238–71.

Justiniano, A. and Primiceri, G. E. (2008). The Time-Varying Volatility of Macroeconomic Fluctuations. *American Economic Review*, 98(3):604–41.

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. *Review of Economic Studies*, 65(3):361–393.

Kiyotaki, N. and Moore, J. (1997). Credit Cycles. *Journal of Political Economy*, 105(2):211–248.

Lopez-Salido, D. and Nelson, E. (2010). Postwar Financial Crises and Economic Recoveries in the United States. *Working Paper, FED Board*.

McKay, A. and Reis, R. (2008). The Brevity and Violence of Contractions and Expansions. *Journal of Monetary Economics*, 55(4):738–751.

Mendoza, E. G. (2010). Sudden Stops, Financial Crises and Leverage. *The American Economic Review*, 100(5):1941–1966.

Metiu, N., Hilberg, B., and Grill, M. (2015). Financial Frictions and Global Spillovers. *Bundesbank Discussion Paper*, 04/2015.

Monacelli, T., Quadrini, V., and Trigari, A. (2011). Financial Markets and Unemployment. *NBER Working Paper*, 17389.

Mortensen, D. T. and Pissarides, C. A. (1994). Job Creation and Job Destruction in the Theory of Unemployment. *The Review of Economic Studies*, 61(3):397–415.

Omori, Y., Chib, S., Shephard, N., and Nakajima, J. (2007). Stochastic Volatility with Leverage: Fast and Efficient Likelihood Inference. *Journal of Econometrics*, 140(2).

Petrosky-Nadeau, N. (2013). Credit, Vacancies and Unemployment Fluctuations. *Review of Economic Dynamics*, 17(2):191–205.

Petrosky-Nadeau, N. and Wasmer, E. (2013). The Cyclical Volatility of Labor Markets under Frictional Financial Markets. *American Economic Journal: Macroeconomics*, 5(1):193–221.

Pizzinelli, C. and Zanetti, F. (2017). State Dependence in Labor Market Fluctuations: Evidence, Theory and Policy Implications. *Mimeo*.

Prieto, E., Eickmeier, S., and Marcellino, M. (2016). Time Variation in Macro-Financial Linkages. *Journal of Applied Econometrics*.

Primiceri, G. E. (2005). Time Varying Structural Vector Autoregressions. *The Review of Economic Studies*, 72(3):821–852.

Silvestrini, A. and Zaghini, A. (2015). Financial Shocks and the Real Economy in a Nonlinear World: From Theory to Estimation. *Journal of Policy Modeling*, 37(6):915–929.

Stock, J. H. and Watson, M. W. (2012). Disentangling the Channels of the 2007–09 Recession. *Brookings Papers on Economic Activity*, 2012(1):81–135.

Zanetti, F. (2017). Financial Shocks and Labor Market Fluctuations. *Macroeconomic Dynamics*, (forthcoming).

A Additional Figures

A.1 Baseline Specification

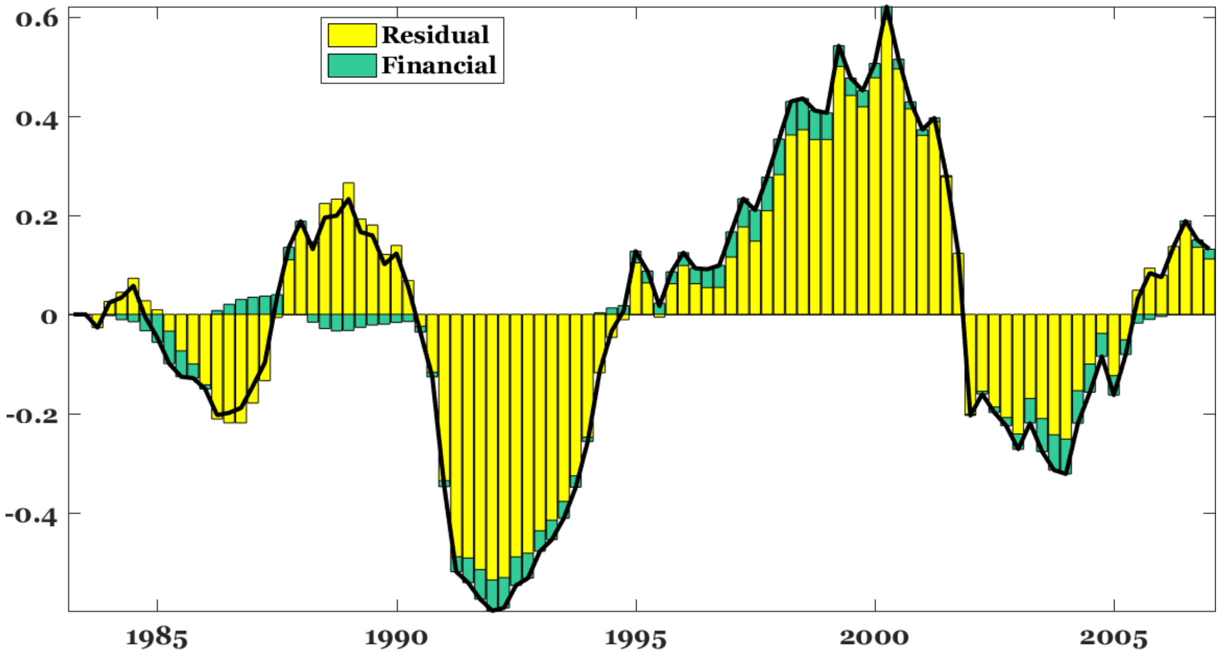


Figure A.1: CVAR historical decomposition of vacancy rate, sample 1983Q1-2006Q4.

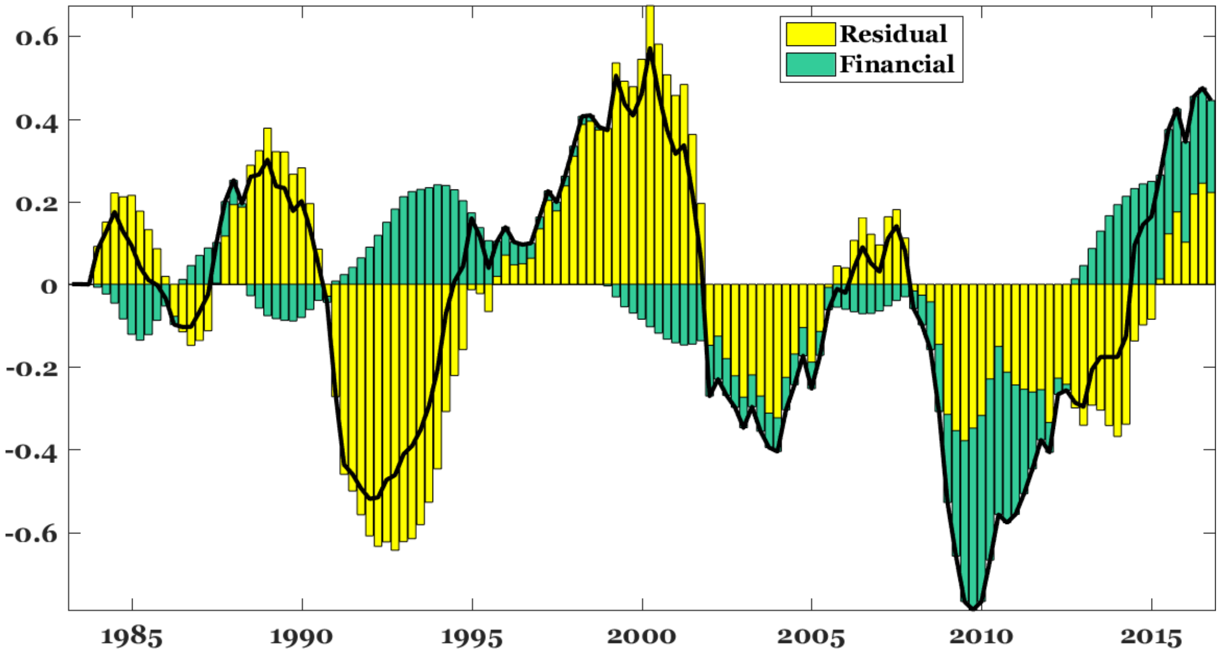


Figure A.2: CVAR historical decomposition of vacancy rate, sample 1983Q1-2016Q3.

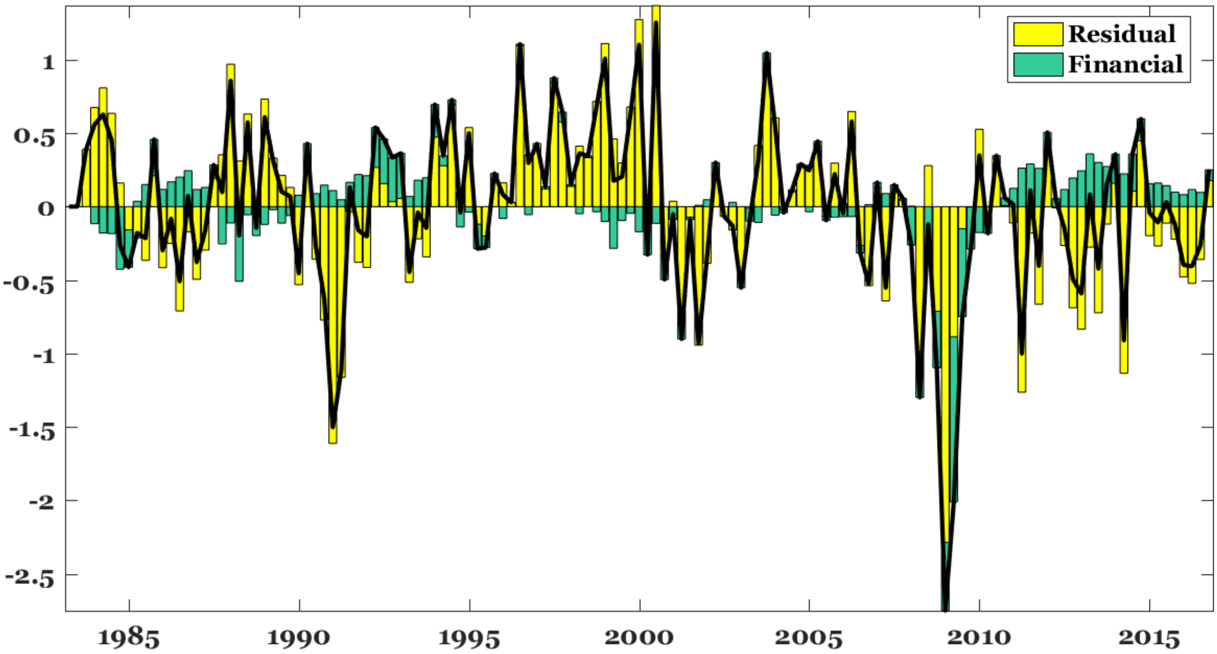


Figure A.3: Historical Decomposition of GDP (CVAR).

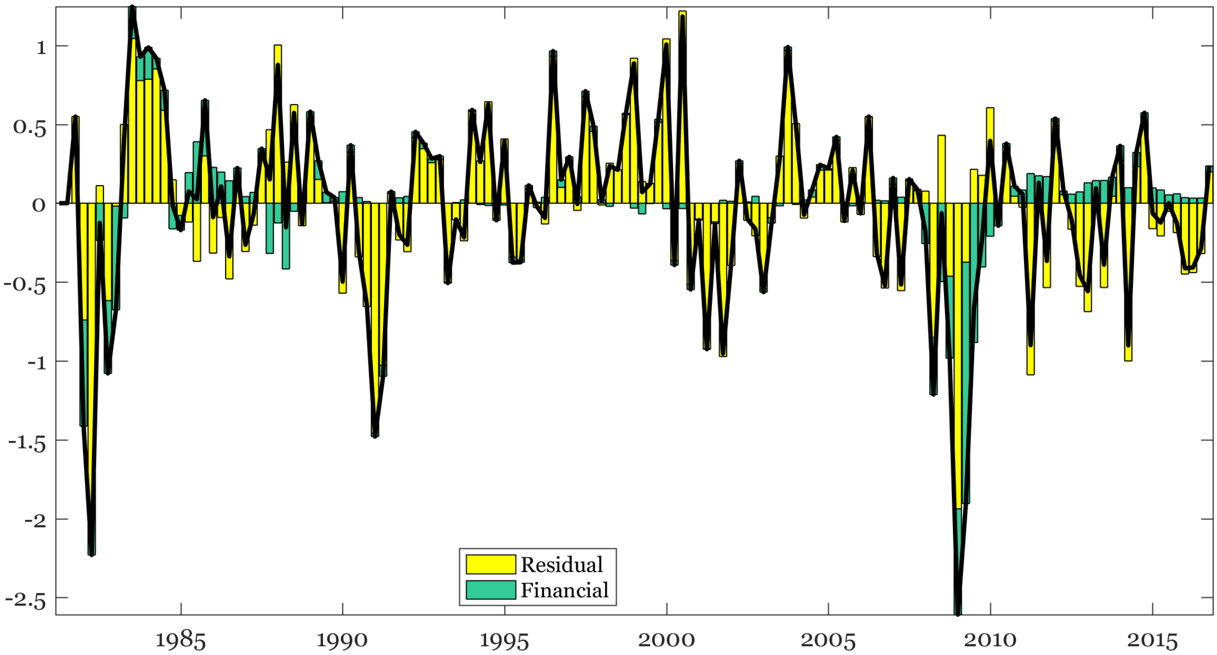


Figure A.4: Historical Decomposition of GDP (TV-VAR-SV).

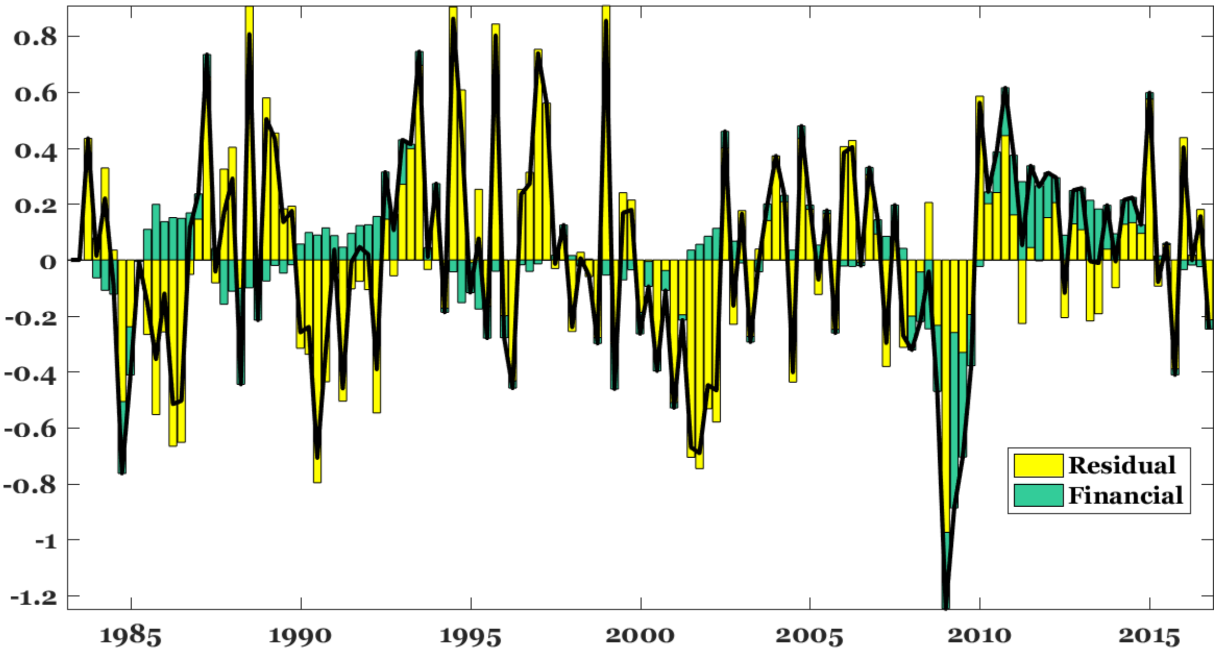


Figure A.5: Historical Decomposition of hours (CVAR).

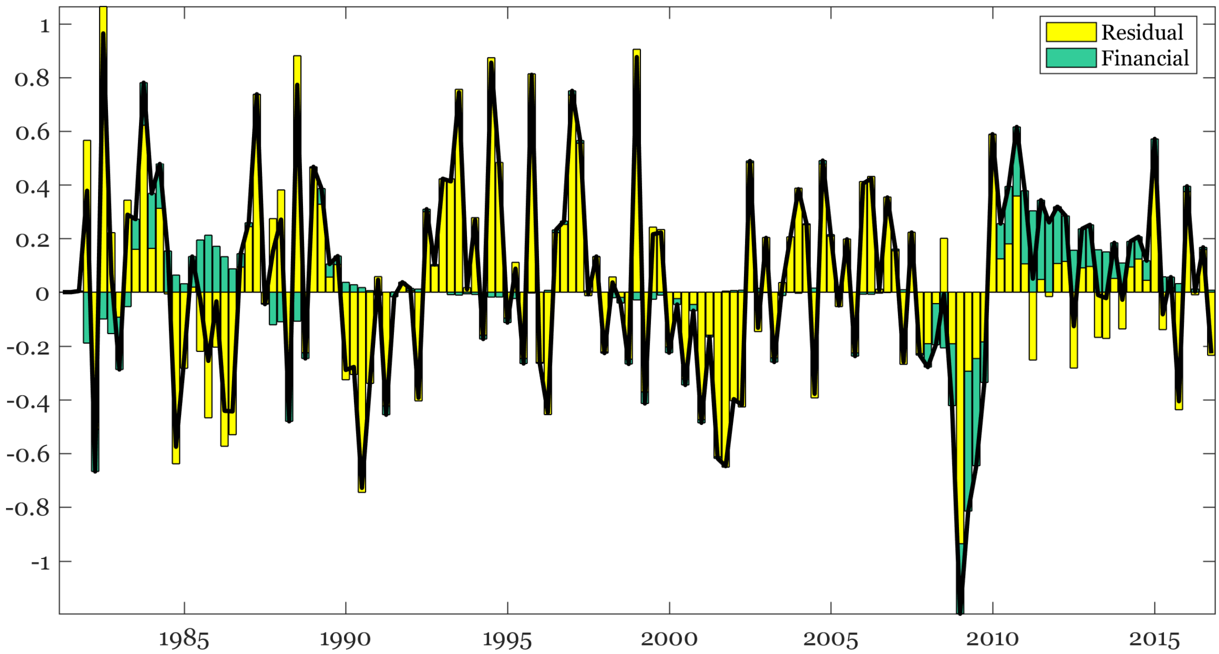


Figure A.6: Historical Decomposition of hours (TV-VAR-SV).

Chapter 3

The Economic Drivers of Volatility and Uncertainty

co-authored with Andrea Carriero and Massimiliano Marcellino

1 Introduction

The seminal paper by Engle (1982) emphasized the importance of changing conditional volatilities in economic and financial variables. The ensuing literature on this topic is massive, starting with time-series contributions but then entering also the mainstream economic literature, with the debate on the "Great Moderation's" origin, see, e.g., Fernández-Villaverde and Rubio-Ramırez (2013). Supporters of the "good luck" rather than "good policy" explanation emphasized the importance of allowing for stochastic volatility in common econometric models, such as VARs, see e.g. Primiceri (2005), Cogley and Sargent (2005) and Sims and Zha (2006).[1] The debate was hot again after the financial crisis hit, and there seems to be growing consensus on the importance of allowing for economic and financial shocks with particularly large variance during crisis periods, see for example Abbate et al. (2016) and references therein.

A common feature of papers modeling volatilities as stochastic and time-varying is that different shocks can drive both levels and volatilities of the variables, but the nature of shocks driving volatilities is left unexplained. This contrasts with the use of structural economic shocks to explain changes in conditional means of the variables. In this paper we would like to close this gap, introducing a model where the same structural shocks can drive the evolution of both conditional mean and variance of the variables.

The first step is the identification of structural shocks. The literature has somewhat converged on the importance of allowing for a large enough information set when identifying structural shocks, in order to avoid omitted variable bias. This can be achieved by using either large BVARs, see e.g. Bańbura et al. (2010), or large factor models, see e.g. Bernanke et al. (2005). In both types of models it is possible to introduce changing volatilities, see e.g. Carriero et al. (2016a) and Eickmeier et al. (2015), respectively. However, neither model is ideal. The VAR requires the number of structural and reduced form shocks to be the same, which complicates identification, in particular when many variables are jointly considered. Actually, in the large VAR literature typically only a monetary shock is identified, see e.g. Bańbura et al. (2010) or Carriero et al. (2016b). The factor model, or the factor augmented VAR model (FAVAR), allows for a small number of structural economic shocks common to all variables (plus variable specific idiosyncratic shocks), which is more

[1]From a reduced form perspective, Clark and Ravazzolo (2015) reports evidence that introducing stochastic volatilities improves the forecasting accuracy for macro-financial variables, strengthening the evidence already gathered by D'Agostino et al. (2013) and Carriero et al. (2015).

in line with economic theory. However, being the factors unobservable and estimated as a linear combination of all variables, the identification, estimation and interpretation of structural shocks has a relevant amount of complexity.

Our proposed model has features of both VARs and factor models. It is a large Multivariate Autoregressive Index (MAI) Model, which is a reduced rank VAR where each variable is driven by (lags of) linear combinations of groups of variables, and the latter can be interpreted as observable factors, or "Indexes", see e.g. Reinsel (1983), Geweke (1996) and Carriero et al. (2016b) [CKM16 henceforth] for details. Specifically, we will have real, nominal and financial factors combining, respectively, real nominal and financial variables. The indexes follow themselves a VAR, which is an implication of the MAI model rather than a priori assumption as in factor models. Hence, structural demand, supply and monetary/financial shocks can be identified from the VAR of observable factors, using common techniques.

We include Stochastic Volatility (SV) in the MAI model, while existing MAI models are homoskedastic. In the MAI-SV model, errors in both the reduced rank VAR and the VAR of observable factors feature SV. More specifically, applying a decomposition similar to that introduced by Johansen (1995) in the context of cointegrated models, we derive a time-varying decomposition of the estimated volatility of each variable into two orthogonal components: one common to all the variables, the other variable specific. Moreover, we show that the common volatility component can be further decomposed into three orthogonal components driven by volatilities of the structural demand, supply and monetary / financial shocks driving the common factors, and therefore also the conditional means of all variables.

Papers such as Jurado et al. (2015) and Carriero et al. (2017) interpret the common volatility as an uncertainty measure, as when it increases it becomes more difficult to accurately predict the future evolution of the variables. In their case, there are one or two volatility factors, interpreted as macro and financial uncertainty. In our case, there can be more factors, each driven by structural economic shocks. Hence, our results can also shed light on the economic drivers of uncertainty.

We develop a Gibbs Sampling algorithm for (Bayesian) inference in the MAI-SV model, allowing for time-variation both in variances and covariances across shocks, making use of efficient strategies to reduce the computational burden resulting from the large number of unobservable state variables.

Next, we apply the model using 20 macroeconomic and financial variables for the USA, the same analyzed by Bańbura et al. (2010) and CKM16 but for an extended period that covers 1964-2016. We find, not surprisingly and in line with the literature, that allowing for SV is quite important. The identified demand, supply and monetary / financial shocks have effects on the conditional expectations of the variables, qualitatively in line with economic theory and with the results in Bańbura et al. (2010) and CKM16, even though in our context the actual size of the shocks, and therefore their effects, are changing over time due to SV.

Moreover, the common component of volatility is substantial, in line with the results in Carriero et al. (2017), though there are differences across variables. Interestingly, the common component of volatility seems to have diminished its importance over time for employment and, less so, for earnings, personal income, and consumption, with values well below 50% at the end of the sample. The fraction of volatility explained by the common component is much higher for industrial production and capacity utilization, but it drops substantially for both variables after the financial crisis, from values around 80-90% to slightly above 50%. For the unemployment rate and housing starts, the fraction is instead stable at about 50% over the entire sample. The fraction of common conditional time-varying volatility is rather stable and high for nominal variables, much more so for CPI and the PCE deflator than for PPI, but still well above 50% for the latter. A similar pattern emerges also for the financial variables, with values around 50% for the Fed Funds rate and the 10 year T-Bond but, interestingly, much higher for the S&P500, whose time-varying volatility seems to be substantially affected by economic shocks.

Further, when decomposing the common volatility components, it turns out that demand shocks are clearly dominant for industrial production and capacity utilization, in the sense that an increase in the volatility of the former leads to an increase in the volatility of the latter. The same was true for employment at the beginning of the sample, but then the role of demand shocks has progressively declined in favor of supply and monetary/financial shocks, that combined explained more than 50% of the common volatility component of employment at the end of the sample. A decling role for demand shocks is evident, though milder, also for housing starts. Instead, for earning, income, consumption and unemployment the share of demand shocks is rather stable over time, below 50% for the three variables and close to 50% for unemployment. Supply and monetary/financial shocks explain a comparable fraction of the remaning common volatility component of these four variables.

For the common volatility of the nominal variables, supply shocks are dominant, in particular for CPI and the PCE deflator, with demand shocks ranked second and monetary/financial shocks third, though slightly more important for PPI.

For the common volatility of financial variables, the contribution of different structural shocks is overall comparable and rather stable over time. The monetary / financial shock is particularly important to explain the common component of the S&P500 volatility, but even for this variable it only explains about 50%, with the other 50% close to equally split between demand and supply shocks.

Overall, the structural demand, supply and financial shocks appear as important drivers not only of the level of macroeconomic and financial variables but also of their changing conditional volatility. Their volatility explains, in general, more than 50% of the overall volatility, with their relative importance being variable dependent and, sometimes, changing over time.

From an economic point of view, the volatility of the structural demand shock could be related to changes in the inventory mechanism (e.g., McConnell and Perez-Quiros, 2000) or to the globalization process (e.g., Bianchi and Civelli, 2015), while that of the supply shock can depend on the behavior of the oil market (e.g., Lee et al., 1995) or on firms' productivity (e.g., Christiano et al., 2010), and the volatility of the financial/monetary shock can be influenced by financial innovation (e.g., Dynan et al., 2006) or by changes in the conduct of monetary policy (e.g., Clarida et al., 2000). Justiniano and Primiceri (2008) introduce a DSGE model with stochastic volatility in several structural shocks. Specifically, they find that a reduction in the variance of investment shocks, interpreted also as a proxy for unmodeled financial frictions, is the main driver of the US Great Moderation, with a limited role for changes in monetary policy. Their evidence is somewhat in line with our common volatility decomposition that shows the gradually decreasing role of the monetary/financial shock contribution from the 1980s to the Great Financial Crisis. However, our structural decomposition is based on the common part of the volatility which is only a fraction of the total volatility. Moreover, Benati and Surico (2009) show that it can be difficult to map changes in the variance of structural VAR shocks with those in DSGE models.

The remainder of the paper is structured as follows. Section 2 reviews the MAI model and extends it by introducing SV. Section 3 discusses theoretically the SV decomposition. Section 4 presents the choice of priors distributions and develops the MCMC algorithm for estimation, with additional details provided in the appendix. Section 5 presents the empirical results. Section 6 concludes.

2 The MAI-SV Model

This section presents the econometric model used in the paper. In particular, the MAI model proposed by CKM16 is first reviewed, and then augmented with stochastic volatility, with either fixed or time-varying contemporaneous relations among innovations.

2.1 The MAI model

An n-dimensional zero mean weakly stationary process y_t is often represented as a $VAR(p)$:

$$\underbrace{y_t}_{n\times 1} = \sum_{\ell=1}^{p} \underbrace{\Phi_\ell}_{n\times n} y_{t-\ell} + u_t, \qquad u_t \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0}, \underbrace{\Omega}_{n\times n}\right). \tag{1}$$

When the number of variables, n, is large, the $VAR(p)$ is affected by the curse of dimensionality, which could even prevent its estimation when the number of degrees of freedom is too low. One possibility to reduce the number of parameters is to impose reduced rank restrictions on the matrices governing the dynamic evolution of the variables. Specifically, following Reinsel (1983) and CKM16, we assume that $\Phi_\ell = A_\ell \cdot B_0$, $\ell = 1, ..., p$, so that

$$\underbrace{y_t}_{n\times 1} = \sum_{\ell=1}^{p} \underbrace{A_\ell}_{n\times r} \cdot \underbrace{B_0}_{r\times n} y_{t-\ell} + u_t, \qquad u_t \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0}, \underbrace{\Omega}_{n\times n}\right). \tag{2}$$

The specification in (2) is a Multivariate Autoregressive Index (MAI) model. Each of the n variables is driven by (the lags of) r common observable factors ($B_0\, y_{t-\ell}$, the "indexes"), and by variable specific errors, u_t. Defining the obervable factors as

$$F_t \equiv B_0 \cdot Y_t, \tag{3}$$

the MAI model can be also written as:

$$y_t = \sum_{\ell=1}^{p} A_\ell \cdot \; F_{t-\ell} + u_t. \tag{4}$$

Interestingly, and as pointed out by CKM16, the MAI model implies a $VAR(p)$ specification for the factors:

$$\underbrace{F_t}_{r\times 1} = \sum_{\ell=1}^{p} \underbrace{C_\ell}_{r\times r} F_{t-\ell} + \underbrace{\omega_t}_{r\times 1}, \qquad \omega_t \overset{i}{\sim} \mathcal{MN}\left(\mathbf{0},\ \Xi \right), \tag{5}$$

where $C_\ell = B_0 A_\ell$, $\ell = 1, ..., p$, and $\Xi \equiv B_0 \Omega B_0'$.

Grouping (4) and (5), we obtain a factor augmented VAR (FAVAR) - type model, where however the factors are observable.

2.2 The MAI model with Stochastic Volatility (MAI-SV)

To introduce stochastic volatility (SV) in the MAI model, following Cogley and Sargent (2005), we assume that:

$$u_t = G^{-1} \Sigma_t \varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ I_n \right), \tag{6}$$

so that

$$u_t \overset{i}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{\Omega_t}_{n\times n} \right), \qquad \Omega_t = G^{-1}\Sigma_t \Sigma_t \left(G^{-1}\right)', \tag{7}$$

where G is a triangular matrix containing reduced form covariances, and $(\Sigma_t)_{t=1}^{T}$ is the history of diagonal matrices containing the stochastic volatilities:

$$\underbrace{G}_{n\times n} = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ g_1 & 1 & \ddots & \ddots & \vdots \\ g_2 & g_3 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ g_{m-n+2} & g_{m-n+3} & \dots & g_m & 1 \end{bmatrix}, \qquad \underbrace{g}_{m\times 1} \equiv \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{bmatrix}, \qquad m \equiv \frac{n\,(n-1)}{2},$$

$$\underbrace{\Sigma_t}_{n\times n} = Diag\left(\sigma_t\right), \qquad \underbrace{\sigma_t}_{n\times 1} \equiv \begin{bmatrix} \sigma_{1,t} \\ \sigma_{2,t} \\ \vdots \\ \sigma_{n,t} \end{bmatrix}.$$

Following Primiceri (2005), the law of motion for the time-varying (TV) standard deviations σ_t is defined in logarithms as:

$$\log \sigma_t = \log \sigma_{t-1} + \nu_{\sigma,t}, \qquad \nu_{\sigma,t} \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0}, \underbrace{Q_\sigma}_{n \times n} \right). \tag{8}$$

To conclude, note that the errors in the $VAR(p)$ for the observable factors F_t also present SV. Specifically, it is:

$$\omega_t \overset{i}{\sim} \mathcal{MN}\left(\mathbf{0}, \; \Xi_t \right), \quad \Xi_t \equiv B_0 \Omega_t B_0'. \tag{9}$$

The equations in (4) and (5), augmented with the error specifications in (6) and (9), are similar to a FAVAR model with SV in both the (observable) factors and idiosyncratic errors.[2]

2.3 The MAI-SV model with time-varying covariances (MAI-SVCV)

The MAI-SV model can be further generalized by allowing the off-diagonal elements contained in the G matrix to also change over time. The resulting volatility specification is as in Primiceri (2005). The error specification becomes:

$$u_t \overset{i}{\sim} \mathcal{MN}\left(\mathbf{0}, \underbrace{\Omega_t}_{n \times n} \right), \qquad \varepsilon_t \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0}, \; I_n \right), \tag{10}$$

$$\Omega_t = G_t^{-1} \Sigma_t \Sigma_t \left(G_t^{-1} \right)', \qquad u_t = G_t^{-1} \Sigma_t \varepsilon_t,$$

and, following Primiceri (2005), the law of motion for the elements g_t of the lower triangular matrix G_t is:

$$g_t = g_{t-1} + \nu_{g,t}, \qquad \nu_{g,t} \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0}, \underbrace{Q_g}_{m \times m} \right). \tag{11}$$

[2] A small factor model with common stochastic volatility in the innovations to both factors and idiosyncratic errors is presented by Marcellino et al. (2016).

When the model has a large number of variables, n, the off-diagonal elements contained in each G_t create a computational bottleneck, as $m = n(n-1)/2$. For example, in a medium-large model with 20 variables as in our empirical application, g_t becomes an $m = 190$ dimensional stochastic process. To tackle this issue, we may follow the strategy proposed by de Wind and Gambetti (2014), that is to decompose the m-dimensional process g_t as:

$$g_t = \Lambda_g \widetilde{g}_t + M_g g_0, \qquad Q_g = \underbrace{\Lambda_g}_{m\times q_g} \Lambda_g',$$

where

$$\underbrace{\Lambda_g}_{m\times q_g} = \underbrace{V_g}_{m\times q_g} \underbrace{D_g}_{q_g\times q_g}.$$

and D_g is the diagonal matrix containing the square roots of the q_g non-zero eigenvalues of Q_g, while V_g is the matrix whose columns are the associated eigenvectors (normalized to unit length).

3 A Decomposition of the Stochastic Volatility

3.1 Common and idiosyncratic volatilities

CKM16 introduce the following decomposition:[3]

$$I_n = \Omega B_0' \Xi^{-1} B_0 + B_{0\perp}' \left(B_{0\perp} \Omega^{-1} B_{0\perp}'\right)^{-1} B_{0\perp} \Omega^{-1}, \tag{12}$$

where $B_{0\perp}$ is the $(n-r)\times n$ orthogonal matrix of B_0, such that $B_0 B_{0\perp}' = \mathbf{0}_{r\times(n-r)}$.

Using a similar decomposition in the context of the MAI-SV or MAI-SVCV models, we can then split the residuals u_t in their projection over the subspace generated by the r common factors, i.e. the factors' residuals ω_t, and their projection over the subspace

[3]See also Johansen (1995) and Centoni and Cubadda (2003).

generated by the $n - r$ idiosyncratic errors ψ_t (orthogonal to ω_t):

$$y_t = \sum_{\ell=1}^{p} A_\ell \cdot F_{t-\ell} + \Omega_t B_0' \Xi_t^{-1} \underbrace{B_0 u_t}_{\omega_t} + B_{0\perp}' \left(B_{0\perp} \Omega_t^{-1} B_{0\perp}' \right)^{-1} \underbrace{B_{0\perp} \Omega_t^{-1} u_t}_{\psi_t}, \tag{13}$$

$$y_t = \sum_{\ell=1}^{p} A_\ell \cdot F_{t-\ell} + \Omega_t B_0' \Xi_t^{-1} \underbrace{\omega_t}_{\text{Common}} + B_{0\perp}' \left(B_{0\perp} \Omega_t^{-1} B_{0\perp}' \right)^{-1} \underbrace{\psi_t}_{\text{Idio}}.$$

We can then decompose the total error volatility Ω_t into the volatility of the common component and that of the idiosyncratic component:[4]

$$\Omega_t = \Omega_t^{com} + \Omega_t^{idio},$$
$$\Omega_t^{com} = \Omega_t B_0' \Xi_t^{-1} B_0 \Omega_t \ , \quad \Omega_t^{idio} = B_{0\perp}' \left(B_{0\perp} \Omega_t^{-1} B_{0\perp}' \right)^{-1} B_{0\perp}.$$

3.2 Additional decomposition of the common volatility

Starting from the residuals of the VAR for the observed factors, ω_t, we can recover structural economic shocks, ε_t^ω, using standard techniques. The simplest possibility is to apply a Choleski decomposition, which yields:

$$F_t = B_0 \sum_{\ell=1}^{p} A_\ell F_{t-\ell} + \omega_t, \qquad \omega_t \overset{i}{\sim} \mathcal{MN}\left(\mathbf{0},\ \Xi_t \right),$$
$$\Xi_t = S_t^\omega \cdot S_t^{\omega\prime}, \qquad \omega_t = S_t^\omega \cdot \varepsilon_t^\omega, \qquad \varepsilon_t^\omega \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ I_r \right).$$

[4]The complete derivation is the following:

$$\Omega_t^{com} = \Omega_t B_0' \Xi_t^{-1} \mathbb{E}\left(\omega_t \omega_t'\right) \Xi_t^{-1} B_0 \Omega_t \ = \ \Omega_t B_0' \Xi_t^{-1} \Xi_t \Xi_t^{-1} B_0 \Omega_t \ = \ \Omega_t B_0' \Xi_t^{-1} B_0 \Omega_t.$$

$$\begin{aligned}
\Omega_t^{idio} &= B_{0\perp}' \left(B_{0\perp} \Omega_t^{-1} B_{0\perp}' \right)^{-1} B_{0\perp} \Omega_t^{-1} \mathbb{E}\left(u_t u_t'\right) \Omega_t^{-1} B_{0\perp}' \left(B_{0\perp} \Omega_t^{-1} B_{0\perp}' \right)^{-1} B_{0\perp} \\
&= B_{0\perp}' \left(B_{0\perp} \Omega_t^{-1} B_{0\perp}' \right)^{-1} \left(B_{0\perp} \Omega_t^{-1} B_{0\perp}' \right) \left(B_{0\perp} \Omega_t^{-1} B_{0\perp}' \right)^{-1} B_{0\perp} \\
&= B_{0\perp}' \left(B_{0\perp} \Omega_t^{-1} B_{0\perp}' \right)^{-1} B_{0\perp}.
\end{aligned}$$

We can now insert the structural shocks ε_t^ω into the decomposition in (13), obtaining:

$$y_t = \sum_{\ell=1}^{p} A_\ell \cdot F_{t-\ell} + \sum_{j=1}^{r} \Omega_t B_0' \Xi_t^{-1} S_t^\omega \varepsilon_{t,j}^\omega \cdot e_j + B_{0\perp}' \left(B_{0\perp} \Omega_t^{-1} B_{0\perp}' \right)^{-1} \psi_t$$

where e_j is the $r \times 1$ column vector of the r-dimensional canonical base, i.e.

$$e_1 = \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}', \quad e_2 = \begin{bmatrix} 0 & 1 & \dots & 0 \end{bmatrix}', \quad \dots, \quad , e_r = \begin{bmatrix} 0 & 0 & \dots & 1 \end{bmatrix}'.$$

Hence, we can decompose the common volatility Ω_t^{com} into orthogonal components, each associated with a structural shock:

$$\Omega_t^{com} = \sum_{j=1}^{r} \Omega_t B_0' \Xi_t^{-1} \left(S_t^\omega e_j \cdot e_j' S_t^{\omega\prime} \right) \Xi_t^{-1} B_0 \Omega_t.$$

In the empirical section we will apply these formulae to decompose the estimated conditional volatilities of US macro and financial variables into their common and idiosyncratic components, further studying the contribution of structural shocks when decomposing the common volatility component. Section 4 discusses estimation of the MAI-SV model. Readers not interested in technical details can go directly to Section 5.

4 Estimation of MAI-SV and MAI-SVCV models

As it is common in Bayesian inference, we need to specify first prior distributions for the model's elements, and then an MCMC algorithm to draw from the conditional posterior distributions, as the joint posterior distribution is not analytically available.

4.1 Specification of the prior distributions

The set of prior distributions is constructed in various steps, often using a training sample $\{-T^*, \dots, -1, 0\}$.

4.1.1 Prior on B_0 for the Metropolis step

Prior knowledge for the unrestricted elements of B_0 is elicited with a Normal distribution. To define these prior distributions, let us decompose the n variables in r blocks, so to have as many blocks as factors (r).

$$\underbrace{y_t}_{n\times 1} = \begin{bmatrix} \underbrace{y_t^{1\prime}}_{1\times n_1} & \underbrace{y_t^{2\prime}}_{1\times n_2} & \vdots & \underbrace{y_t^{r\prime}}_{1\times n_r} \end{bmatrix}', \qquad n = \sum_{j=1}^{r} n_j.$$

For each $j \in \{1, \ldots, r\}$, we compute the largest eigenvalue score from the Principal components analysis, so to obtain a final set of r score series $\left(S_t^j\right)_{t\in\{1,\ldots,T\}}^{j\in\{1,\ldots,r\}}$. Once obtained the scores, we consider the following $n - r$ univariate regression models:

$$\forall j \in \{1, \ldots, r\}, \quad \forall k \in \{2, \ldots, n_j\}, \quad S_t^j = B_{0,j,k} \cdot y_{t,k}^j + u_{j,k,t}, \quad u_{j,k,t} \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_{j,k}^2\right)$$

To normalize the first element of each $B_{0,j}$, $B_{0,j,1}$ is set at 1. Defining:

$$\forall j \in \{1, \ldots, r\}, \qquad \underbrace{\widetilde{B}_{0,j}}_{1\times(n_j-1)} \equiv \begin{bmatrix} B_{0,j,2} & \ldots & B_{0,j,n_j} \end{bmatrix},$$

for each $\widetilde{B}_{0,j}$, we compute the OLS estimate and its variance.

The prior distribution for $\underbrace{B_0}_{r\times n}$ can be then centered at

$$\underbrace{B_0}_{r\times n} = \begin{bmatrix} 1 & \widetilde{B}_{0,1} & 0 & \mathbf{0}_{1\times(n_2-1)} & \ldots & 0 & \mathbf{0}_{1\times(n_r-1)} \\ 0 & \mathbf{0}_{1\times(n_1-1)} & 1 & \widetilde{B}_{0,2} & \ldots & 0 & \mathbf{0}_{1\times(n_r-1)} \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \mathbf{0}_{1\times(n_1-1)} & 0 & \mathbf{0}_{1\times(n_2-1)} & \ldots & 1 & \widetilde{B}_{0,r} \end{bmatrix}$$

and the respective variances are coming from each separate regression. Prior covariances among elements are set to zero.

4.1.2 Prior on the loadings A

Defining $A \equiv \begin{bmatrix} A_1 & \dots & A_p \end{bmatrix}$, the prior on $a = vec\,(A')$ is multivariate Normal, centered on $\mathbf{0}$, and with diagonal variance V_a resembling a Minnesota prior.

$$V_a = \begin{bmatrix} \widehat{\sigma}^2_{y,1} & 0 & \dots & 0 \\ 0 & \widehat{\sigma}^2_{y,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \widehat{\sigma}^2_{y,n} \end{bmatrix} \otimes \begin{bmatrix} \Upsilon_1 & 0 & \dots & 0 \\ 0 & \Upsilon_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \Upsilon_p \end{bmatrix},$$

$$\forall \ell \in \{1, \dots, p\}, \qquad \Upsilon_\ell = \frac{\lambda^2}{\ell^2} \cdot \begin{bmatrix} \frac{1}{\widehat{\sigma}^2_{F,1}} & 0 & \dots & 0 \\ 0 & \frac{1}{\widehat{\sigma}^2_{F,2}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \frac{1}{\widehat{\sigma}^2_{F,r}} \end{bmatrix},$$

where $\widehat{\sigma}^2_{y,j}$ and $\widehat{\sigma}^2_{F,s}$ are the residual variances of a univariate AR(1) for, respectively, each variable j and each factor s (computed using the prior mean for B_0).

4.1.3 Prior for the elements of the residual variance

In the MAI-SV case, the prior for the elements of G is a multivariate Normal distribution centered at zero, with large diagonal covariance matrix. In the MAI-SVCV case, the prior of the initial vector g_0 is a multivariate normal centered at zero with identity matrix. The prior for σ_0 is a multivariate normal, centered at $\begin{bmatrix} \widehat{\sigma}^2_{y,1} & \widehat{\sigma}^2_{y,2} & \dots & \widehat{\sigma}^2_{y,n} \end{bmatrix}'$, with identity covariance matrix, as in Primiceri (2005). Prior distributions for innovation covariance matrices Q_g and Q_σ are calibrated as in de Wind and Gambetti (2014) and Primiceri (2005).

4.2 Gibbs Sampler

This subsection describes each step of the Gibbs Sampler (GS) used to simulate from the joint posterior distribution of both parameters $\{A, B_0, G, Q_\sigma\}$ and unobservable states $(\sigma_t)_{t=1}^T$ of the MAI-SV model. Moreover, the Omori et al. (2007) procedure requires drawing the indexes of Normal components of the mixture approximating the $\log \chi_1^2$, contained in the matrix S. The GS steps are then:

1. Draw the loadings $A \,\Big|\, B_0, G, Q_\sigma, (\sigma_t)_{t=1}^T$,
2. Draw the factor weights $B_0 \,\Big|\, A, G, Q_\sigma, (\sigma_t)_{t=1}^T$,
3. Draw the off-diagonal elements in $G \,\Big|\, A, B_0, Q_\sigma, (\sigma_t)_{t=1}^T$,
4. Draw the indexes of the mixture in $S \,\Big|\, A, B_0, G, Q_\sigma, (\sigma_t)_{t=1}^T$,
5. Draw a history of volatilities $(\sigma_t)_{t=1}^T \,|S, A, B_0, G, Q_\sigma$,
6. Draw the covariance of volatilities' innovations $Q_\sigma \,\Big|\, A, B_0, G, (\sigma_t)_{t=1}^T$.

In the MAI-SVCV model, instead of drawing the time invariant off-diagonal elements in G, we need to draw the unobservable states $(g_t)_{t=1}^T$ and the covariance matrix of their innovations Q_g. The augmented GS encompasses the following steps:

1. Draw loadings $A \,\Big|\, B_0, Q_g, Q_\sigma, (\sigma_t, g_t)_{t=1}^T$,
2. Draw the factor weights $B_0 \,\Big|\, A, Q_g, Q_\sigma, (\sigma_t, g_t)_{t=1}^T$,
3. Draw the TV off-diagonal elements $(g_t)_{t=1}^T \,\Big|\, A, B_0, Q_g, Q_\sigma, (\sigma_t)_{t=1}^T$,
4. Draw the covariance of TV off-diagonal elements $Q_g \,\Big|\, A, B_0, Q_\sigma, (\sigma_t, g_t)_{t=1}^T$,
5. Draw the indexes of the mixture in $S \,\Big|\, A, B_0, Q_g, Q_\sigma, (\sigma_t, g_t)_{t=1}^T$,
6. Draw a history of volatilities $(\sigma_t)_{t=1}^T \,\Big|\, S, A, B_0, Q_g, Q_\sigma, (g_t)_{t=1}^T$
7. Draw the covariance of volatilities' innovations $Q_\sigma \,\Big|\, A, B_0, Q_g, (\sigma_t, g_t)_{t=1}^T$.

The Gibbs Sampler steps for the MAI-SV and the MAI-SVCV are described in detail in in section A of the Appendix.

5 Empirical Results

5.1 Data and model specification

Using an updated version of the dataset in Bańbura et al. (2010) and CKM16, for comparison, the MAI-SVCV model is estimated using 20 monthly time series, covering the period from January 1964 to December 2016. All data are referred to the US economy and can be downloaded from the FRED database. The first 84 observations, from January 1964 to December 1970, are used as training sample to calibrate the prior distributions, while the rest of the sample is used in estimation. The variables are used in differences, standardized and demeaned. Figure 1 reports the transformed data used for estimation.

The number of lags ($p = 13$) and shrinkage parameters, as well as the number ($r = 3$) and factor composition (structure of B_0), follow CKM16. In particular, table 1 lists the variable and their factor grouping. The first factor is a real factor, gathering information from the real economic activity variables; the second factor is a nominal factor, representing changes in price dynamics; the third factor is a monetary/financial factor.

Carriero et al. (2016a) estimate a BVAR with independent stochastic volatilities for 125 variables (including macroeconomic indicators, an array of interest rates, some stock return measures, and exchange rates). A factor analysis of the volatilities indicates two components to account for the vast majority of innovations to volatilities. Here, we use three rather than two factors, even if the number of variables is smaller, to get robustness to potential omission of a third factor, as the third factor is significant in the conditional means, and the three factors can be given a meaningful economic interpretation as real, nominal and financial factors. Computing the first three principal components of the median volatility of residuals, we are able to explain more than 90% of time variation across the sample, in line with Carriero et al. (2016a).

5.2 Identifying the structural shocks and their effects

A Markov Chain of 10 thousand draws (with additional 30% initial draws discarded as burn-in) is simulated to obtain the posterior distribution of the MAI-SVCV parameters and unobservable states, implementing the estimation procedure described in Section 4.2. Convergence statistics confirm reliability of the resulting estimates.

We now use the estimated MAI-SVCV to identify demand, supply and monetary/financial shocks, and their effects on the variables under analysis. In the next subsection we will instead study the estimated stochastic volatilities, and assess the contributions of individual structural shocks' volatility.

The MAI-SVCV model can be used to perform structural analysis as in FAVARs applications. Indeed, as shown before, the factors' law of motion is:

$$F_t = B_0 \sum_{\ell=1}^{p} A_\ell F_{t-\ell} + \omega_t, \qquad \omega_t \overset{i}{\sim} \mathcal{MN}\,(\,\mathbf{0},\ \Xi_t\,)\,, \quad \Xi_t \equiv B_0 \Omega_t B_0',$$

where the factors are ordered as real, nominal and monetary / financial.

The peculiarity of this model comes from the observability of factors ($F_t = B_0 y_t$) and the heteroskedasticity of factor innovations, ω_t. Operating via a Choleski decomposition of the covariance matrix Ξ_t, we identify structural demand, supply and monetary/financial shocks. The estimated factors are reported in Figure 2, along with NBER recessions (as shaded areas) and the uncertainty events highlighted in Bloom (2009) (as black lines). Due to the definition of factors and estimated coefficients, improved economic conditions are associated with higher demand factor and lower supply and monetary/financial factors. The factors are reaching extreme values during the Oil crisis in the 1970s and in correspondence of the Great Financial Crisis, while from the Great moderation to the early 2000s they appear less volatile, even in case of recessions and rare events. It's interesting to notice how the demand and supply factors seem negatively correlated in case of supply-driven recessions, as the Oil crises, and positively correlated in other recessions as the last global Crisis.

Using the simulated posterior distribution, the estimated volatilities of the identified structural shocks and their posterior bands are reported in Figure 3, along with NBER recessions and Bloom (2009) uncertainty events. The Great Moderation is evident for all shocks, but the volatility of supply and financial shocks seems to increase again from the mid-90s, and all volatilities peak during the recent Global Financial Crisis.

We can also compare these volatilities with the uncertainty estimates of Carriero et al. (2017). Our demand factor volatility appears much correlated with their macro uncertainty estimate, even though the only large volatility spike after 1970s is observed in correspondence of the last Great Financial Crisis, and our measure of demand factor volatility reaches record lows in the end of 2016. The supply factor volatility shows a smaller degree of time variation than the demand factor during the Great Moderation, but it has been changing more since the spike of 2008, probably because of large fluctuations in commodity prices in the last years. As for the financial/monetary factor volatility, the largest spike throughout the sample is reasonably observed in coincidence of the years 2008-2009, differently from Carriero et al. (2017) financial uncertainty estimate, for which there is a comparable spike also in the early 2000s. It's interesting to observe how volatility of the monetary/financial factor is increasing in the years 2013-2016, most likely because of increased volatilities of bond yields and exchange rates.

To better understand the identified shocks and their effects, Figure 4 reports the dynamic response of each factor after each structural shock (posterior median), measured at the NBER troughs: a (positive) demand shock boosts the output factor, the financial factor and, to a lesser extent, the price factor; a (negative) supply shock increases prices, reduces the real output factor, and overall has non significant effects on the financial factor; finally, a (negative) financial shock lowers the real output factor and has non-significant effects on the price factor. The temporal heterogeneity in the responses is due to different shocks' sizes (i.e. their standard deviations) and also to the time-varying simultaneous relations across variables. The price factor displays the larger degree of time variation in median posterior responses, especially at larger horizon, even though these differences among NBER troughs are not significant when estimation uncertainty is taken into account.

Using the factor representation, it is now possible to compute the response of variables y_t to the structural shocks, again as in FAVAR models. Figures 5, 6 and 7 report the responses of all variables to, respectively, the identified demand, supply and financial shocks, at the NBER troughs. The responses are broadly in line with those in CKM16: a demand shock increases output variables and decreases unemployment, and triggers an increase of both prices and Fed Funds rate, along with an appreciation of the real exchange rate and a reduction in monetary indicators. The negative supply shock increases prices and worsens the real economy indicators. Interestingly, the response of the Fed Funds rate has an opposite sign in the last two recessions with respect to the previous recessions. Instead, in almost all cases there is a depreciation of the real exchange rate. Responses to the financial/monetary shock are also in line with economic theory: real variables decrease along with monetary aggregates, while the effects on prices are not significant. In terms of time variation, there are some differences in the responses at different NBER troughs, because of different shocks' size and changing simultaneous relationships. A demand shock in November 2009 has a larger effect on Industrial Production, Capacity Utilization and

Housing Starts, while a monetary/financial shock in November 2001 has a large effect on stocks and macro variables. However, when considering estimation uncertainty, such time variation is not much significant.

To provide a different measure of time variation of the responses to the structural shocks, we have also computed the temporal evolution of the responses at fixed horizons of 1, 12 and 48 months. The estimated median responses, together with 68% bands, reported in Figures 11 to 19 in section C of the Appendix, show responses broadly in line with those commented above.

5.3 Decomposing the Stochastic Volatilities

Figure 8 reports the stochastic volatilities for the 20 variables, estimated using the MAI-SVCV model. Stochastic volatilities appear to be significantly changing over time for all variables, highlighting the importance of allowing for a time-varying conditional variance in macro-financial applications. The Great Moderation is evident in variables such as employment, earnings, consumption, CPI and the PCE deflator. The recent financial crisis is associated with volatility peaks in many real and nominal variables, including employment, earning, consumption, industrial production, housing starts, CPI, PPI. However, often these peaks are global rather than local, with volatility reaching the highest (or close to highest) levels over the entire sample for personal income, industrial production, unemployment, capacity utilization, housing starts, CPI, and PPI industrials. About the monetary/financial variables, volatility peaks during the last global crisis are particularly evident for the Fed Funds rate, stock prices, money measures and reserves.

Using the posterior distribution of the volatilities from the MAI-SVCV model, and applying the decomposition in Section 3, we split each of the reduced form volatilities discussed above into the common component, driven by innovations to the three factors, and the idiosyncratic components, driven by the idiosyncratic innovations, ψ_t.

Figure 8 plots also the decomposition of the total volatility for each variable into the two orthogonal components. From a graphical inspection, it seems that the common component of volatility is dominant for most variables, though the idiosyncratic part cannot be neglected.

To have a more precise measure of how much volatility is explained by each component, Figure 9 reports the time-varying percentage shares of explained volatility by the common and idiosyncratic components. Starting from the real variables, the common component of volatility seems to have diminished its importance over time for employment and, less so, for earnings, personal income, and consumption, with values well below 50% at the end of the sample. The fraction of volatility explained by the common component is much higher for industrial production and capacity utilization, but it drops substantially for both variables after the financial crisis, from values around 80-90% to slightly above 50%. For the unemployment rate and housing starts, the fraction is instead stable at about 50% over the entire sample.

Moving to nominal variables, the fraction of their conditional time-varying volatility explained by the common shocks is rather stable and high, much more so for CPI and the PCE deflator than for PPI, but still well above 50% for the latter. A similar pattern emerges also for the financial variables, with values around 50% for the Fed Funds rate and the 10 year T-Bond yield but, interestingly, much higher for the S&P500, whose time-varying volatility seems to be substantially affected by economic shocks.

Next, we further decompose the common component share of volatility into three orthogonal sub-components, each driven by the volatility of structural demand, supply and monetary/financial shocks identified in the previous subsection. Figure 10 reports the time-varying percentage contribution of each structural shock in determining the common component volatility. Starting again with the real variables, demand shocks are clearly dominant for industrial production and capacity utilization. The same was true for employment at the beginning of the sample, but then the role of demand shocks has progressively declined in favor of supply and monetary/financial shocks, that combined explained more than 50% of the common volatility component of employment at the end of the sample. A declining role for demand shocks is evident, though milder, also for housing starts. Instead, for earning, income, consumption and unemployment the share of demand shocks is rather stable over time, below 50% for the three variables and close to 50% for unemployment. Supply and monetary/financial shocks explain a comparable fraction of the remaning common volatility component of these four variables.

Regarding price variables, supply shocks are dominant, in particular for CPI and the PCE deflator, with demand shocks ranked second and monetary/financial shocks third, though slightly more important for PPI.

For the financial variables, the contribution of the three types of structural shocks is overall comparable and rather stable over time. The volatility of the monetary / financial shock is particularly important to explain the common component of the S&P500 volatility, but even for this variable it only explains about 50%, with the other 50% close to equally split between demand and supply shocks.

From an economic point of view, the volatility of the structural demand shock could be related to changes in the inventory mechanism (e.g., McConnell and Perez-Quiros, 2000) or to the globalization process (e.g., Bianchi and Civelli, 2015), while that of the supply shock can depend on the behavior of the oil market (e.g., Lee et al., 1995) or on firms' productivity (e.g., Christiano et al., 2010), and the volatility of the financial/monetary shock can be influenced by financial innovation (e.g., Dynan et al., 2006) or by changes in the conduct of monetary policy (e.g., Clarida et al., 2000). Justiniano and Primiceri (2008) introduce a DSGE model with stochastic volatility in several structural shocks. Specifically, they find that a reduction in the variance of investment shocks, interpreted also as a proxy for unmodeled financial frictions, is the main driver of the US Great Moderation, with a limited role for changes in monetary policy. Their evidence is somewhat in line with the gradually decreasing role of the monetary/financial shock contribution from the 1980s to the Great Financial Crisis in our common volatility decomposition. However, our structural decomposition is based on the common part of the volatility which is only a fraction of the total volatility. Moreover, Benati and Surico (2009) show that it can be difficult to map changes in the variance of structural VAR shocks with those in DSGE models.

Overall, the structural demand, supply and financial shocks appear as important drivers not only of the level of macroeconomic and financial variables but also of their changing conditional volatility. In fact, the volatility of demand, supply and monetary / financial shocks explains, in general, more than 50% of the overall volatility, with their relative importance being variable dependent and, sometimes, changing over time.

6 Conclusions

Many economic variables feature changes in their conditional volatility, and stochastic volatility specifications are commonly used in macroeconomic applications to model this feature. In models with stochastic volatility, different shocks drive the levels and volatilities of the variables, with volatility shocks left unexplained. In this paper we decompose the shocks driving the volatility processes into the volatility of structural economic shocks, in order to understand the relative importance over time of demand, supply and monetary

/ financial shocks as drivers of volatility.

The structural shocks are obtained from a Multivariate Autoregressive Index (MAI) model, a particular reduced rank VAR that can be also interpreted as a factor model, featuring stochastic volatility, over a large set of real, nominal and financial indicators.

We have developed a Gibbs Sampling algorithm for (Bayesian) inference, introducing efficient strategies to reduce the computational burden.

Using the model estimated with US data for the period 1964-2016, we have found that the common component of volatility is substantial, it explains at least 50% of the overall volatility for most variables, though the share is declining over time for some real variables. Moreover, a large fraction of the common volatility component is driven by the volatility of structural demand, supply and financial shocks, in general more than 50%, with their relative importance being variable dependent and, sometimes, changing over time.

Bibliography

Abbate, A., Eickmeier, S., Lemke, W., and Marcellino, M. (2016). The changing international transmission of financial shocks: Evidence from a classical time-varying favar. *Journal of Money, Credit and Banking*, 48(4):573–601.

Ataei, A. (2014). Improved qrginv algorithm for computing moore-penrose inverse matrices. *ISRN Applied Mathematics*.

Bańbura, M., Giannone, D., and Reichlin, L. (2010). Large bayesian vector auto regressions. *Journal of Applied Econometrics*, 25(1):71–92.

Benati, L. and Surico, P. (2009). Var analysis and the great moderation. *The American Economic Review*, 99(4):1636–1652.

Bernanke, B. S., Boivin, J., and Eliasz, P. (2005). Measuring the effects of monetary policy: a factor-augmented vector autoregressive (favar) approach. *The Quarterly journal of economics*, 120(1):387–422.

Bianchi, F. and Civelli, A. (2015). Globalization and inflation: Evidence from a time-varying var. *Review of Economic Dynamics*, 18(2):406–433.

Bloom, N. (2009). The impact of uncertainty shocks. *econometrica*, 77(3):623–685.

Bodnar, T. and Okhrin, Y. (2008). Properties of the singular, inverse and generalized inverse partitioned wishart distributions. *Journal of Multivariate Analysis*, 99(10):2389–2405.

Carriero, A., Clark, T. E., and Marcellino, M. (2015). Bayesian vars: specification choices and forecast accuracy. *Journal of Applied Econometrics*, 30(1):46–73.

Carriero, A., Clark, T. E., and Marcellino, M. (2016a). Large vector autoregressions with stochastic volatility and flexible priors. *Federal Reserve Bank of Cleveland Working Paper, no. 16-17*.

Carriero, A., Clark, T. E., and Marcellino, M. (2017). Measuring uncertainty and its impact on the economy. *The Review of Economics and Statistics*.

Carriero, A., Kapetanios, G., and Marcellino, M. (2011). Forecasting large datasets with bayesian reduced rank multivariate models. *Journal of Applied Econometrics*, 26(5):735–761.

Carriero, A., Kapetanios, G., and Marcellino, M. (2016b). Structural analysis with multivariate autoregressive index models. *Journal of Econometrics*, 192(2):332–348.

Carter, C. K. and Kohn, R. (1994). On Gibbs Sampling for State Space Models. *Biometrika*, 81(3):541–553.

Centoni, M. and Cubadda, G. (2003). Measuring the business cycle effects of permanent and transitory shocks in cointegrated time series. *Economics Letters*, 80(1):45–51.

Christiano, L. J., Motto, R., and Rostagno, M. (2010). Financial factors in economic fluctuations. *ECB Working Paper*, (No. 1192).

Clarida, R., Gali, J., and Gertler, M. (2000). Monetary policy rules and macroeconomic stability: evidence and some theory. *The Quarterly journal of economics*, 115(1):147–180.

Clark, T. E. and Ravazzolo, F. (2015). Macroeconomic forecasting performance under alternative specifications of time-varying volatility. *Journal of Applied Econometrics*, 30(4):551–575.

Cogley, T. and Sargent, T. J. (2005). Drifts and volatilities: monetary policies and outcomes in the post wwii us. *Review of Economic dynamics*, 8(2):262–302.

Corsello, F. (2018). Rank Reduction in BVARs with Time Varying Coefficients and Stochastic Volatility: Specification and Estimation. *Bocconi University, mimeo.*

D'Agostino, A., Gambetti, L., and Giannone, D. (2013). Macroeconomic forecasting and structural change. *Journal of Applied Econometrics*, 28(1):82–101.

de Wind, J. and Gambetti, L. (2014). Reduced-Rank Time-Varying Vector Autoregressions. *CPB Discussion Paper*, (270).

Del Negro, M. and Primiceri, G. E. (2015). Time varying structural vector autoregressions and monetary policy: a corrigendum. *The review of economic studies*, 82(4):1342–1345.

Dıaz-Garcıa, J. A., Jáimez, R. G., and Mardia, K. V. (1997). Wishart and pseudo-wishart distributions and some applications to shape theory. *Journal of Multivariate Analysis*, 63(1):73–87.

Dynan, K. E., Elmendorf, D. W., and Sichel, D. E. (2006). Can financial innovation help to explain the reduced volatility of economic activity? *Journal of monetary Economics*, 53(1):123–150.

Eickmeier, S., Lemke, W., and Marcellino, M. (2015). Classical time varying factor-augmented vector auto-regressive models—estimation, forecasting and structural analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(3):493–533.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007.

Fernández-Villaverde, J. and Rubio-Ramırez, J. (2013). Macroeconomics and volatility: Data, models, and estimation. econometrics. *Cambridge University Press. Ferreira da Silva, G.(2002). The impact of financial system development on business cycles volatility: Cross-country evidence. Journal of Macroeconomics*, 24(2):233–253.

Fuller, W. A. (2009). *Introduction to statistical time series*, volume 428. John Wiley & Sons.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2014). *Bayesian data analysis*, volume 2. CRC press Boca Raton, FL.

Geweke, J. (1996). Bayesian reduced rank regression in econometrics. *Journal of econometrics*, 75(1):121–146.

Johansen, S. (1995). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press on Demand.

Jurado, K., Ludvigson, S. C., and Ng, S. (2015). Measuring uncertainty. *The American Economic Review*, 105(3):1177–1216.

Justiniano, A. and Primiceri, G. E. (2008). The Time-Varying Volatility of Macroeconomic Fluctuations. *American Economic Review*, 98(3):604–41.

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with arch models. *The review of economic studies*, 65(3):361–393.

Lee, K., Ni, S., and Ratti, R. A. (1995). Oil shocks and the macroeconomy: the role of price variability. *The Energy Journal*, pages 39–56.

Litterman, R. B. (1986). Forecasting with bayesian vector autoregressions—five years of experience. *Journal of Business & Economic Statistics*, 4(1):25–38.

Marcellino, M., Porqueddu, M., and Venditti, F. (2016). Short-term gdp forecasting with a mixed-frequency dynamic factor model with stochastic volatility. *Journal of Business & Economic Statistics*, 34(1):118–127.

McConnell, M. M. and Perez-Quiros, G. (2000). Output fluctuations in the united states:

What has changed since the early 1980's? *American Economic Review*, 90(5):1464–1476.

Omori, Y., Chib, S., Shephard, N., and Nakajima, J. (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics*, 140(2).

Primiceri, G. E. (2005). Time Varying Structural Vector Autoregressions. *The Review of Economic Studies*, 72(3):821–852.

Reinsel, G. (1983). Some results on multivariate autoregressive index models. *Biometrika*, 70(1):145–156.

Reinsel, G. and Velu, R. (1998). Multivariate reduced-rank regression, volume 136 of lectures notes in statistics. New York: Springer-Verlag.

Sims, C. A. and Zha, T. (2006). Were there regime switches in us monetary policy? *The American Economic Review*, 96(1):54–81.

Table 1: Composition of Factors

Variable	**F1**	**F2**	**F3**
Employees (Total nonfarm)	1	0	0
Average hourly earnings	$B_0^{(1,2)}$	0	0
Personal Income	$B_0^{(1,3)}$	0	0
Real Consumption	$B_0^{(1,4)}$	0	0
Industrial Production	$B_0^{(1,5)}$	0	0
Capacity Utilization	$B_0^{(1,6)}$	0	0
Unemployment rate	$B_0^{(1,7)}$	0	0
Housing starts	$B_0^{(1,8)}$	0	0
CPI all items	0	1	0
Producer Price Index (Farm & Foods)	0	$B_0^{(2,10)}$	0
Implicit price deflator for personal cons. exp.	0	$B_0^{(2,11)}$	0
Producer Price Index (Industrials)	0	$B_0^{(2,12)}$	0
Federal Funds, effective	0	0	1
M1 money stock	0	0	$B_0^{(3,14)}$
M2 money stock	0	0	$B_0^{(3,15)}$
Total reserves of depository institutions	0	0	$B_0^{(3,16)}$
Nonborrowed reserves of depository institutions	0	0	$B_0^{(3,17)}$
S&P's common stock price index	0	0	$B_0^{(3,18)}$
Interest rate on treasury bills, 10 year constant maturity	0	0	$B_0^{(3,19)}$
Effective Exchange rate	0	0	$B_0^{(3,20)}$

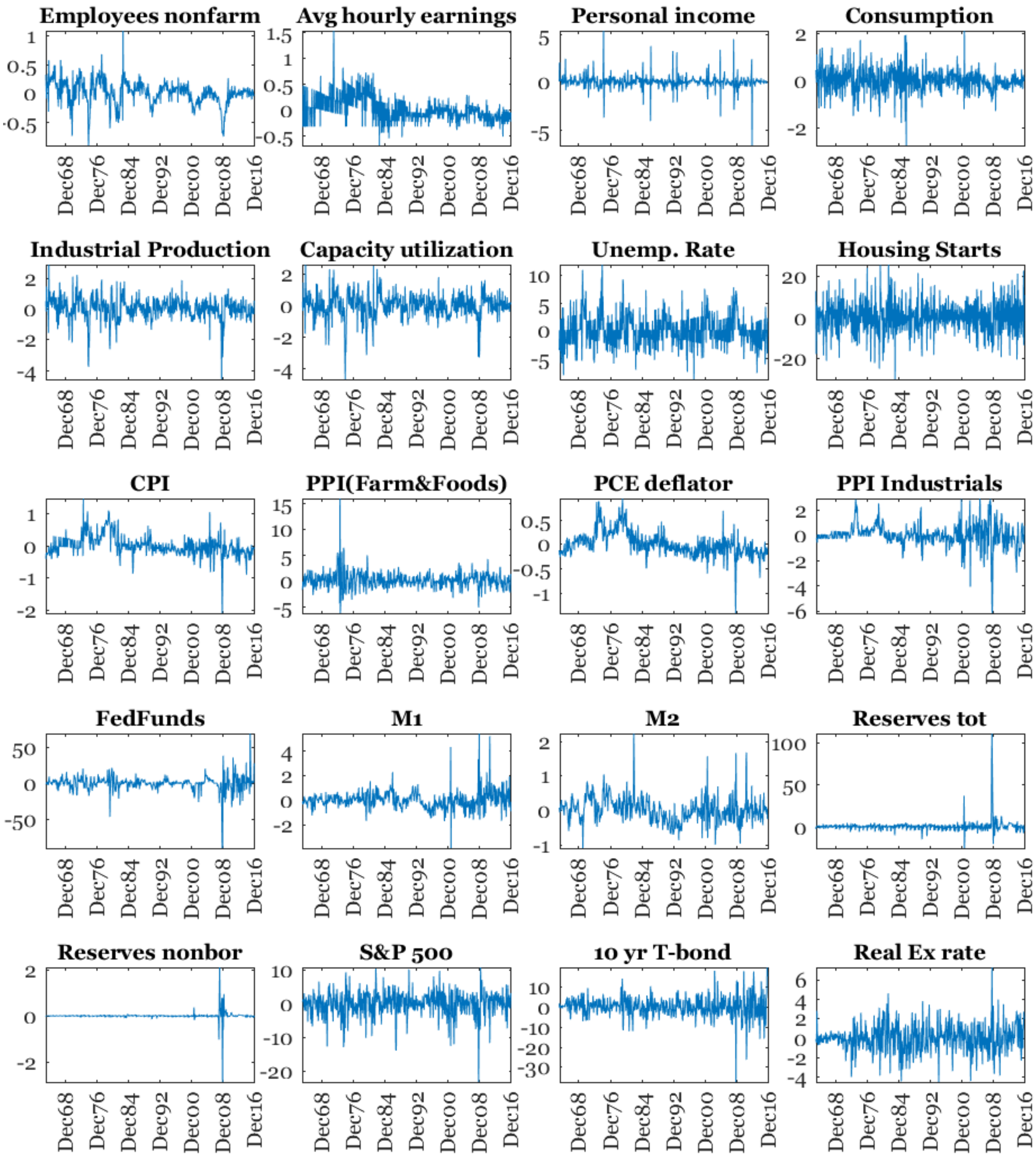


Figure 1: Transformed variables

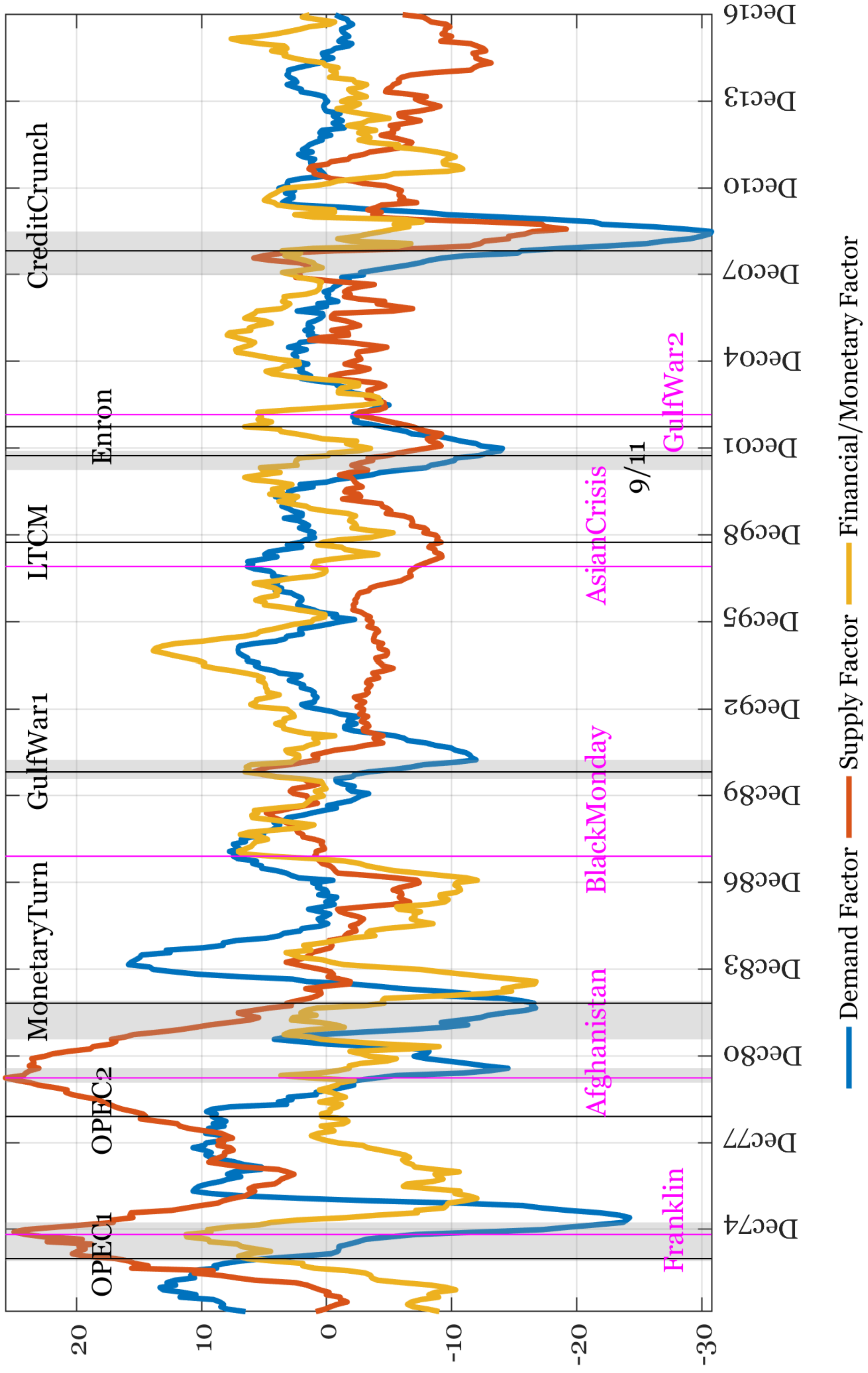


Figure 2: Estimated observable factors, with NBER recessions (shaded areas) and uncertainty events (black vertical lines).

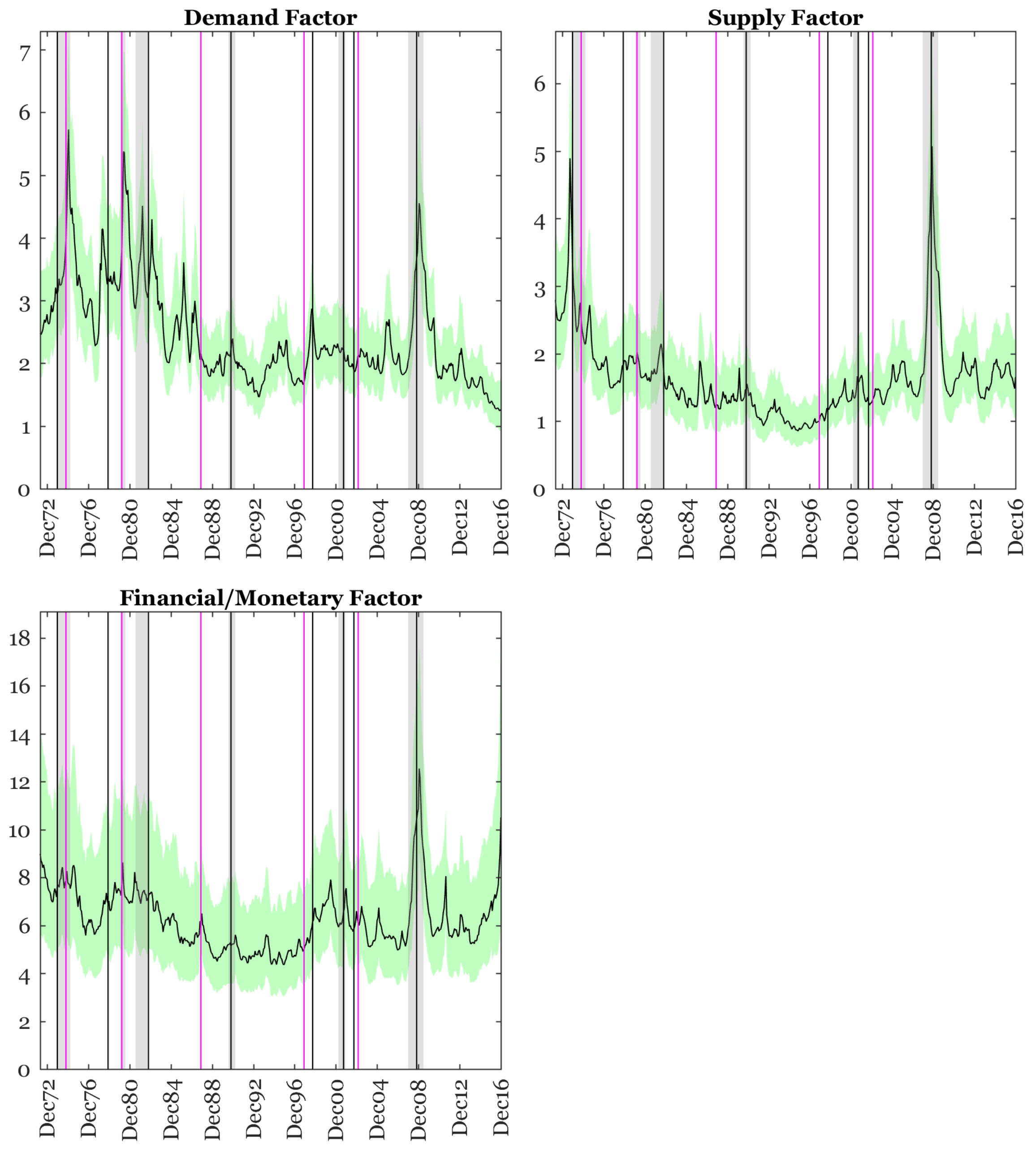


Figure 3: Identified structural shocks volatilities. Green Bands correspond to the 68% credible regions of the posterior distribution.

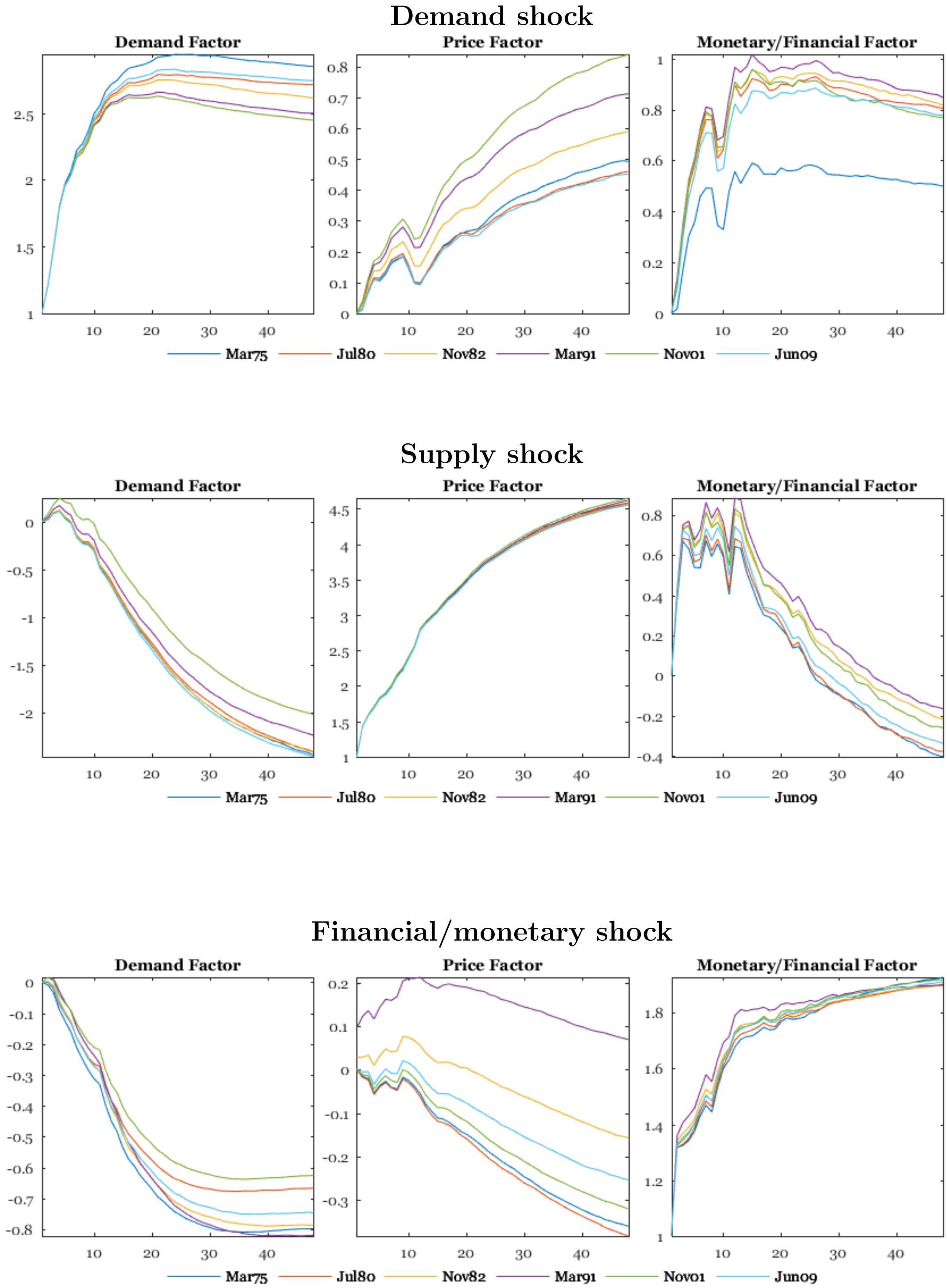


Figure 4: Response of each factor to all shocks at NBER troughs. Mean posterior distribution.

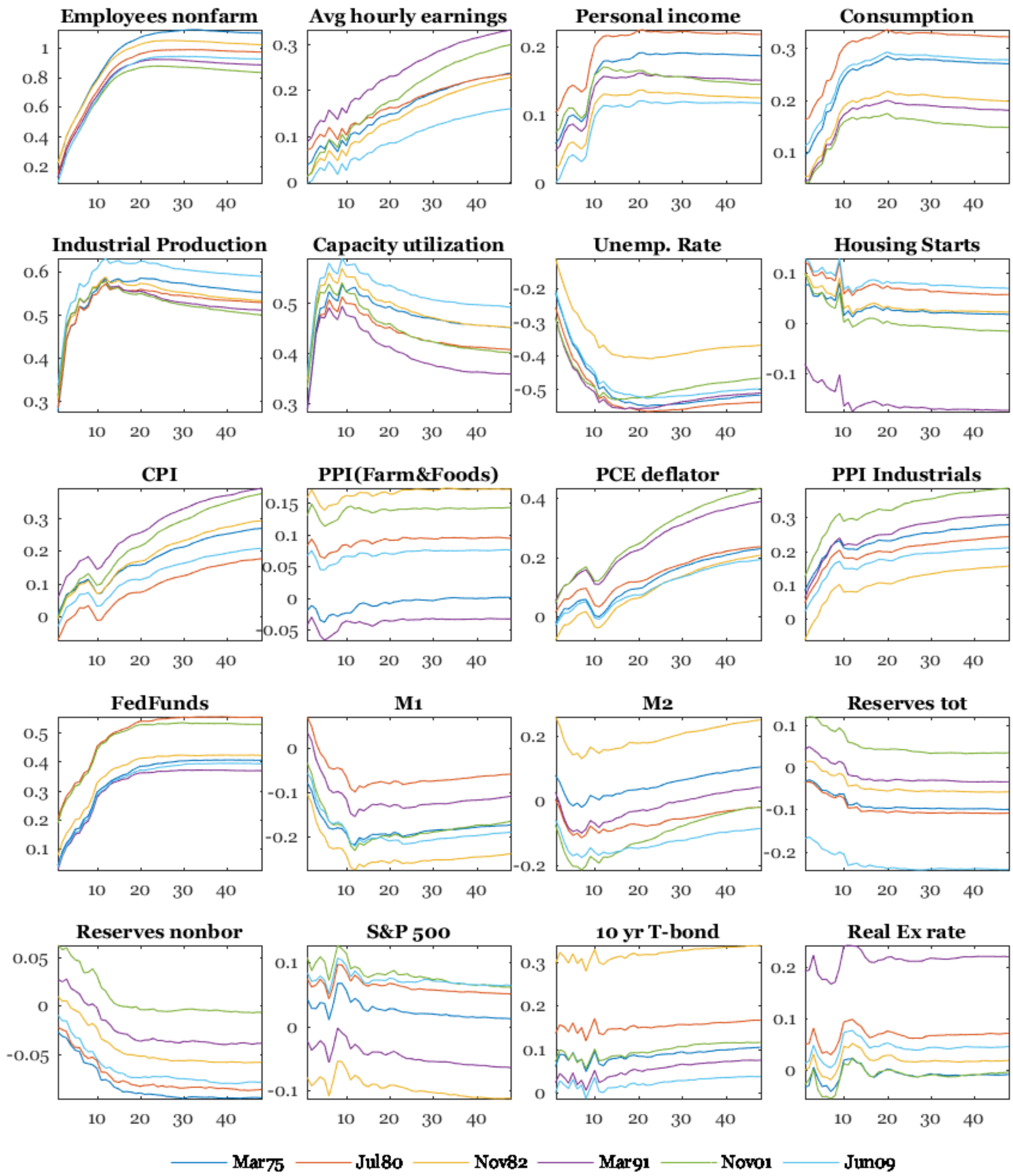


Figure 5: Response to a permanent Demand shock at NBER troughs. Mean posterior distribution.

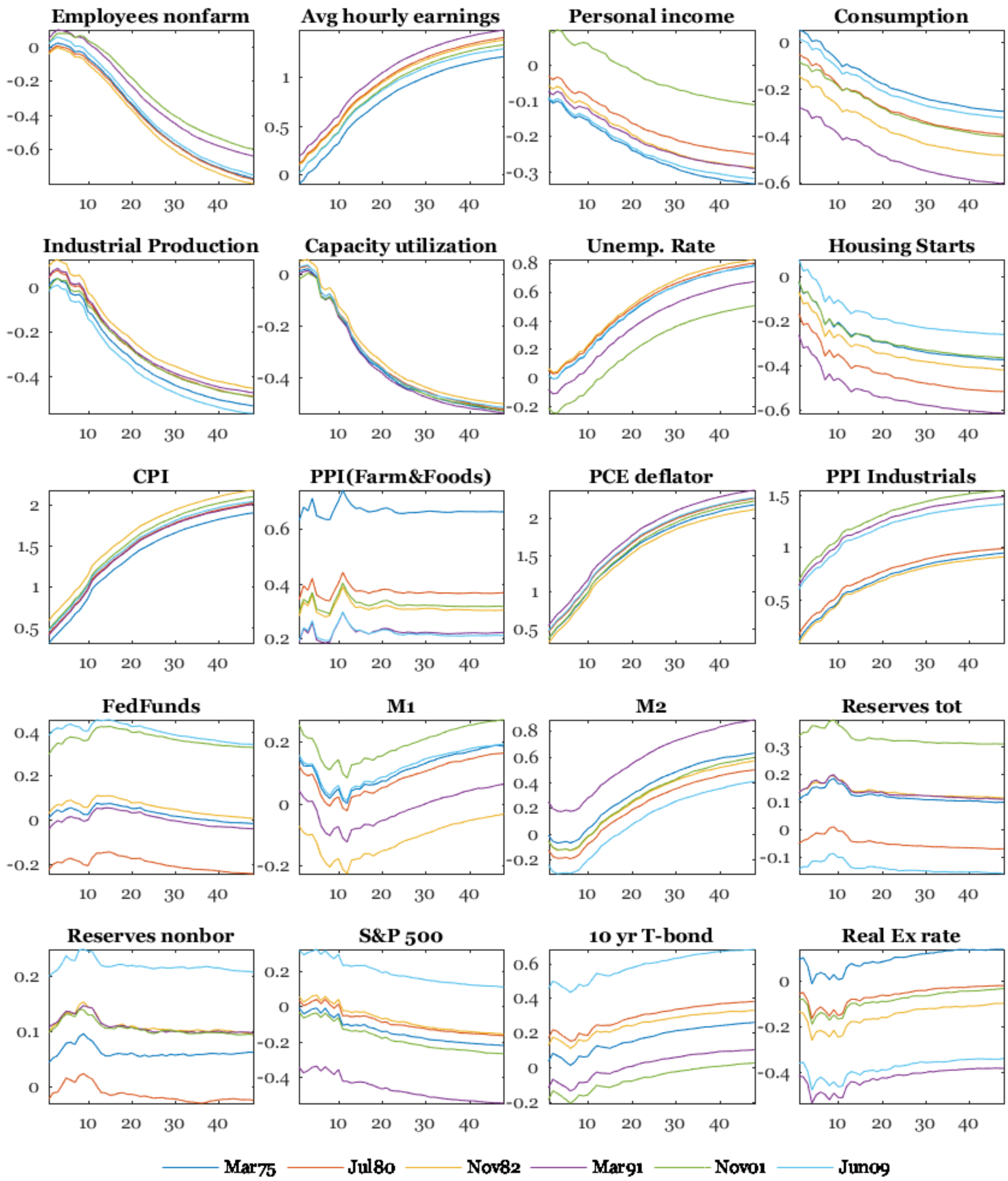


Figure 6: Response to a permanent Supply shock at NBER troughs. Mean posterior distribution.

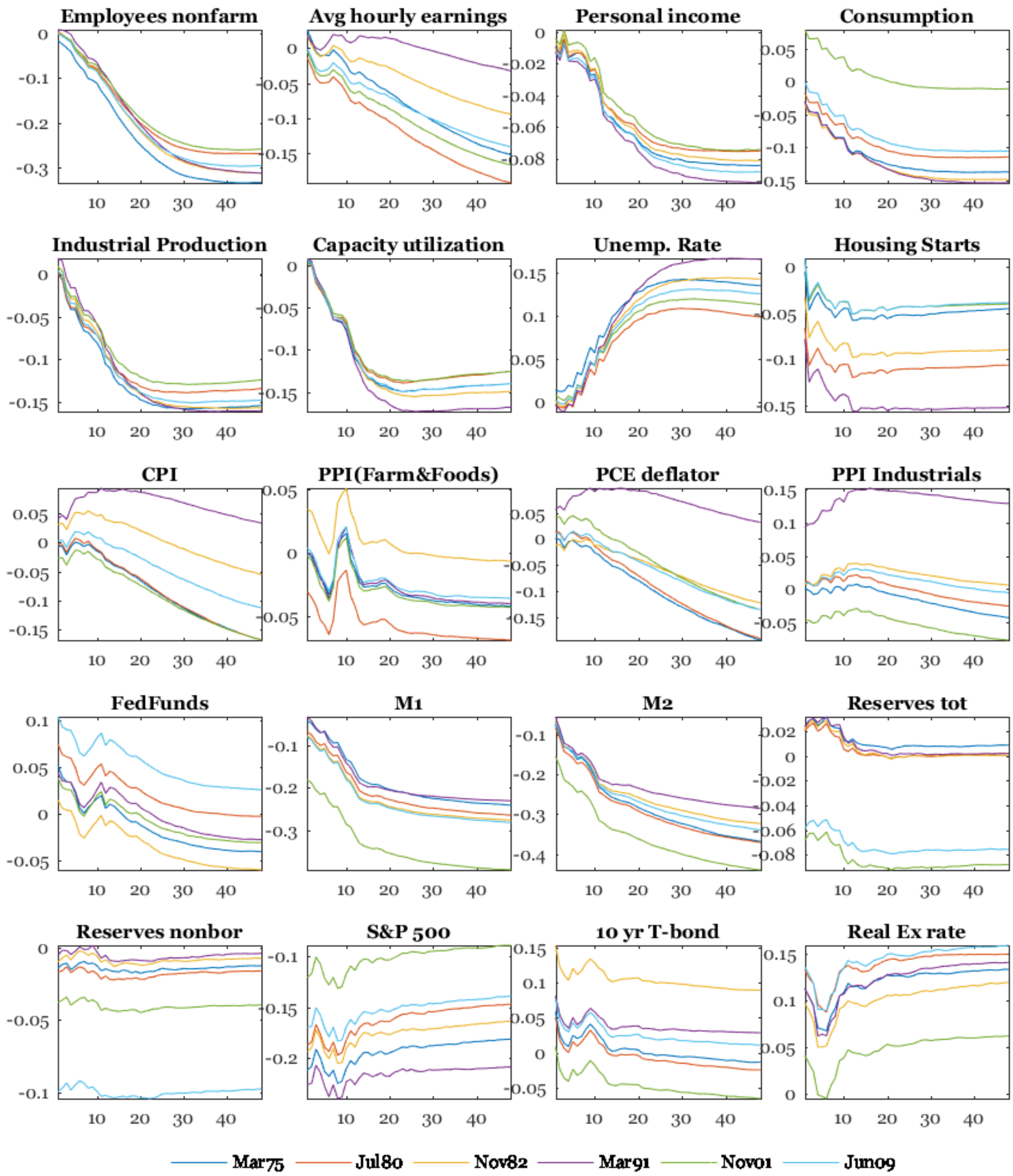


Figure 7: Response to a permanent Financial shock at NBER troughs. Mean posterior distribution.

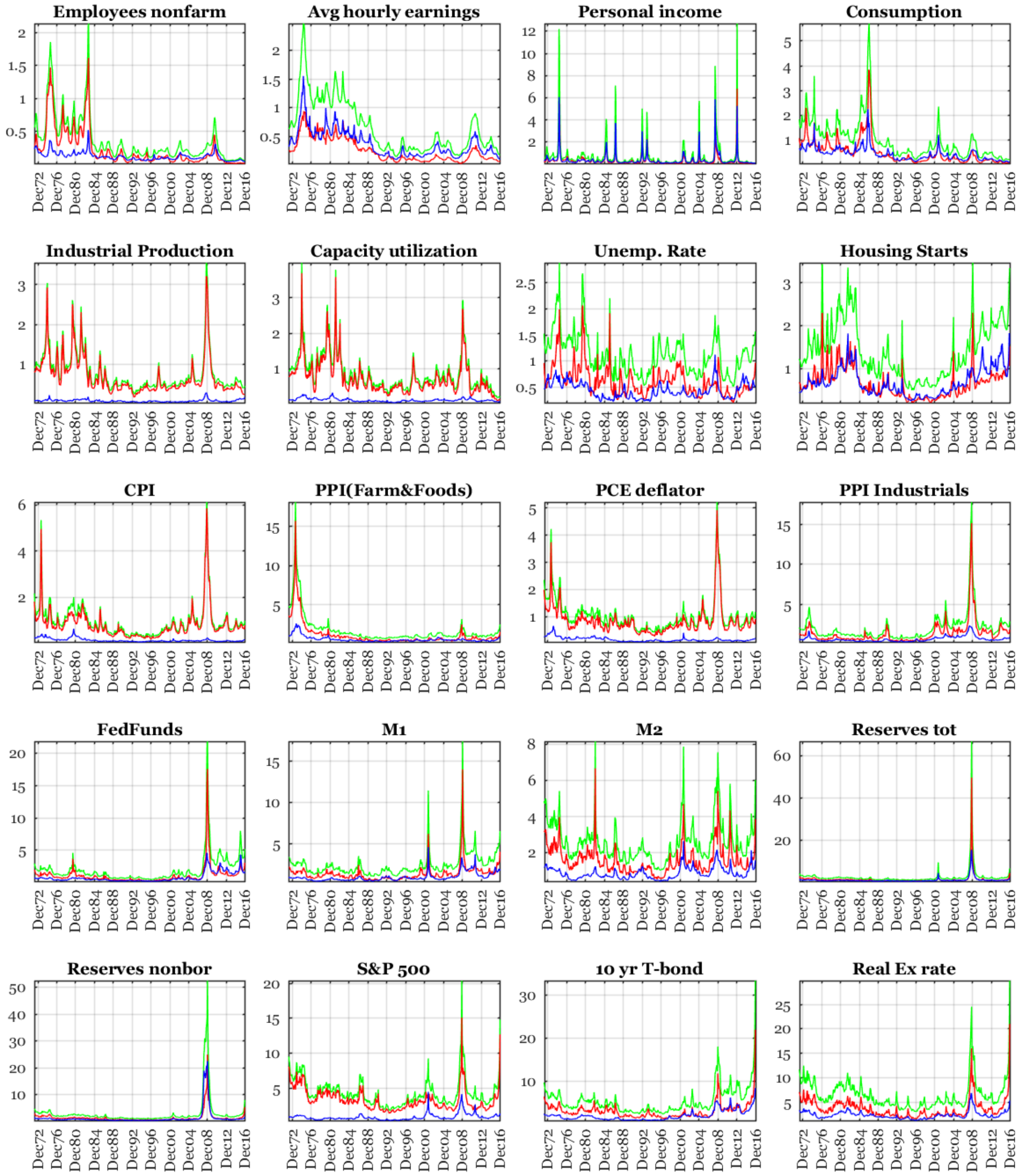


Figure 8: Residual volatility and its decomposition: total (green), common factors (red), Idiosyncratic (blue)

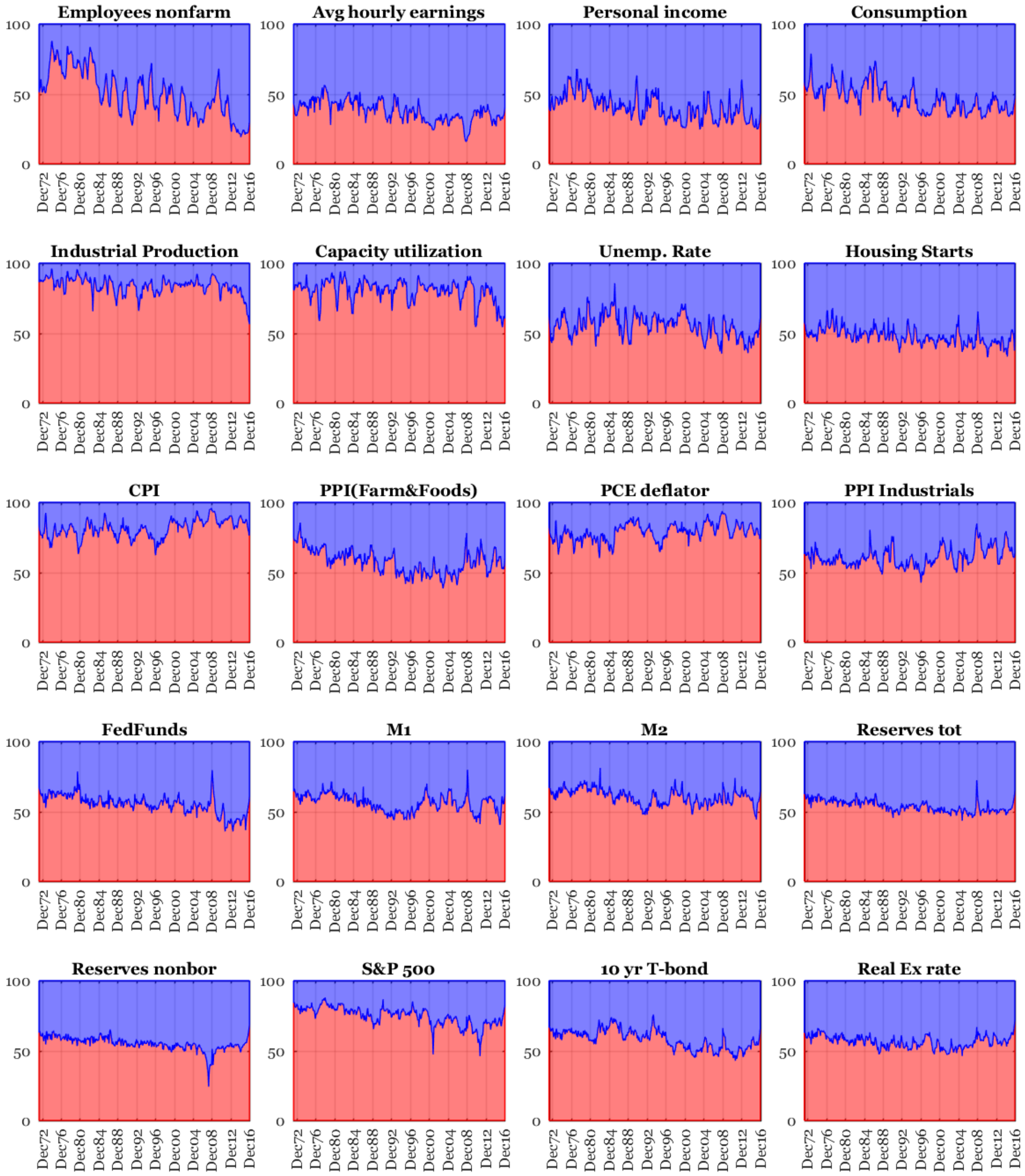


Figure 9: Residual volatility shares (%): common factors (red), Idiosyncratic (blue)

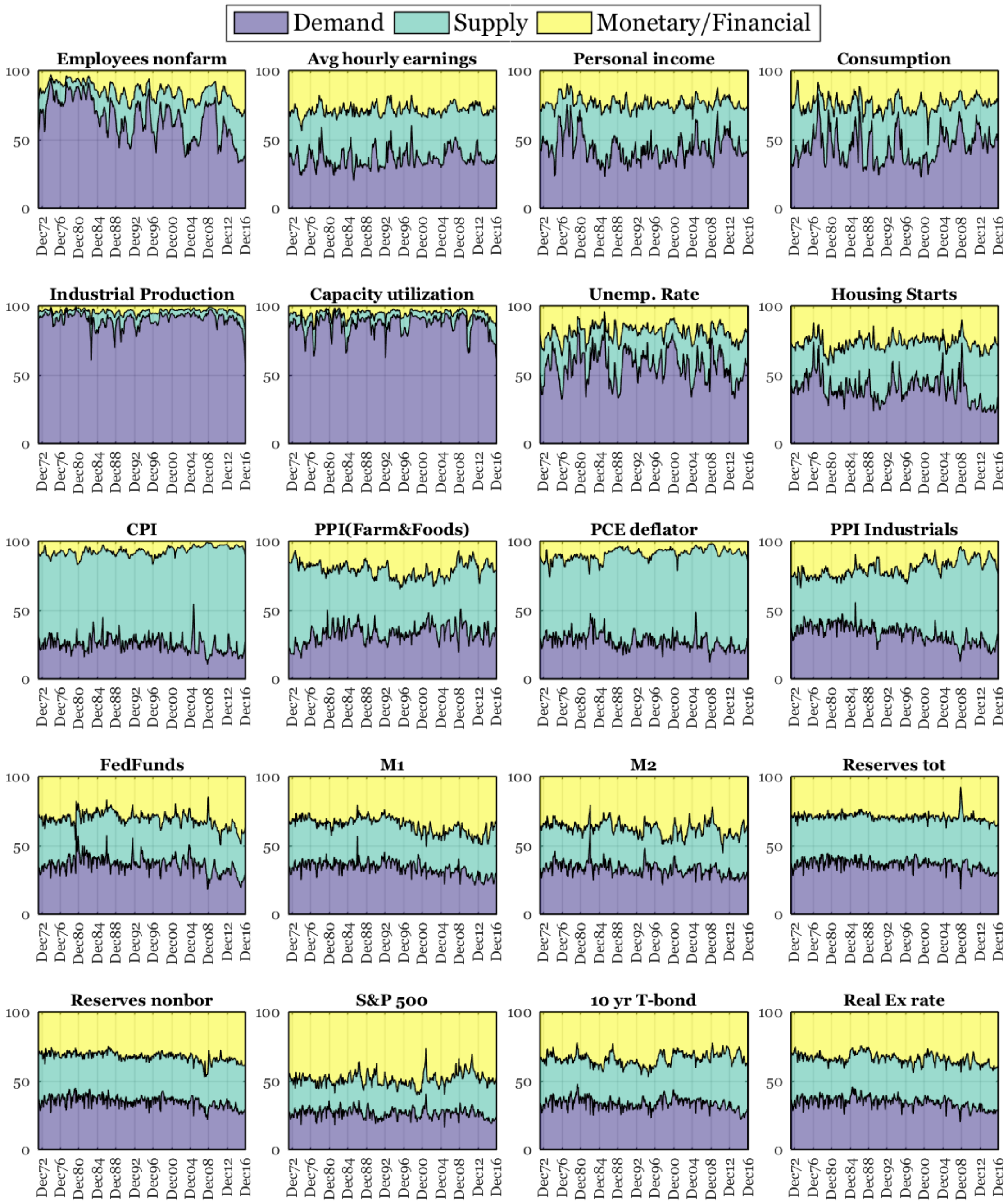


Figure 10: Contribution (%) of the identified shocks over the common factor component of residual volatility

A Gibbs Sampler for estimation of the MAI-SV and MAI-SVCV models

A.1 Step 1: Draw loadings A

The first step of the GS aims at drawing the loadings contained in A. Recalling that:

$$A \equiv \begin{bmatrix} A_1 & \dots & A_p \end{bmatrix}, \qquad x_t \equiv vec\left(x_t^\bullet\right), \qquad \underbrace{x_t^\bullet}_{n\times p} \equiv \begin{bmatrix} y_{t-1} & \dots & y_{t-p} \end{bmatrix},$$

$$Z_t \equiv \begin{bmatrix} B_0 y_{t-1} \\ \vdots \\ B_0 y_{t-p} \end{bmatrix} = \left(I_p \otimes B_0\right) \cdot x_t = vec\left(B_0 \cdot x_t^\bullet\right),$$

the model can be restated as:

$$y_t = \begin{bmatrix} A_1 & \dots & A_p \end{bmatrix} \begin{bmatrix} B_0 y_{t-1} \\ \vdots \\ B_0 y_{t-p} \end{bmatrix} + u_t \ = \underbrace{A}_{n\times rp} \cdot \underbrace{Z_t}_{rp\times 1} + u_t.$$

It can then be stacked as:

$$\begin{bmatrix} y_1' \\ y_2' \\ \vdots \\ y_T' \end{bmatrix} = \begin{bmatrix} Z_1' \\ Z_2' \\ \vdots \\ Z_T' \end{bmatrix} A' + \begin{bmatrix} u_1' \\ u_2' \\ \vdots \\ u_T' \end{bmatrix},$$

$$\underbrace{y}_{T\times n} = \underbrace{Z}_{T\times rp} \cdot \underbrace{A'}_{rp\times n} + u.$$

Defining $a \equiv vec\left(A'\right)$, and exploiting the Kronecker properties, the stacked form can be vectorized and transformed into:

$$vec\left(y\right) = vec\left(Z \cdot A' \cdot I_n\right) + vec\left(u\right),$$

$$\underbrace{Y}_{nT\times 1} = \underbrace{\left(I_n \otimes Z\right)}_{n\times nrp} \cdot \underbrace{a}_{nrp\times 1} + U,$$

where U has the following distribution:

$$\underbrace{U}_{nT\times 1} \sim \mathcal{MN}\left(\mathbf{0},\ \underbrace{V_u}_{nT\times nT} \right)$$

and

$$V_u \equiv \begin{bmatrix} \Omega_1^{(1,1)} & 0 & \cdots & 0 & \cdots & \cdots & \Omega_1^{(1,n)} & 0 & \cdots & 0 \\ 0 & \Omega_2^{(1,1)} & \ddots & \vdots & \cdots & \cdots & 0 & \Omega_2^{(1,n)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & \cdots & \cdots & \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \Omega_T^{(1,1)} & \cdots & \cdots & 0 & \cdots & 0 & \Omega_T^{(1,n)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ \Omega_1^{(n,1)} & 0 & \cdots & 0 & \cdots & \cdots & \Omega_1^{(n,n)} & 0 & \cdots & 0 \\ 0 & \Omega_2^{(n,1)} & \ddots & \vdots & \cdots & \cdots & 0 & \Omega_2^{(n,n)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & \cdots & \cdots & \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \Omega_T^{(n,1)} & \cdots & \cdots & 0 & \cdots & 0 & \Omega_T^{(nn)} \end{bmatrix}$$
$$= \sum_{t=1}^{T} \left[\Omega_t \otimes (e_t \cdot e_t') \right].$$

To use an informative prior on a, we follow the approach in Gelman et al. (2014). The strategy incorporates the prior as observations. Considering a multivariate normal prior with the following moments:

$$a \sim \mathcal{MN}\left(\ \bar{a},\ V_a\ \right),$$

it is possible to augment the model with nrp observations that express the prior information:

$$\begin{bmatrix} Y \\ \bar{a} \end{bmatrix} = \begin{bmatrix} I_n \otimes Z \\ I_{nrp} \end{bmatrix} a + \begin{bmatrix} U \\ U_a \end{bmatrix},$$
$$Y^\diamond = Z^\diamond a + U^\diamond, \quad U^\diamond \sim \mathcal{MN}\left(\ \mathbf{0}_{nT+nrp},\ V^\diamond\ \right),$$
$$V^\diamond = \begin{bmatrix} V_u & \mathbf{0}_{nT\times nrp} \\ \mathbf{0}_{nrp\times nT} & V_a \end{bmatrix}.$$

A draw for a then comes from the following posterior:

$$a \sim \mathcal{MN}\left(\, \widetilde{a},\ \left(Z^{\diamond\prime} V^{\diamond -1} Z^{\diamond}\right)^{-1} \,\right),$$

$$\widetilde{a} = \left(Z^{\diamond\prime} V^{\diamond -1} Z^{\diamond}\right)^{-1} Z^{\diamond\prime} V^{\diamond -1} Y^{\diamond}.$$

In order to decrease the computational burden of this step throughout the sampling, we adopt the strategy proposed by Carriero et al. (2016a). In particular, the triangular representation of the system is exploited, and coefficients are drawn equation by equation. The approach proposed in Carriero et al. (2016a) and its generalization to allow for time-varying covariances is analytically documented in the appendix (section B.2).

A.2 Step 2: Draw the (factor weights) elements in B_0

Given the restrictions and the non-linear role of B_0, a Random Walk Metropolis step on the posterior kernel of each element of B_0 is implemented, nested into the Gibbs Sampling algorithm. In order to do this, we first write the likelihood of the model.

Given the reduced form VAR written as:

$$y_t = A \cdot Z_t + u_t\ , \qquad u_t \overset{i}{\sim} \mathcal{MN}\left(\,\mathbf{0},\ \Omega_t\,\right),$$

conditioning on all other elements and using the chain rule, we can write the likelihood kernel as:

$$f\left((y_t)_{t=1}^{T} \middle|\, A, (\Omega_t)_{t=1}^{T}\,,\ B_0 \,\right) \propto \left(\prod_{t=1}^{T} |\Omega_t|^{-\frac{1}{2}} \right) \exp\left\{ -\frac{1}{2} \sum_{t=1}^{T} \left(y_t - A\cdot Z_t\right)' \Omega_t^{-1} \left(y_t - A \cdot Z_t\right) \right\}.$$

Next, let us consider the $r^* \equiv n - r$ scalar unrestricted elements of B_0, i.e. $\left(b_{0,j}\right)_{j=1}^{r^*}$. Then, $\forall j \in \{1, \ldots, r^*\}$, we can define the set $b_{0,j-} \equiv \left(b_{0,s}\right)_{s \neq j}$.

For a given prior $f\left(b_{0,j}\right)$ on each element $b_{0,j}$, we can write the kernel of the conditional posterior of $b_{0,j}$ as:

$$f_{post}\left(b_{0,j} |\, (y_t, \Omega_t)_{t=1}^{T}\,,\ A,\ b_{0,j-} \right) \propto f\left((y_t)_{t=1}^{T} \middle|\, A,\ B_0,\ (\Omega_t)_{t=1}^{T} \right) \cdot f\left(b_{0,j}\right).$$

We are now ready to design the Metropolis step, separately for each j. Given the last step B_0^{i-1}, a random walk candidate is computed as:

$$b_{0,j}^{*} = b_{0,j}^{i-1} + c_j \cdot \eta_t,$$

where c_j is a scaling factor calibrated to have an acceptance rate of approximately 30%-35% and $\eta_t \overset{iid}{\sim} \mathcal{N}(0, v_j)$, where v_j is the variance of the prior $f(b_{0,j})$. The candidate draw is accepted with probability:

$$\alpha_j = \min \left\{ 1, \frac{f_{post}\left(b_{0,j}^{*} \middle| \left(y_t, \Omega_t^{i-1}\right)_{t=1}^{T}, \; A, \; b_{0,j-}^{i-1} \right)}{f_{post}\left(b_{0,j}^{i-1} \middle| \left(y_t, \Omega_t^{i-1}\right)_{t=1}^{T}, \; A, \; b_{0,j-}^{i-1} \right)} \right\}.$$

When the candidate is accepted, then $b_{0,j-}^{i} = b_{0,j}^{*}$, otherwise $b_{0,j-}^{i} = b_{0,j-}^{i-1}$. Repeating this procedure for $\forall j \in \{1, \ldots, r^*\}$, we obtain a draw B_0^i from the distribution of interest.

A.3 Step 3: draw the off-diagonal elements in G

To draw the off-diagonal elements in G, we restate the reduced form as:

$$y_t = \underbrace{A}_{n \times rp} \cdot \underbrace{Z_t}_{rp \times 1} + u_t,$$

$$y_t - A \cdot Z_t = G^{-1} \Sigma_t \varepsilon_t,$$

$$\widehat{y_t} = G^{-1} \Sigma_t \varepsilon_t,$$

$$G \cdot \widehat{y_t} = \Sigma_t \varepsilon_t.$$

Removing ones from the diagonal of G, and bringing off-diagonal elements on the right hand side, produces:

$$G = I_n + G^*.$$

This can be combined in the model to obtain:

$$\left(I_n + G^*\right) \; \widehat{y_t} = \Sigma_t \; \varepsilon_t,$$

$$\widehat{y_t} = -G^* \; \widehat{y_t} + \Sigma_t \; \varepsilon_t.$$

Exploiting the Kronecker properties, we then get:

$$- I_n \underbrace{G^*}_{n\times n} \underbrace{\widehat{y}_t}_{n\times 1} = -\underbrace{(I_n \otimes \widehat{y}_t')}_{n\times n^2} \underbrace{vec\,(G^{*\prime})}_{n^2\times 1},$$

where $vec\,(G^{*\prime})$ have zeros in positions $[(i-1)\,n + j]_{j\in\{1,\dots,n\}}^{i\in\{1,\dots,n\}}$. By removing the zeros, we obtain exactly the elements below the main diagonal of G, gathered in the m-dimensional vector g. Removing the corresponding columns in $-\,(I_n \otimes \widehat{y}_t')$, we construct the matrix W_t, which has the following form:

$$\underbrace{W_t}_{n\times m} = -1 \cdot \begin{bmatrix} 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ \widehat{y}_{1,t} & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \vdots \\ 0 & \widehat{y}_{1,t} & \widehat{y}_{2,t} & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \vdots \\ 0 & 0 & 0 & \widehat{y}_{1,t} & \widehat{y}_{2,t} & \widehat{y}_{3,t} & 0 & \dots & \dots & \dots & \dots & \dots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 & \widehat{y}_{1,t} & \widehat{y}_{2,t} & \widehat{y}_{3,t} & \dots & \widehat{y}_{n-1,t} \end{bmatrix}.$$

We can then rewrite the model as:

$$\begin{aligned} \widehat{y}_t &= -G^*\widehat{y}_t + \Sigma_t\varepsilon_t, \\ \widehat{y}_t &= -\,(I_n \otimes \widehat{y}_t')\,vec\,(G^{*\prime}) + \Sigma_t\varepsilon_t, \\ \widehat{y}_t &= W_t\; g + \varepsilon_t^*, \qquad \varepsilon_t^* \sim \mathcal{MN}\left(\, \mathbf{0}_{n\times 1},\ \Sigma_t^2 \,\right). \end{aligned}$$

Next, we stack the model as:

$$\begin{bmatrix} \widehat{y}_1 \\ \widehat{y}_2 \\ \vdots \\ \widehat{y}_T \end{bmatrix} = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_T \end{bmatrix} g + \begin{bmatrix} \varepsilon_1^* \\ \varepsilon_2^* \\ \vdots \\ \varepsilon_T^* \end{bmatrix},$$

$$\underbrace{\widehat{y}}_{nT\times 1} = \underbrace{W}_{Tn\times m} \cdot \underbrace{g}_{m\times 1} + \varepsilon^*, \qquad \varepsilon^* \sim \mathcal{MN}\left(\, \mathbf{0}_{nT\times 1},\ \Sigma^2 \,\right),$$

where Σ is the diagonal matrix containing all the stacked stochastic volatilities vectors in the main diagonal:

$$\Sigma = Diag\left(\begin{bmatrix}\sigma_1' & \sigma_2' & \dots & \sigma_T'\end{bmatrix}'\right).$$

We can then use a similar approach as the one implemented for a, following Gelman et al. (2014). Specifically, given the prior :

$$g \sim \mathcal{MN}\left(\ \bar{g},\ V_g\ \right),$$

we augment the model with r observations that express the prior information:

$$\begin{aligned}
\begin{bmatrix}\widehat{y}\\ \bar{g}\end{bmatrix} &= \begin{bmatrix}W\\ I_m\end{bmatrix} g + \begin{bmatrix}\varepsilon^*\\ \varepsilon_g\end{bmatrix},\\
\widehat{Y}^\diamond &= W^\diamond g + \varepsilon^\diamond, \quad \varepsilon^\diamond \sim \mathcal{MN}\left(\ \mathbf{0}_{nT+m},\ V_\varepsilon^\diamond\ \right),\\
V_\varepsilon^\diamond &= \begin{bmatrix}\Sigma^2 & \mathbf{0}_{nT\times m}\\ \mathbf{0}_{m\times nT} & V_g\end{bmatrix}.
\end{aligned}$$

A draw for g then comes from the following posterior:

$$\begin{aligned}
g &\sim \mathcal{MN}\left(\ \widetilde{g},\ \left(W^{\diamond\prime}V_\varepsilon^{\diamond-1}W^\diamond\right)^{-1}\ \right),\\
\widetilde{g} &= \left(W^{\diamond\prime}V_\varepsilon^{\diamond-1}W^\diamond\right)^{-1} W^{\diamond\prime}V_\varepsilon^{\diamond-1}\widehat{Y}^\diamond.
\end{aligned}$$

A.4 Step 4: Draw a matrix S_i conditional on $\left(\sigma_t^{i-1}\right)_{t=1}^T$

Before drawing the (unobservable) stochastic volatilities is necessary to draw the matrix S containing the indexes of Normal components. To start building the necessary form, we recall the model formulation used previously and transform it as:

$$y_t = A \cdot Z_t + G^{-1}\Sigma_t\varepsilon_t,$$

$$\underbrace{G\left(y_t - A \cdot Z_t\right)}_{\widetilde{y}_t} = \Sigma_t\varepsilon_t,$$

$$\widetilde{y}_t = \Sigma_t\varepsilon_t.$$

Recalling that

$$\underbrace{\Sigma_t}_{n\times n} = Diag\left(\underbrace{\sigma_t}_{n\times 1}\right),$$

$$\Sigma_t\varepsilon_t = \Sigma_t\begin{bmatrix}\varepsilon_{1,t}\\ \varepsilon_{2,t}\\ \vdots\\ \varepsilon_{n,t}\end{bmatrix} = \begin{bmatrix}\sigma_{1,t}\ \varepsilon_{1,t}\\ \sigma_{2,t}\ \varepsilon_{2,t}\\ \vdots\\ \sigma_{n,t}\ \varepsilon_{n,t}\end{bmatrix} = \sigma_t \odot \varepsilon_t,$$

we take the square element by element on both sides of $\widetilde{y}_t = \Sigma_t\varepsilon_t$, which yields:

$$(\widetilde{y}_t)^{\cdot 2} = (\Sigma_t\varepsilon_t)^{\cdot 2} = \sigma_t^{\cdot 2} \odot \varepsilon_t^{\cdot 2} \quad\Longleftrightarrow\quad \begin{bmatrix}\widetilde{y}_{1,t}^2\\ \widetilde{y}_{2,t}^2\\ \vdots\\ \widetilde{y}_{n,t}^2\end{bmatrix} = \begin{bmatrix}\sigma_{1,t}^2\ \varepsilon_{1,t}^2\\ \sigma_{2,t}^2\ \varepsilon_{2,t}^2\\ \vdots\\ \sigma_{n,t}^2\ \varepsilon_{n,t}^2\end{bmatrix}.$$

We add on the left hand-side a small constant[5] $\bar{c} = 10^{-3}$ and applying the logarithm on both sides[6] in order to break the non-linearity, we obtain:

$$\underbrace{\begin{bmatrix}\log\ \left(\widetilde{y}_{1,t}^2 + \bar{c}\right)\\ \log\ \left(\widetilde{y}_{2,t}^2 + \bar{c}\right)\\ \vdots\\ \log\ \left(\widetilde{y}_{n,t}^2 + \bar{c}\right)\end{bmatrix}}_{\widetilde{y}_t^*} = 2\begin{bmatrix}\log\ \sigma_{1,t}\\ \log\ \sigma_{2,t}\\ \vdots\\ \log\ \sigma_{n,t}\end{bmatrix} + \begin{bmatrix}\log\ \varepsilon_{1,t}^2\\ \log\ \varepsilon_{2,t}^2\\ \vdots\\ \log\ \varepsilon_{n,t}^2\end{bmatrix},$$

$$\widetilde{y}_t^* = 2\log\sigma_t + \log\ \left[(\varepsilon_t)^{\cdot 2}\right].$$

Now, as $\varepsilon_t \overset{iid}{\sim} \mathcal{MN}(\mathbf{0}, I_n)$, each term $\log\varepsilon_{j,t}^2$ has a $\log\chi_1^2$ distribution, and $\log\ \left[(\varepsilon_t)^{\cdot 2}\right]$ is a vector of independent $\log\chi_1^2$ random variables.

Omori et al. (2007), improving upon Kim et al. (1998), show that the $\log\chi_1^2$ is very well

[5]The addition of a small constant term has numerical stability purposes, as explained in Fuller (2009) and Primiceri (2005).

[6]Noticing that $\log\begin{bmatrix}a\\ b\end{bmatrix} = \begin{bmatrix}\log a\\ \log b\end{bmatrix}$.

approximated by a mixture of ten Normal distributions:

$$f_{\log \chi_1^2}(x) \approx \sum_{j=1}^{10} m_j^p \; f_{\mathcal{N}}\left(x \middle| \, m_j^m, m_j^v\right),$$

where m_j^p, m_j^m and m_j^v are contained in the following table:

j	1	2	3	4	5	6	7	8	9	10
m_j^p	0.00609	0.04775	0.13057	0.20674	0.22715	0.18842	0.12047	0.05591	0.01575	0.00115
m_j^m	1.92677	1.34744	0.73504	0.02266	−0.85173	−1.97278	−3.46788	−5.55246	−8.68384	−14.65000
m_j^v	0.11265	0.17788	0.26768	0.40611	0.62699	0.98583	1.57469	2.54498	4.16591	7.33342

Therefore, in order to have a conditionally Gaussian measurement equation, we should condition on a specific history of normal components of the mixtures. Indeed, let us define the matrix S, whose element (h,t) is such that:

$$\underbrace{S}_{n\times T} = [s_{h,t}]_{h\in\{1,\dots,n\}}^{t\in\{1,\dots,T\}}, \qquad \left[\log \; \varepsilon_{h,t}^2 \middle| \, s_{h,t} = j\right] \sim \mathcal{N}\left(m_j^m, m_j^v\right).$$

It follows that we can estimate a history of stochastic volatilies by conditioning on indexes specific Normal components contained in the matrix S. The procedure is then divided in the substeps listed below (following the ordering in Del Negro and Primiceri (2015)).

Taking the volatilities of the previous point of the chain, $\left(\sigma_t^{i-1}\right)_{t=1}^T$, we can use the fact that

$$\log \; \left[(\varepsilon_t)^{\cdot 2}\right] = \widetilde{y}_t^* - 2\log\sigma_t,$$

where $\left(\log \; \left[(\varepsilon_t)^{\cdot 2}\right]\right)_{t=1}^T$ is a sequence of vectors of independent (across time and equations) and identically distributed $\log\chi_1^2$ random variables, which are approximated by mixtures of normal components. Then, the element $s_{h,t}$ that indexes the specific component from which $\log \; \varepsilon_{h,t}^2$ is drawn, has support $J = \{1,\dots,10\}$, and the following discrete probability distribution:

$$\forall j \in J, \qquad Pr\left[s_{h,t} = j \middle| \, \widetilde{y}_{h,t}^*, \sigma_t\right] = \frac{m_j^p \; f_{\mathcal{N}}\left(\widetilde{y}_{h,t}^* - 2\log\sigma_{h,t} \middle| \, m_j^m, m_j^v\right)}{\sum_{\iota=1}^{10} m_\iota^p \; f_{\mathcal{N}}\left(\widetilde{y}_{h,t}^* - 2\log\sigma_{h,t} \middle| \, m_\iota^m, m_\iota^v\right)}.$$

Independent draws for all variables $h \in \{1,\dots,n\}$ at all periods $t \in \{1,\dots,T\}$ from this distribution will form the new matrix S_i.

A.5 Step 5: Draw a history of volatilities $(\sigma_t)_{t=1}^T$

Let us consider the transformation of the model derived in the previous substep:

$$\widetilde{y}_t^* = 2\log\sigma_t + \log\ \left[(\varepsilon_t)^{\cdot 2}\right].$$

Now, conditioning on the just drawn matrix S_i specifying the indexes of Normal components of the mixture, the vector log $\left[(\varepsilon_t)^{\cdot 2}\right]$ has the following Gaussian distribution:

$$\log\ \left[(\varepsilon_t)^{\cdot 2}\right]\Big|\, S\ \sim \mathcal{MN}\left(\underbrace{\begin{bmatrix} m^m_{s_{1,t}} \\ m^m_{s_{2,t}} \\ \vdots \\ m^m_{s_{n,t}} \end{bmatrix}}_{\varphi_t},\ \underbrace{\begin{bmatrix} m^v_{s_{1,t}} & 0 & \dots & 0 \\ 0 & m^v_{s_{2,t}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & m^v_{s_{n,t}} \end{bmatrix}}_{\Upsilon_t}\right)$$

Hence, given the distribution of log $\left[(\varepsilon_t)^{\cdot 2}\right]\Big|\, S$, we can define the needed state space form as:

$$\widetilde{y}_t^* = \varphi_t + 2\log\sigma_t + \zeta_t, \qquad \zeta_t \overset{i}{\sim} \mathcal{MN}\left(\,\mathbf{0},\ \Upsilon_t\,\right),$$

$$\log\sigma_t = \log\sigma_{t-1} + \nu_{\sigma,t}, \qquad \nu_{\sigma,t} \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{Q_\sigma}_{n\times n}\right).$$

At this point, the Forward Filtering Backward Sampling procedure, introduced by Carter and Kohn (1994) can be implemented to draw an history $(\log\sigma_t^i)_{t=1}^T$. The filter can be initialized at:

$$\sigma_{0|0} = \bar{\sigma}, \qquad P_{\sigma,0|0} = \bar{P}_\sigma.$$

A.6 Step 6: Draw a covariance matrix Q_σ

Conditioning on the new $(\sigma_t^i)_{t=0}^T$, we can draw the covariance matrix Q_σ. Indeed, as:

$$\log \sigma_t = \log \sigma_{t-1} + \nu_{\sigma,t}, \qquad \nu_{\sigma,t} \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{Q_\sigma}_{n\times n}\right),$$

having a complete history $(\sigma_t^i)_{t=0}^T$, given the random walk law of motion, is equivalent to having a complete history of the innovations $\nu_{\sigma,t}$. Stacking the $\nu_{\sigma,t}$ across time, we get:

$$\underbrace{\nu_\sigma^*}_{n\times T} = \begin{bmatrix} \nu_{\sigma,1} & \nu_{\sigma,2} & \dots & \nu_{\sigma,T} \end{bmatrix},$$

and we can easily compute the innovation sum of squares matrix:

$$\underbrace{S_\sigma}_{n\times n} = \underbrace{\nu_\sigma^*}_{n\times T}\underbrace{\nu_\sigma^{*\prime}}_{T\times n}.$$

If the prior on the matrix Q_σ is a $n \times n$ Inverse Wishart with scale matrix $\bar{Q}_\sigma$ and degrees of freedom $\tau_{\sigma,0}$:

$$Q_\sigma \sim \mathcal{IW}_n\left(\bar{Q}_\sigma,\ \tau_{\sigma,0}\right),$$

then the posterior is conjugate and given by:

$$Q_\sigma \big| \left(\sigma_t^i\right)_{t=0}^T \sim \mathcal{IW}_n\left(S_\sigma + \bar{Q}_\sigma,\ \tau_{\sigma,0} + T\right)$$

A.7 Alternative Step 3 for the MAI-SVCV: Draw a history of TV off-diagonal elements $\left(g_t^i\right)_{t=1}^T$

Since $m = \frac{n(n-1)}{2}$ may be large, an efficient strategy to attenuate the curse of dimensionality is the rank reduction strategy of de Wind and Gambetti (2014) to draw the time-varying off-diagonal elements $(g_t)_{t=1}^T$. Within this procedure, TV components and time invariant residuals are drawn separately in different substeps.

The vector g_t is decomposed as:

$$g_t = \Lambda_g \widetilde{g}_t + M_g g_0,$$

with

$$\underbrace{\Lambda_g}_{m \times q_r} = \underbrace{V_g}_{m \times q_g} \underbrace{D_g}_{q_g \times q_g}, \qquad M_g = I_m - \Lambda_g \left(\Lambda_g' \Lambda_g \right)^{-1} \Lambda_g',$$

where D_g is the diagonal matrix containing the square roots of the q_g non-zero eigenvalues of Q_g, while V_g is the matrix whose columns are the associated eigenvectors (normalized to unit length). $M_g g_0$ represents the time invariant residual and $\widetilde{g}_t$ is the q_g-dimensional stochastic vector of TV components, having the following law of motion:

$$\widetilde{g}_t = \widetilde{g}_{t-1} + \widetilde{\nu}_{g,t}, \qquad \underbrace{\widetilde{\nu}_{g,t}}_{q_g \times 1} \overset{iid}{\sim} \mathcal{MN} \left(\, \mathbf{0}, \ I_{q_g} \, \right).$$

As in de Wind and Gambetti (2014), the transformed time-varying components $\Lambda_g \widetilde{g}_t$ and the non-TV component $M_g g_0$ are drawn separately.

In order to build the state space form, we first transform the reduced form in the following way:

$$\begin{aligned} y_t &= A \cdot Z_t + G_t^{-1} \Sigma_t \varepsilon_t, \\ \underbrace{y_t - A \cdot Z_t}_{\widehat{y}_t} &= G_t^{-1} \Sigma_t \varepsilon_t, \\ G_t \ \widehat{y}_t &= \Sigma_t \ \varepsilon_t. \end{aligned}$$

Removing ones from the diagonal of G_t, and bringing off-diagonal elements on the right hand side, produces:

$$G_t = I_n + G_t^*.$$

This can be inserted in the model to obtain:

$$\begin{aligned} (I_n + G_t^*) \ \widehat{y}_t &= \Sigma_t \ \varepsilon_t, \\ \widehat{y}_t &= -G_t^* \ \widehat{y}_t + \Sigma_t \ \varepsilon_t. \end{aligned}$$

Exploiting the Kronecker properties, we get:

$$- I_n \underbrace{G_t^*}_{n \times n} \underbrace{\widehat{y}_t}_{n \times 1} = -\underbrace{(I_n \otimes \widehat{y}_t')}_{n \times n^2} \underbrace{vec\,(G_t^{*\prime})}_{n^2 \times 1},$$

where $vec\left(G_t^{*\prime}\right)$ have zeros in positions $\left[(i-1)\,n+j\right]_{j\in\{1,\dots,n\}}^{i\in\{1,\dots,n\}}$. By removing the zeros we obtain exactly the elements below the main diagonal of G_t, gathered in the vector g_t. Removing the corresponding columns in $-\left(I_n \otimes \widehat{y}_t'\right)$, we construct the matrix W_t, which has the following form:

$$\underbrace{W_t}_{n\times m} = -1 \cdot \begin{bmatrix} 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ \widehat{y}_{1,t} & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \vdots \\ 0 & \widehat{y}_{1,t} & \widehat{y}_{2,t} & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \vdots \\ 0 & 0 & 0 & \widehat{y}_{1,t} & \widehat{y}_{2,t} & \widehat{y}_{3,t} & 0 & \dots & \dots & \dots & \dots & \dots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 & \widehat{y}_{1,t} & \widehat{y}_{2,t} & \widehat{y}_{3,t} & \dots & \widehat{y}_{n-1,t} \end{bmatrix}.$$

Hence, we obtain a measurement equation with g_t as states:

$$\begin{aligned} \widehat{y}_t &= -G_t^* \widehat{y}_t + \Sigma_t \varepsilon_t, \\ \widehat{y}_t &= -\left(I_n \otimes \widehat{y}_t'\right) vec\left(G_t^{*\prime}\right) + \Sigma_t \varepsilon_t, \\ \widehat{y}_t &= W_t \; g_t + \Sigma_t \varepsilon_t. \end{aligned}$$

A.7.1 SubStep: draw a history of TV components $\left(\widetilde{g}_t^i\right)_{t=1}^T$

Let us consider the de Wind and Gambetti (2014) decomposition on the m-dimensional process g_t:

$$g_t = \Lambda_g \widetilde{g}_t + M_g g_0, \qquad Q_g = \underbrace{\Lambda_g}_{m\times q_g} \Lambda_g',$$

and plug it into the measurement equation:

$$\begin{aligned} \widehat{y}_t &= W_t g_t + \Sigma_t \varepsilon_t, \\ \widehat{y}_t - W_t M_g g_0 &= W_t \Lambda_g \cdot \widetilde{g}_t + \Sigma_t \varepsilon_t, \\ \widehat{y}_t^* &= W_t^* \; \widetilde{g}_t + \Sigma_t \varepsilon_t. \end{aligned}$$

The resulting state space form is:

$$\widehat{y}_t^* = W_t^* \widetilde{g}_t + \Sigma_t \varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ I_n \right),$$
$$\underbrace{\widetilde{g}_t}_{q_g \times 1} = \widetilde{g}_{t-1} + \widetilde{\nu}_{g,t}, \qquad \underbrace{\widetilde{\nu}_{g,t}}_{q_g \times 1} \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ I_{q_g} \right).$$

The FFBS procedure can be implemented to draw an history $(\widetilde{g}_t^i)_{t=1}^T$. The filter, following de Wind and Gambetti (2014), can be initialized at:

$$\widetilde{g}_{0|0} = R_g \bar{g} \qquad P_{g,0|0} = R_g \bar{P}_g R_g'.$$

where

$$R_g \equiv \left(\Lambda_g' \Lambda_g\right)^{-1} \Lambda_g'.$$

A.7.2 SubStep: draw a vector of time invariant component $M_g g_0$ of g_t

We consider the following transformation to obtain the regression model of interest:

$$\widehat{y}_t = W_t g_t + \Sigma_t \varepsilon_t,$$
$$\widehat{y}_t = W_t \left(\Lambda_g \widetilde{g}_t + M_g g_0 \right) + \Sigma_t \varepsilon_t,$$
$$\underbrace{\widehat{y}_t - W_t \Lambda_g \widetilde{g}_t}_{\widehat{y}_t^\bullet} = W_t \underbrace{M_g g_0}_{\mu_{g,0}} + \underbrace{\Sigma_t \varepsilon_t}_{\widehat{\varepsilon}_t},$$
$$\widehat{y}_t^\bullet = W_t\ \mu_{g,0} + \widehat{\varepsilon}_t \qquad \widehat{\varepsilon}_t \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ \Sigma_t^{\cdot 2} \right).$$

We want to impose the $q_g < m$ restrictions given by:

$$\underbrace{R_g}_{q_g \times r} \underbrace{\mu_{g,0}}_{r \times 1} = \mathbf{0}_{q_g \times 1}.$$

Stacking the time dimensions in columns, we write the restricted regression model as:

$$\underbrace{\widehat{Y}^{\bullet}}_{nT\times 1} = \underbrace{W}_{nT\times m} \mu_{g,0} + \underbrace{\xi}_{nT\times 1}, \qquad \xi \sim \mathcal{MN}\left(\mathbf{0}_{nT\times 1},\ \Sigma^{\cdot 2} \right),$$

$$\underbrace{R_g}_{q_g\times m} \underbrace{\mu_{g,0}}_{r\times 1} = \mathbf{0}_{q_g\times 1}.$$

Then, given the prior for g_0 and the unrestricted Bayesian regression moments:

$$\begin{aligned} g_0 &\sim \mathcal{MN}\left(\bar{g},\ \bar{P}_g \right), \\ \Psi_g &= \left(W'\Sigma^{\cdot -2}W + \bar{P}_g^{-1}\right)^{-1}, \\ \psi_g &= \Psi_g \left(W'\Sigma^{\cdot -2}\ \widehat{Y}^{\bullet} + \bar{P}_g^{-1}\bar{g}\right), \end{aligned}$$

the sampling posterior distribution for $\mu_{g,0}$ is given by:

$$\begin{aligned} \mu_{g,0} &\sim \mathcal{MN}\left(\widetilde{\mu}_{g,0}\ ,\ P_{\mu_g} \right), \\ \widetilde{\mu}_{g,0} &= \Psi_{\mu_g}\psi_g, \\ P_{\mu_g} &= \Psi_{\mu_g}\Psi_g, \\ \Psi_{\mu_g} &= I_r - \Psi_g R_g' \left(R_g \Psi_g R_g'\right)^{-1} R_g. \end{aligned}$$

A.7.3 SubStep: sum the TV and non-TV draws of g_t

We can now sum the draws obtained in the two previous substeps:

$$g_t = \Lambda_g \widetilde{g}_t + M_g g_0, \qquad \forall t \in \{1, \ldots, T\},$$

finally obtaining a history of time-varying off-diagonal elements $(g_t^i)_{t=1}^T$.

A.8 Additional Step for the MAI-SVCV: Draw a reduced rank covariance matrix Q_g

Conditioning on the drawn $(g_t^i)_{t=1}^T$, we can draw the reduced rank covariance matrix Q_g. Indeed, since:

$$g_t = g_{t-1} + \nu_{g,t}, \qquad \nu_{g,t} \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{Q_g}_{r\times r}\right),$$

having a complete history $(g_t)_{t=1}^T$, given the random walk law of motion, is equivalent to having a complete history of innovations $\nu_{g,t}$. Stacking the $\nu_{g,t}$ across time, we get:

$$\underbrace{\nu_g^*}_{r\times T} = \begin{bmatrix} \nu_{g,1} & \nu_{g,2} & \dots & \nu_{g,T} \end{bmatrix},$$

and we can easily compute the innovation sum of squares matrix:

$$\underbrace{S_g}_{r\times m} = \underbrace{\nu_g^*}_{m\times T} \underbrace{\nu_g^{*\prime}}_{T\times m}.$$

By construction, the rank of S_g would be $q_g < m$. Hence, if the prior on the matrix Q_g is an $m \times m$ Singular Inverse Wishart with rank $q_g < m$, scale matrix $\bar{Q}_g$ and degrees of freedom $\tau_{g,0}$:

$$Q_g \sim \mathcal{SIW}_m^{q_g}\left(\bar{Q}_g,\ \tau_{g,0}\right),$$

then, the posterior is conjugate and given by:

$$Q_g \big| \left(g_t^i\right)_{t=0}^T \sim \mathcal{SIW}_m^{q_g}\left(S_g + \bar{Q}_g,\ \tau_{g,0} + T\right).$$

Section B.1 in the appendix discusses how to compute efficiently a draw from a Singular Inverse Wishart distribution.

B Additional details for the GS

B.1 Procedure to draw from a Singular Inverse Wishart distribution

Consider the $m \times m$ matrix Ξ having the following $\mathcal{SIW}$ distribution:

$$\Xi \sim \mathcal{SIW}_m^{\,q}\left(\underbrace{S}_{m\times m},\ d\right),$$

where S is a scale matrix, d the degrees of freedom, and q is the rank of both Ξ and S.

Following Bodnar and Okhrin (2008) and Dıaz-Garcıa et al. (1997), and applying some modifications tailored at improving the algorithm efficiency, it is possible to draw from the distribution of Ξ through the following steps:

1. Construct the Moore–Penrose pseudoinverse S^+ of the scale matrix S. Moore Penrose computation follows the highly efficient Qrginv algorithm proposed by Ataei (2014).
2. Construct the diagonal matrix $\underbrace{\Lambda_S}_{q\times q}$ containing the non-zero eigenvalues of S^+ in the main diagonal, and the matrix $\underbrace{U_S}_{m\times q}$ whose columns are the associated eigenvectors.
3. Draw d independent vectors $\underbrace{\widetilde{z}}_{q\times 1}$ from the multivariate normal $\mathcal{MN}\left(\,\mathbf{0},\ \Lambda_S\,\right)$ and stack them as columns in the matrix $\underbrace{\widetilde{Z}}_{q\times d}$.
4. Construct the matrices $\underbrace{Z}_{m\times q} = U_S\widetilde{Z}$ and $\underbrace{W}_{m\times m} = ZZ'$. The matrix W is a draw from a Singular Wishart distribution with scale matrix S^+ and d degrees of freedom.
5. Compute the Moore–Penrose pseudoinverse W^+ of W. The matrix $\widetilde{\Xi} = W^+$ is a draw from the Singular Inverse Wishart of interest.

B.2 Procedure to draw variable-specific coefficients with triangular structure of heteroskedastic innovations

This approach generalizes the one proposed by Carriero et al. (2016a). Consider the following multivariate regression, with variable-specific coefficients, common regressors across equations, and heteroskedastic errors, whose time variation can be decomposed through a standard triangular reduction.

We are considering the following model:

$$\forall t \in \{1, \dots, T\}, \qquad \underbrace{Z_t}_{n\times 1} = \underbrace{\Theta}_{n\times k} \cdot \underbrace{F_t}_{k\times 1} + \underbrace{u_t}_{n\times 1}, \qquad u_t \overset{i}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{\Omega_t}_{n\times n}\right),$$

$$\Omega_t = W_t \Sigma_t \Sigma_t W_t', \qquad u_t = W_t \Sigma_t \varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} \mathcal{MN}\left(\ \mathbf{0},\ I_n\ \right),$$

where

$$\underbrace{W_t}_{n\times n} = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ w_{1,t} & 1 & \ddots & \ddots & \vdots \\ w_{2,t} & w_{3,t} & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ w_{m-n+2,t} & w_{m-n+3,t} & \dots & w_{m,t} & 1 \end{bmatrix}, \qquad \underbrace{w_t}_{m\times 1} \equiv \begin{bmatrix} w_{1,t} \\ w_{2,t} \\ \vdots \\ w_{m,t} \end{bmatrix}, \qquad m \equiv \frac{n\,(n-1)}{2},$$

$$\underbrace{\Sigma_t}_{n\times n} = \begin{bmatrix} \sigma_{1,t} & 0 & \dots & 0 \\ 0 & \sigma_{2,t} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sigma_{n,t} \end{bmatrix}, \qquad \underbrace{\sigma_t}_{n\times 1} \equiv \begin{bmatrix} \sigma_{1,t} \\ \sigma_{2,t} \\ \vdots \\ \sigma_{n,t} \end{bmatrix}.$$

When $n \cdot k$ gets large, it is convenient to split the previous system in n sets of univariate equations. Indeed, splitting Θ in row vectors it is possible to write the system as:

$$\Theta = \begin{bmatrix} \theta_{[1,:]} \\ \theta_{[2,:]} \\ \vdots \\ \theta_{[n,:]} \end{bmatrix} \quad \rightarrow \quad \begin{cases} z_{1,t} = \theta_{[1,:]} \cdot F_t + \sigma_{1,t}\varepsilon_{1,t} \\ z_{2,t} = \theta_{[2,:]} \cdot F_t + w_{1,t}\sigma_{1,t}\varepsilon_{1,t} + \sigma_{2,t}\varepsilon_{2,t} \\ \vdots \\ z_{n,t} = \theta_{[n,:]} \cdot F_t + w_{m-n+2,t}\sigma_{1,t}\varepsilon_{1,t} + \dots + w_{m,t}\sigma_{n-1,t}\varepsilon_{n-1,t} + \sigma_{n,t}\varepsilon_{n,t} \end{cases}$$

that, defining

$$\underbrace{W_t^*}_{n\times n} = \begin{bmatrix} 0 & 0 & \dots & \dots & 0 \\ w_{1,t} & 0 & \ddots & \ddots & \vdots \\ w_{2,t} & w_{3,t} & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ w_{m-n+2,t} & w_{m-n+3,t} & \dots & w_{m,t} & 0 \end{bmatrix} = W_t - I_n,$$

and

$$\xi_t = W_t^* \Sigma_t \varepsilon_t,$$

can be written as:

$$\begin{cases} z_{1,t} - \xi_{1,t} = \theta_{[1,:]} \cdot F_t + \sigma_{1,t}\varepsilon_{1,t} \\ z_{2,t} - \xi_{2,t} = \theta_{[2,:]} \cdot F_t + \sigma_{2,t}\varepsilon_{2,t} \\ \vdots \\ z_{n,t} - \xi_{n,t} = \theta_{[n,:]} \cdot F_t + \sigma_{n,t}\varepsilon_{n,t} \end{cases}$$

This system is constituted by the following set of n separate equations:

$$\forall j \in \{1, \dots, n\}, \qquad z_{j,t} - \xi_{j,t} = \theta_{[j,:]} \cdot F_t + \varepsilon^*_{j,t}, \quad \varepsilon^*_{j,t} \sim \mathcal{N}\left(0, \sigma^2_{j,t}\right)$$

Notice that for $j \geq 2$ each $\xi_{j,t}$ depends on $\left\{\varepsilon^*_{1,t}, \dots, \varepsilon^*_{j-1,t}\right\}$, since only the first $j-1$ elements of the j-th row of W_t^* are non-zero.

Then, conditioning on $(W_t^*, \sigma_t, F_t)_{t=1}^T$, and given the following priors on the elements of Θ:

$$\forall j \in \{1, \dots, n\}, \qquad vec\left(\theta_{[j,:]}\right) \sim \mathcal{MN}\left(\ \bar{\theta}_{[j,:]},\ V_{j,\theta}\ \right)$$

it is possible to design the following recursive procedure to draw from the posterior of Θ. Starting from $j = 1$:

$$z_{1,t} - \xi_{1,t} = \theta_{[1,:]} \cdot F_t + \varepsilon^*_{1,t}, \quad \varepsilon^*_{1,t} \sim \mathcal{N}\left(0, \sigma^2_{1,t}\right)$$

and given that $\forall t, \xi_{1,t} = 0$, it is possible to draw from the posterior of $\theta_{[1,:]}$ using the following stacked regression:

$$\begin{bmatrix} z_{1,1} \\ \vdots \\ z_{1,T} \end{bmatrix} = \begin{bmatrix} F_1' \\ \vdots \\ F_T' \end{bmatrix} \cdot vec\left(\theta_{[1,:]}\right) + \varepsilon_1^*, \quad \varepsilon_1^* \sim \mathcal{MN}\left(\mathbf{0}, \begin{bmatrix} \sigma_{1,1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{1,2}^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sigma_{1,T}^2 \end{bmatrix}\right),$$

using the prior

$$vec\left(\theta_{[1,:]}\right) \sim \mathcal{MN}\left(\ \bar{\theta}_{[1,:]},\ V_{1,\theta}\ \right).$$

Then, setting the initial condition $\widetilde{\xi}_{j-1,t} = 0$ and the posterior draw $\widetilde{\theta}_{[1,:]}$, the recursive step for any $j \geq 2$ is:

1. The posterior draw $vec\left(\widetilde{\theta}_{[j-1,:]}\right)$ is used to compute the following $\forall t$:

$$\widetilde{\varepsilon}_{j-1,t}^* = z_{j-1,t} - \widetilde{\xi}_{j-1,t} - \widetilde{\theta}_{[j-1,:]} F_t,$$

$$\widetilde{\xi}_{j,t} = W_{t,[j,:]}^* \cdot \begin{bmatrix} \widetilde{\varepsilon}_{1,t}^* \\ \vdots \\ \widetilde{\varepsilon}_{j-1,t}^* \end{bmatrix}.$$

2. The following stacked regression is used to draw from the posterior of $\theta_{[j,:]}$:

$$\begin{bmatrix} z_{j,1} - \widetilde{\xi}_{j,1} \\ \vdots \\ z_{1,T} - \widetilde{\xi}_{j,T} \end{bmatrix} = \begin{bmatrix} F_1' \\ \vdots \\ F_T' \end{bmatrix} \cdot vec\left(\theta_{[j,:]}\right) + \varepsilon_j^*, \quad \varepsilon_j^* \sim \mathcal{MN}\left(\mathbf{0}, \begin{bmatrix} \sigma_{j,1}^2 & 0 & \dots & 0 \\ 0 & \sigma_{j,2}^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \sigma_{j,T}^2 \end{bmatrix}\right),$$

   using the following prior distribution: $\quad vec\left(\theta_{[j,:]}\right) \sim \mathcal{MN}\left(\ \bar{\theta}_{[j,:]},\ V_{j,\theta}\ \right).$

Repeating this iteration until $j = n$, we have drawn from the conditional posterior:

$$\left[\theta_{[1,:]}' \quad \theta_{[2,:]}' \quad \dots \quad \theta_{[n,:]}'\right]' \Big| \left(W_t^*, \sigma_t, F_t, Z_t\right)_{t=1}^T.$$

If W_t^* is non time-varying, the above procedure is implemented in the same way as in the TV case, simply by considering that $\forall t$ we have $W_t^* = W^*$.

C Additional Figures

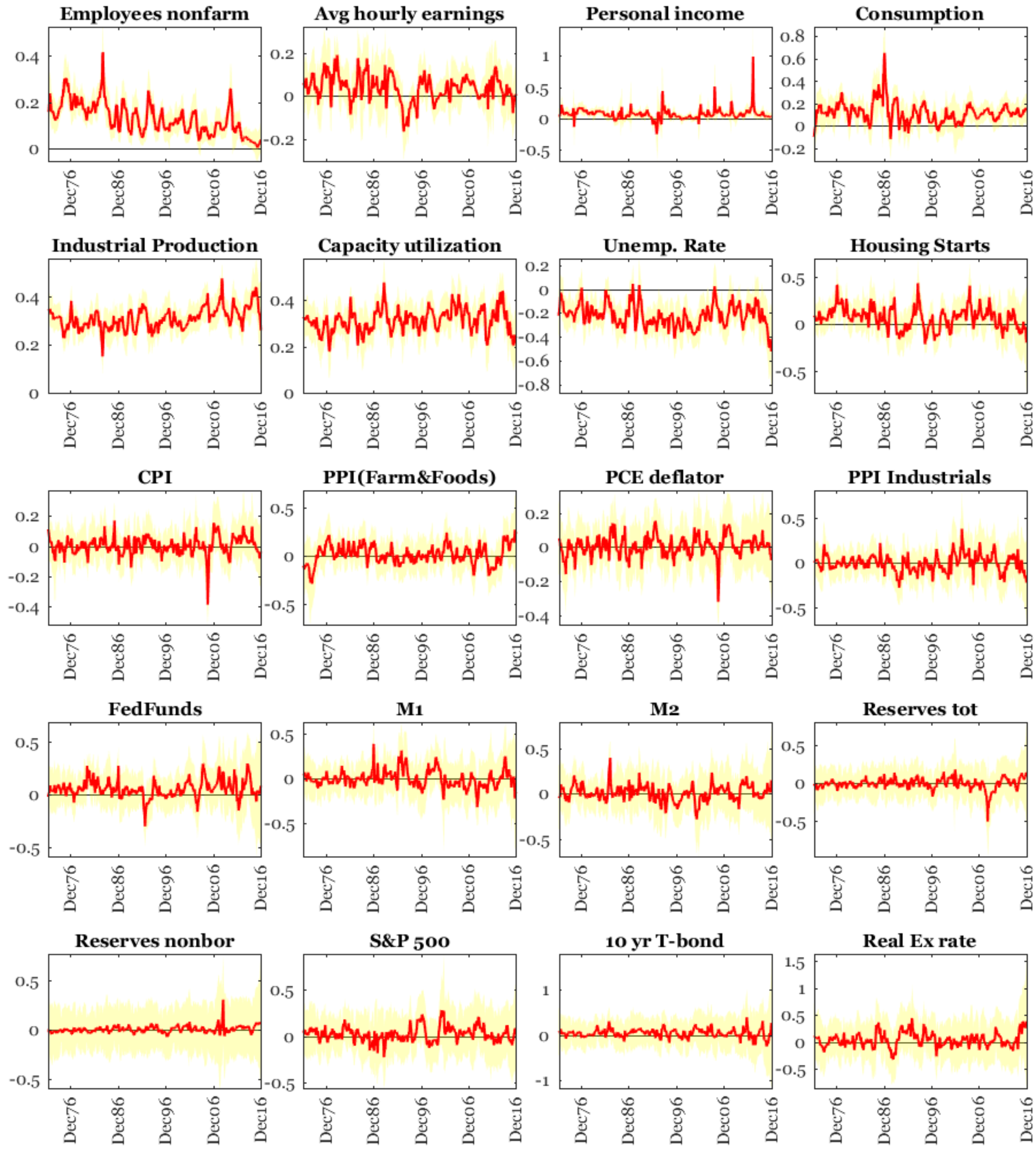


Figure 11: Response to a permanent Demand shock at horizon $h = 1$. Posterior bands.

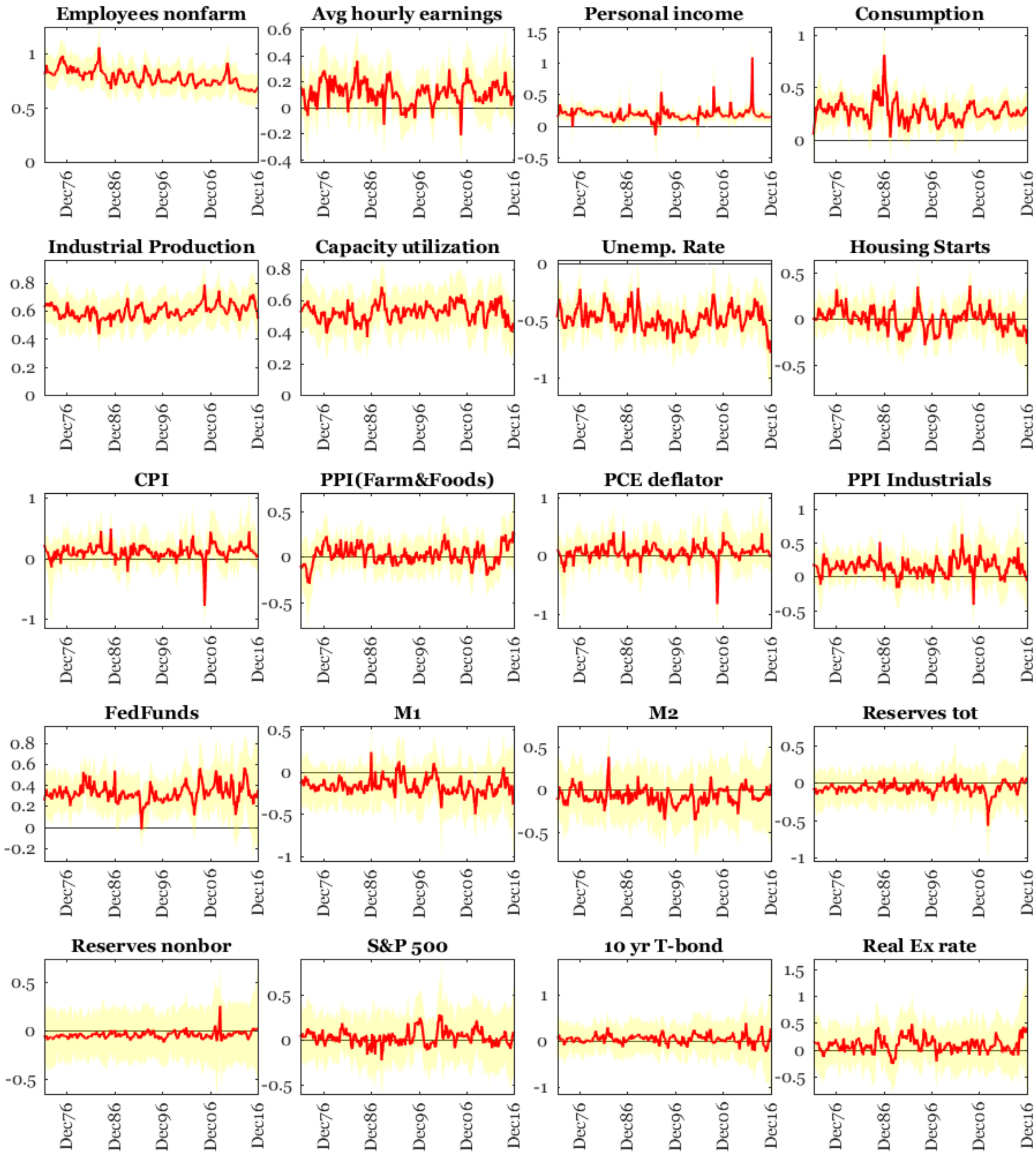


Figure 12: Response to a permanent Demand shock at horizon $h = 12$. Posterior bands.

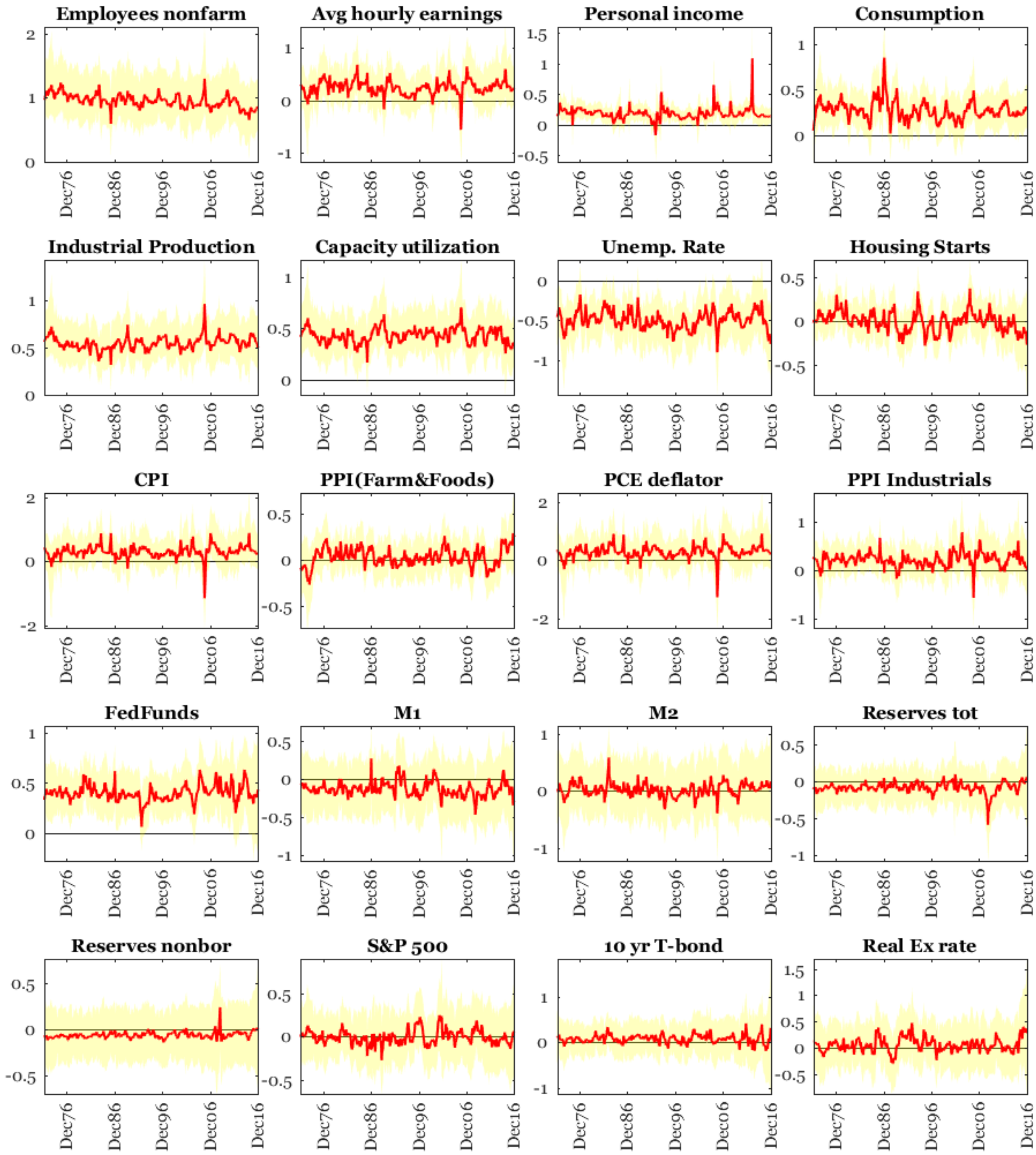


Figure 13: Response to a permanent Demand shock at horizon $h = 48$. Posterior bands.

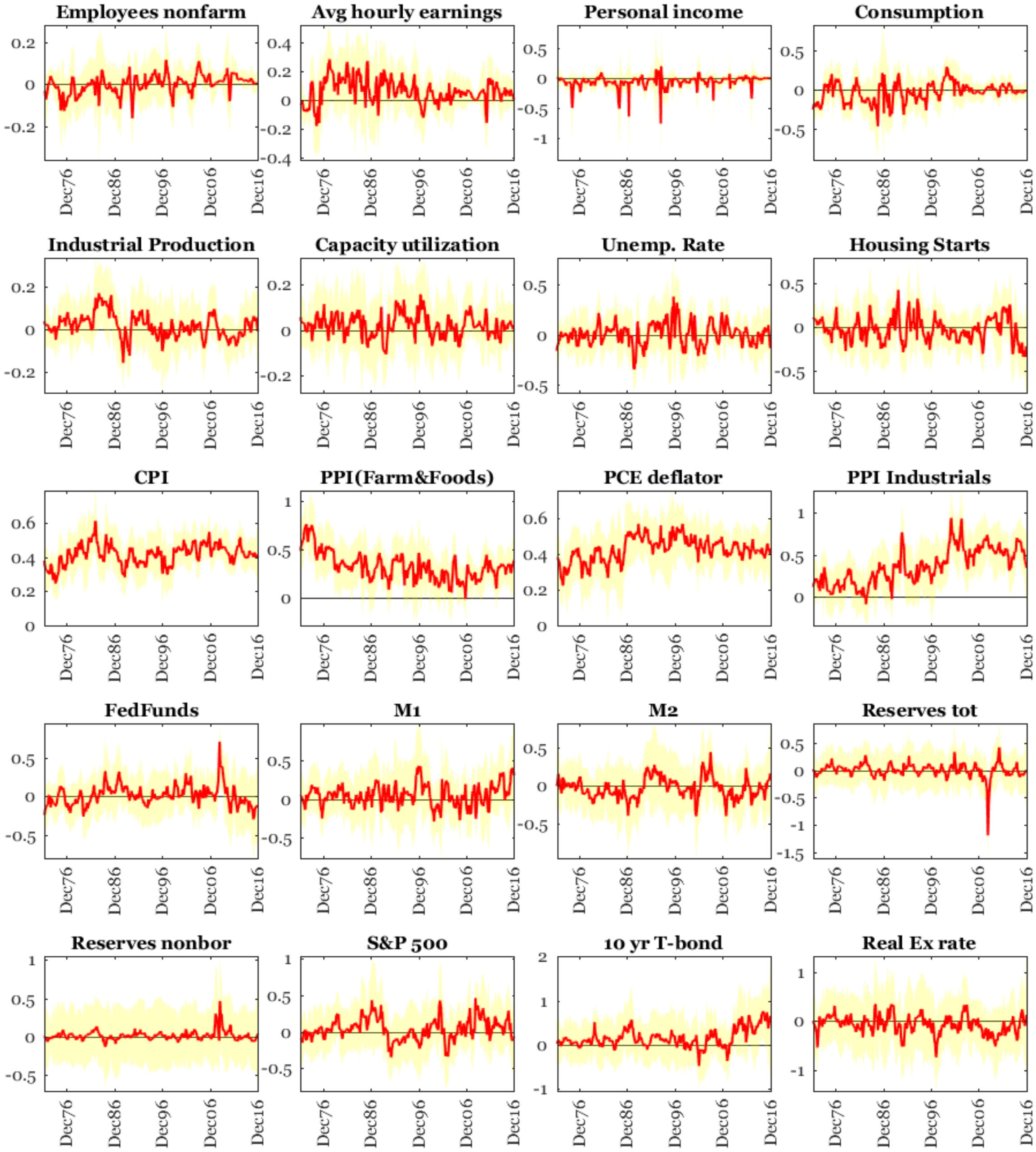


Figure 14: Response to a permanent Supply shock at horizon $h = 1$. Posterior bands.

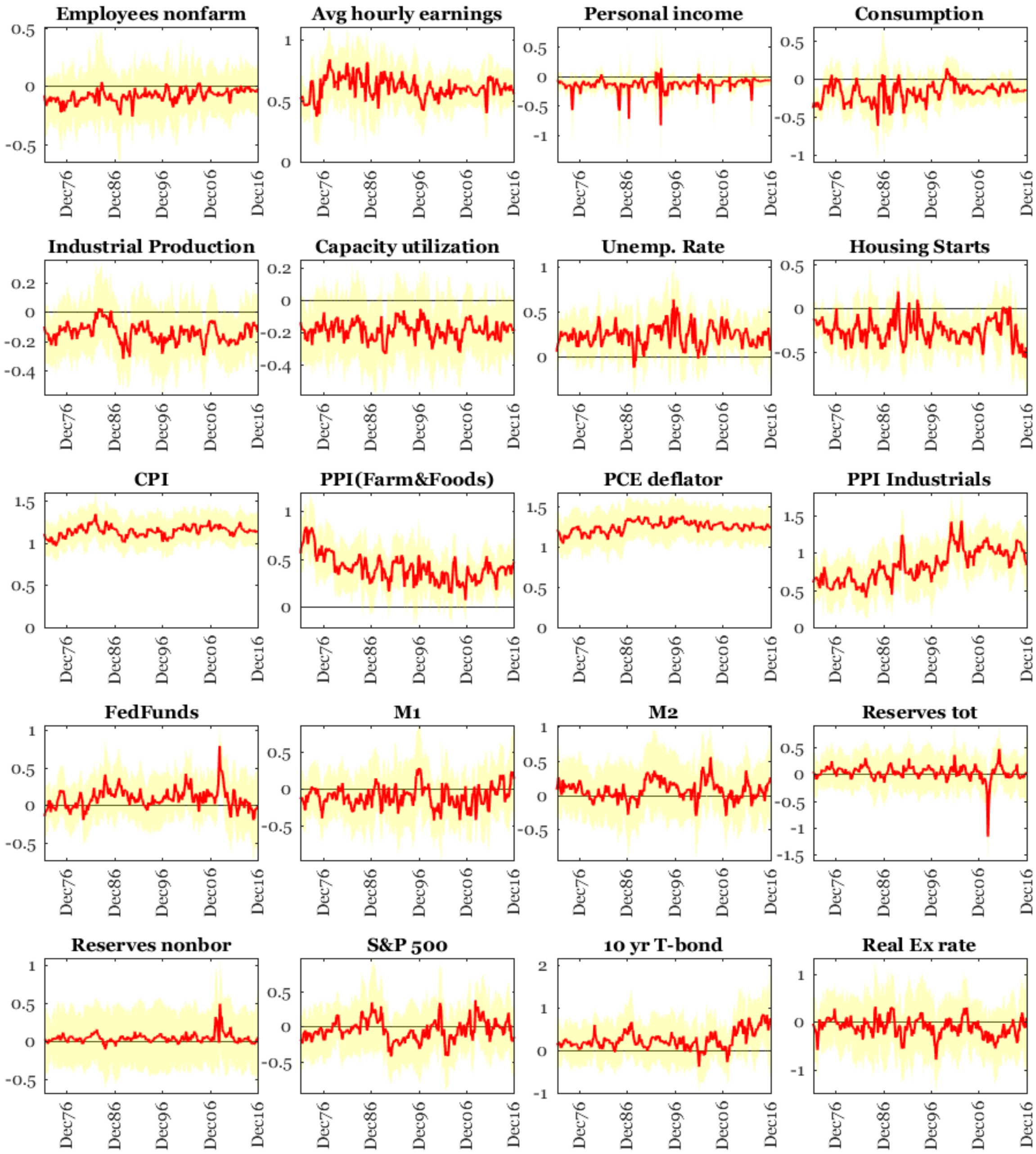


Figure 15: Response to a permanent Supply shock at horizon $h = 12$. Posterior bands.

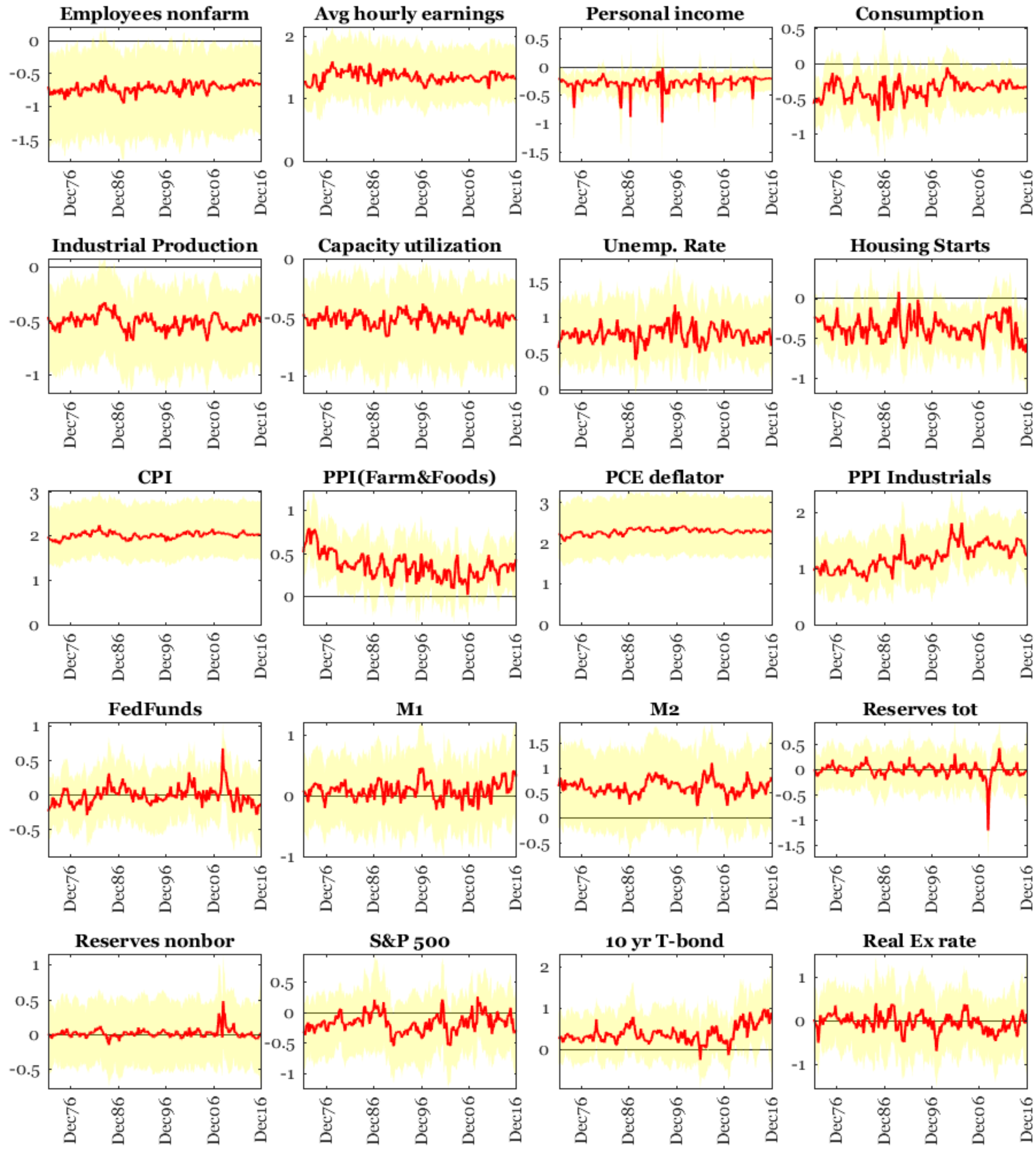


Figure 16: Response to a permanent Supply shock at horizon $h = 48$. Posterior bands.

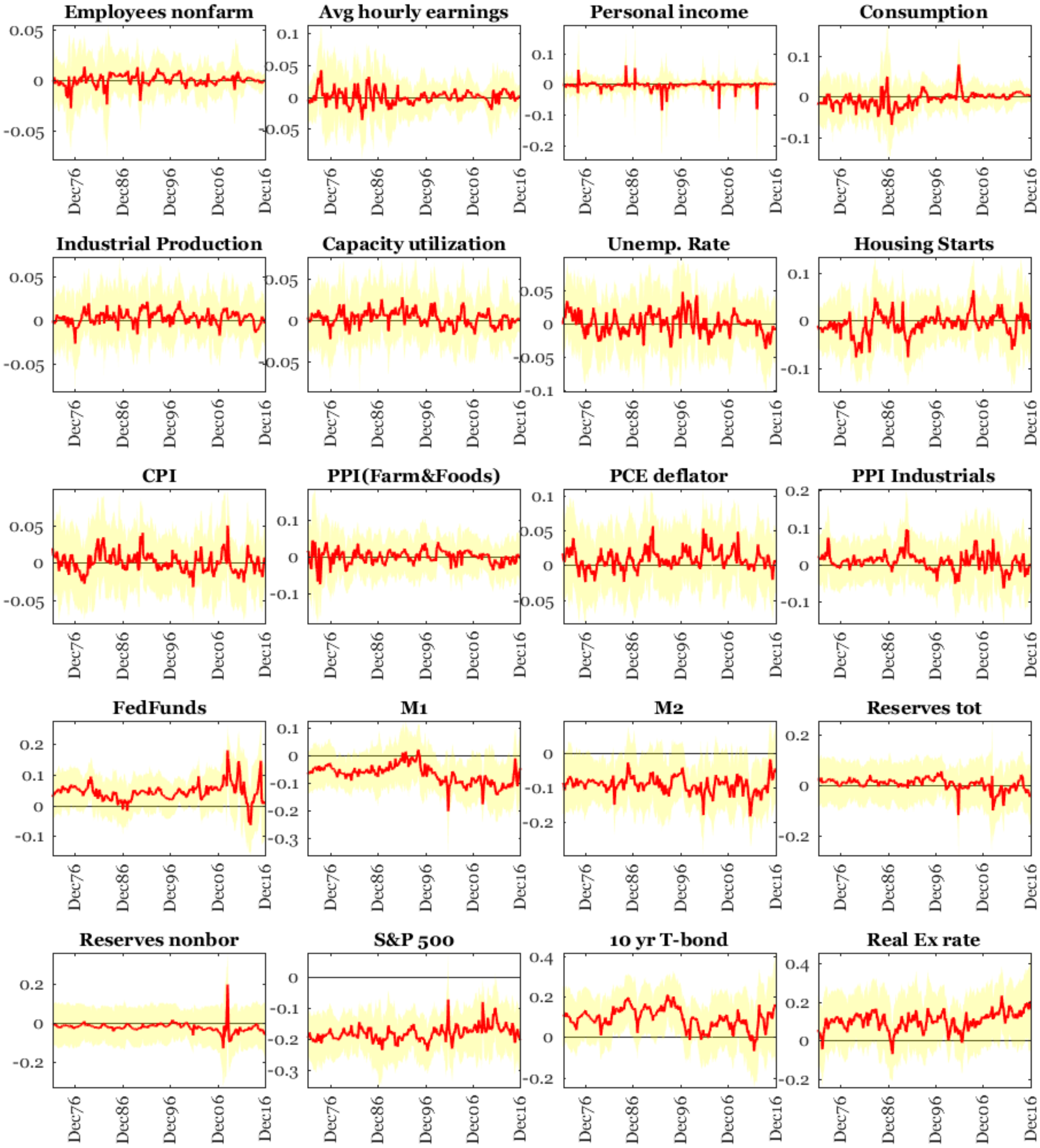


Figure 17: Response to a permanent Financial shock at horizon $h = 1$. Posterior bands.

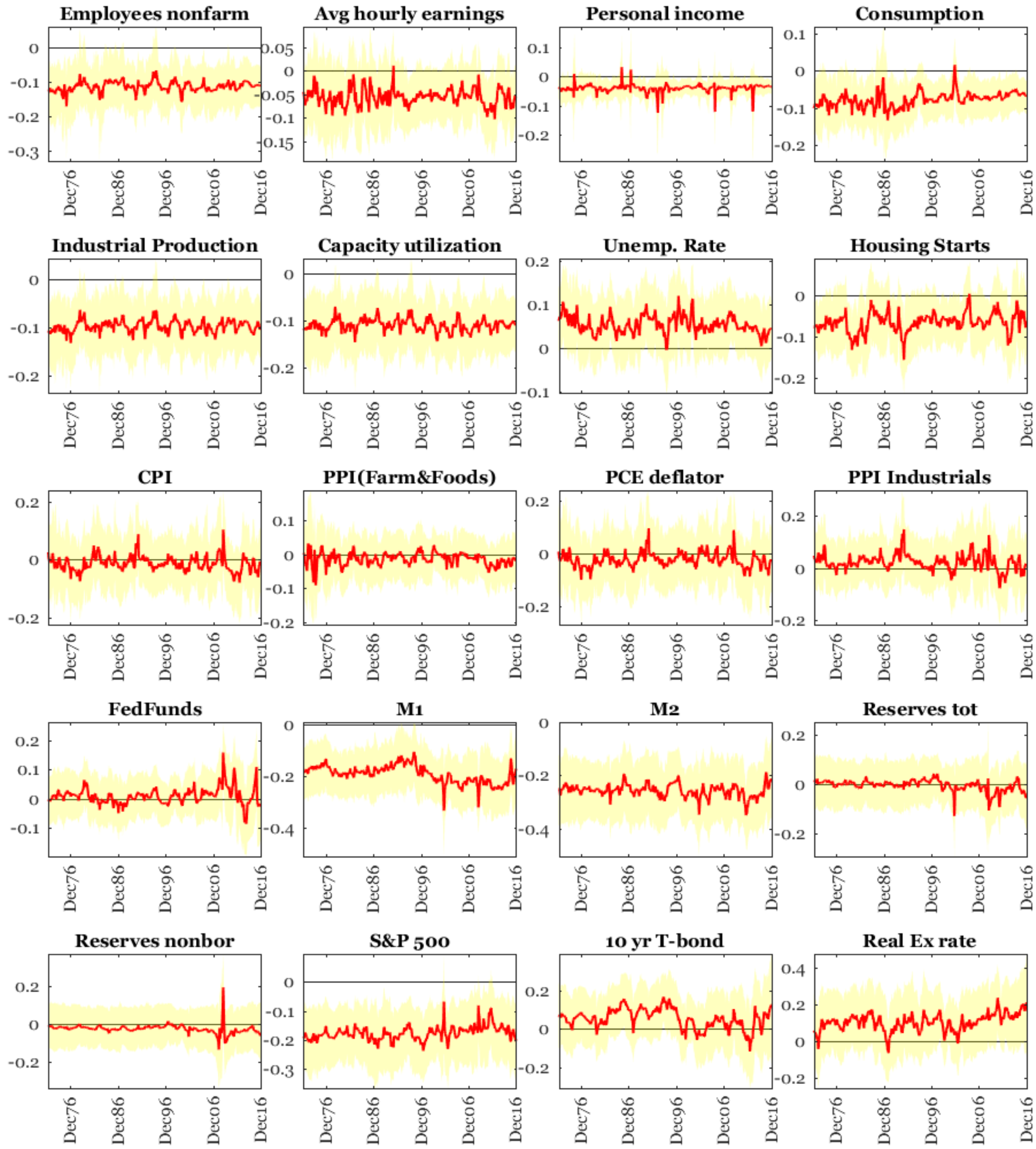


Figure 18: Response to a permanent Financial shock at horizon $h = 12$. Posterior bands.

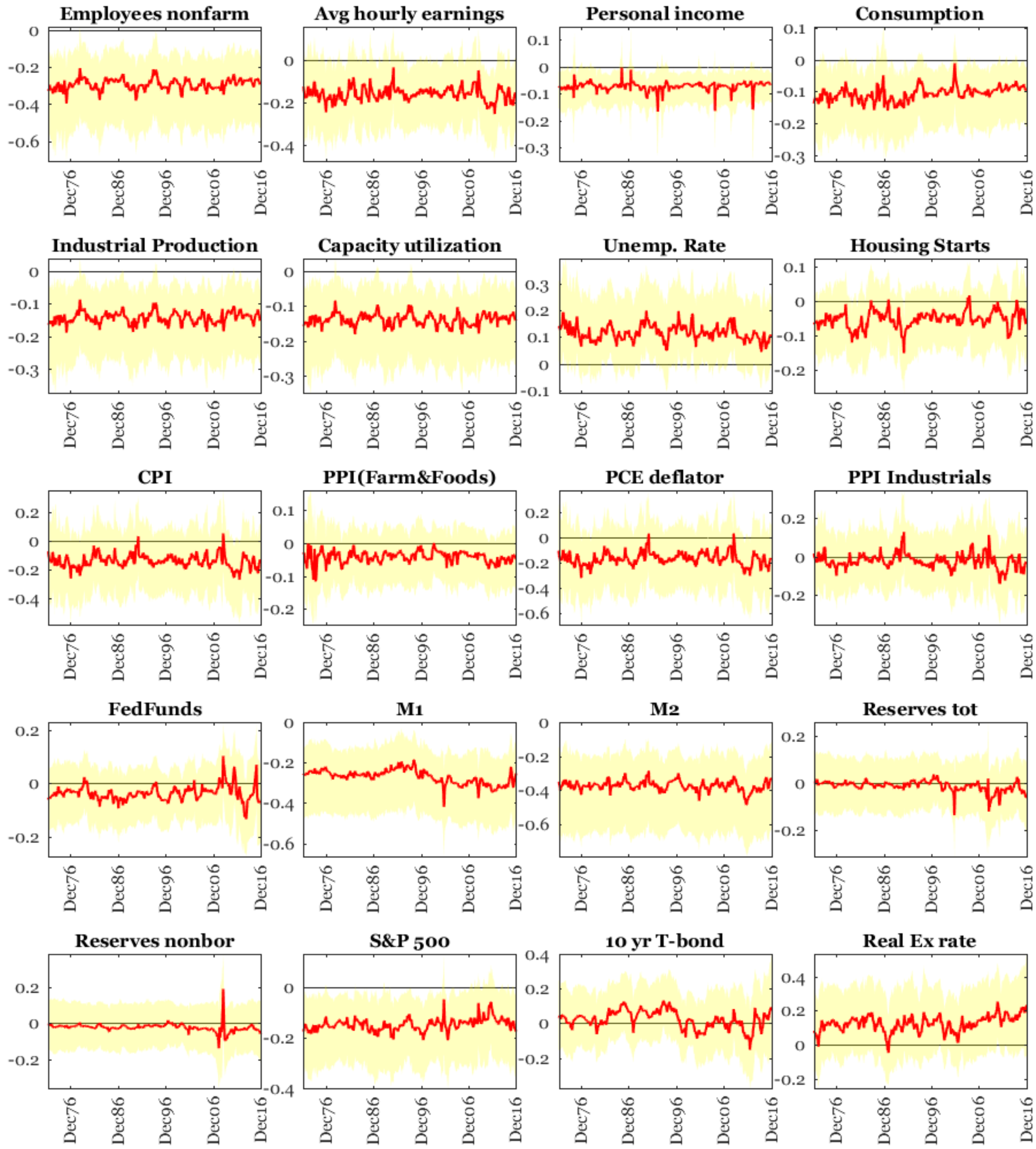


Figure 19: Response to a permanent Financial shock at horizon $h = 48$. Posterior bands.

Chapter 4

The Global Component of Inflation Volatility

co-authored with Andrea Carriero and Massimiliano Marcellino

1 Introduction

Global developments play an important role in the determination of inflation rates. Papers such as Borio and Filardo (2007) and Ciccarelli and Mojon (2010) find that a substantial amount of variation in a large set of national inflation rates can be explained by global factors. Quoting Borio and Filardo (2007): "*...proxies for global economic slack add considerable explanatory power to traditional benchmark inflation rate equations, even allowing for the influence of traditional indicators of external influences on domestic inflation, such as import and Oil prices. Moreover, the role of such global factors has been growing over time, especially since the 1990s. And in a number of cases, global factors appear to have supplanted the role of domestic measures of economic slack.*" This evidence has been recently challenged by Lodge and Mikolajun (2016), whose results suggest that the relevance of global factors for forecasting domestic inflation is related to their ability to capture slow-moving trends, like those emphasized by Stock and Watson (2007) in their decomposition of US inflation into trend and cyclical components. Other empirical contributions, as Bianchi and Civelli (2015) and Auer et al. (2017), show that financial openness and Global Value Chains are positively related to the effects of global slack on inflation. We do not take an a priori stance on this point, but we will use an econometric model where the relative contribution of global and country-specific factors as drivers of inflation developments is estimated and can vary over time and across countries.

Another point stressed by Stock and Watson (2007), which however dates back to at least Engle (1982), is the importance of allowing for conditional time-varying volatility when modelling inflation. While Engle introduced the ARCH specification as a model for inflation volatility, Stock and Watson (2007) used stochastic volatility, which is indeed more common in macroeconomics applications and more flexible since it permits to have different shocks as drivers of the level and volatility of an economic variable. Stock and Watson found that the introduction of SV improves the out of sample forecasting power of their model for US inflation, and it is preferable to both rolling estimation and regime switching to allow for heteroskedasticity. Besides forecasting, inflation volatility is also relevant for policy making as, for example, in periods of high volatility it is more difficult to understand whether inflation movements are temporary or persistent.

Volatility needs to be modeled properly in multi-country studies on inflation determinants. In particular, it seems interesting to understand whether and to what extent the cross-country commonality among inflation levels is also present among inflation volatilities. For example, coordinated cross-country expansionary monetary policy after the crisis could have shrunk the expected inflation paths in each country, leading in the end also to a common reduction in actual inflation volatility. Furthermore, recent macro-financial literature has considered stochastic volatility as a basis to construct measures of macro and financial uncertainty (see Jurado et al., 2015, and Carriero et al., 2017). From this perspective, it may be important for a policymaker to disentangle whether inflation uncertainty originates locally or globally.

Mumtaz and Surico (2008) investigate co-movements in an unbalanced panel of inflation rates from the 1970s to early 2000s for 11 countries, using a large dynamic factor model that incorporates time-varying coefficients and stochastic volatility in the unobservable factors' law of motions. Their decomposition does not show a large role for common components, since most of the time variation in levels and volatilities seems captured by the country-specific component and the residuals, which are left unexplained. They conclude that there has been a fall in level, persistence and volatility of inflation across countries, but with the drop in volatility not synchronized across nations.

Delle Monache et al. (2016) extend the model of Stock and Watson (2007) to a multivariate inflation setting for the euro area, where the permanent component is common among inflation rates of EMU members and the cyclical components are modeled as country-specific autoregressive processes with time-varying parameters. They document that the common permanent component has driven the general disinflation within the euro area, and the importance of common shocks to euro area inflation has increased relatively to idiosyncratic disturbances.

We have collected inflation rates for 20 OECD countries, over the period 1960Q1-2016Q4. Figure 1 reports the time series of CPI inflation rates for each country. From visual inspection, there emerges a non-trivial degree of commonality at low-medium frequencies, as pointed out by Lodge and Mikolajun (2016). A plot of the inflation rates together with their first principal component (PC), Figure 2, provides more evidence on their co-movement (the first PC explains about 70% of the variability of all inflation rates). However, the figure also highlights some country-specific movements in inflation rates, and changes in the volatility of inflation, which seems overall smaller in the later part of the sample. To provide descriptive evidence on commonality in inflation volatility, we have estimated AR-SV models for each inflation rate, and in Figure 3 we report the estimated volatilities together with their first principal component, which explains almost 60% of their time variation.

This evidence motivates the choice of decomposing inflation rates into a common component driven by a single global inflation factor, a country-specific component, and an error term featuring, in turn, common and idiosyncratic time-varying volatility.

Hence, we introduce a novel multivariate autoregressive index (MAI) model, with stochastic volatility (SV), and autoregressive (AR) terms. A MAI model is a VAR with a particular reduced rank structure imposed on the coefficient matrices, such that each variable is driven by the lags of a limited number of linear combinations of all variables (so called Indexes), which can be considered as observable common factors. The MAI model was introduced by Reinsel (1983) and further extended by Carriero et al. (2016b) to allow for a large number of variables. Stochastic volatility (SV) was introduced in the MAI model by Carriero et al. (2018), while Cubadda and Guardabascio (2017) allowed for the possibility of autoregressive (AR) terms to capture idiosyncratic components. We combine all these features into the MAI-AR-SV model, and develop a novel Bayesian MCMC estimation algorithm.

The proposed methodological framework is considerably different from Mumtaz and Surico (2008), who build upon the dynamic factor model of Stock and Watson (1989) and Forni et al. (2000), and estimate their model's stochastic volatilities using the univariate method of Jacquier et al. (2004). Our methodology is also substantially different from Delle Monache et al. (2016), who model multi-country inflation rates with a common permanent component and its own changing volatility, estimated in a non-Bayesian setting in which time variation is driven by likelihood scores.

We work with a single index model where the index (a linear combination of all the national inflation rates) represents the global factor that drives both levels and volatilities of all national inflation rates. Inflation levels and volatilities also have an idiosyncratic, country-specific, component, whose relative importance with respect to the global component is time-varying and empirically determined.

We find that the single common factor in the MAI-SV model explains on average about 70% of the variability of all inflation rates. Moreover, there is also substantial commonality in the inflation volatilities, increased in the last two decades. The average (across countries) share of stochastic volatility explained by the global component spans from 20% to 65% throughout the sample. While various sources can be behind the global inflation factor, it turns out that since the early '90s it is strongly correlated with Chinese PPI and Oil inflation.

We also find that the global inflation factor is highly persistent, and this persistence is transmitted to the global component of the national inflation rates, in line with Ciccarelli and Mojon (2010). Level components explained by the common factor show a larger degree of persistence than idiosyncratic components.

We then repeat the same analysis on a a panel of non-Food and non-Energy inflation rates for the same set of OECD countries, using data available for the period 1979Q1-2016Q4, finding a smaller but non-negligible degree of commonality. The global core inflation factor explains roughly 25% of the variability of core CPI inflation levels and the average (across countries) share of stochastic volatility explained by the global component spans from 10% to 20% throughout the sample, without displaying sizable variation over time as in the case of headline inflation rates. The remaining substantial national component of core inflation level and volatility leaves scope for national monetary policies.

The evidence provided in this paper also contributes to the long standing debate on globalisation, inflation and monetary policy. Rogoff (2003) and Rogoff (2006) discuss how various structural elements accompanying the globalisation since the early 1990s may have lowered the global long term equilibrium of inflation rates, fostering the strong global co-movement of CPI and somehow diminishing the role of domestic slack and monetary policy in determining national inflation. However, as highlighted also in the recent speech by Carney (2017), core inflation seems to be less affected by global dynamics, already when looking at simple pairwise correlation. Our work and methodology allow to measure separately the degree of cross-country commonality in first and second moments of both headline and core inflation rates, providing precious information to monetary policy makers pursuing their inflation mandate in an increasingly global context.

Finally, point and density forecast evaluation shows that the MAI-AR-SV model has very good out of sample properties for inflation rates, when compared with a set of multivariate and univariate competitors, and the SV specification is particularly relevant for the proper calibration of density forecasts. These results hold for both all items inflation and core inflation rates, and provide further empirical support for our proposed model.

The paper is structured as follows. Section 2 introduces the econometric models and the volatility decomposition. Section 3 discusses the choice of prior distributions. Section 4 develops the MCMC estimation methodology, with additional details in the Appendix. Section 5 presents data and empirical results on the commonality in inflation rate levels and volatilities. Section 6 assesses the point and density forecasting performance of the MAI-AR-SV inflation model. Section 7 concludes.

2 The econometric model

2.1 The MAI-AR-SV model

We assume that the model for the n-dimensional zero mean process[1] y_t containing the inflation rates of interest is:

$$\underbrace{y_t}_{n\times 1} = \sum_{\ell=1}^{q} \underbrace{\Gamma_\ell}_{n\times n} \cdot y_{t-\ell} + \sum_{\ell=1}^{p} \underbrace{A_\ell}_{n\times r} \cdot \underbrace{B_0}_{r\times n} y_{t-\ell} + u_t, \tag{1}$$

where Γ_ℓ is a diagonal matrix:

$$\Gamma_\ell = \mathrm{Diag}(\gamma_\ell), \qquad \gamma_\ell = \begin{bmatrix}\gamma_{1,\ell} & \gamma_{2,\ell} & \dots & \gamma_{n,\ell}\end{bmatrix}'.$$

We can rewrite the model more compactly as

$$y_t = \sum_{\ell=1}^{s} (\Gamma_\ell + A_\ell \cdot B_0)\ y_{t-\ell} + u_t,$$

or

$$(I - C_1 L - ... C_s L^s) y_t = C(L) y_t = u_t, \tag{2}$$

where $s = \max(p, q)$ and $C_\ell = \Gamma_\ell + A_\ell \cdot B_0$, $\ell = 1, ..., s$.

Moreover, we assume that

$$u_t = G^{-1}\Sigma_t \varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} \mathcal{MN}\left(\, \mathbf{0},\ I_n \,\right), \tag{3}$$

[1] A non-zero mean can be easily allowed by inserting an intercept in the model.

so that

$$u_t \overset{i}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{\Omega_t}_{n\times n}\right), \qquad \Omega_t = G^{-1}\Sigma_t\Sigma_t\left(G^{-1}\right)', \tag{4}$$

where G is a triangular matrix containing reduced form covariances[2], and $(\Sigma_t)_{t=1}^T$ is the history of diagonal matrices containing the stochastic volatilities:

$$\underbrace{G}_{n\times n} = \begin{bmatrix} 1 & 0 & \dots & \dots & 0 \\ g_1 & 1 & \ddots & \ddots & \vdots \\ g_2 & g_3 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ g_{m-n+2} & g_{m-n+3} & \dots & g_m & 1 \end{bmatrix}, \qquad \underbrace{g}_{m\times 1} \equiv \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{bmatrix}, \qquad m \equiv \frac{n\,(n-1)}{2},$$

$$\begin{aligned} \underbrace{\Sigma_t}_{n\times n} &= \mathrm{Diag}(\sigma_t), \qquad \underbrace{\sigma_t}_{n\times 1} \equiv \begin{bmatrix} \sigma_{1,t} \\ \sigma_{2,t} \\ \vdots \\ \sigma_{n,t} \end{bmatrix}, \\ \log\sigma_t &= \log\sigma_{t-1} + \nu_{\sigma,t}, \qquad \nu_{\sigma,t} \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{Q_\sigma}_{n\times n}\right). \end{aligned}$$

The specification in (1)-(4) is a Multivariate Autoregressive Index (MAI) model with stochastic volatility (SV) and autoregressive (AR) terms, MAI-AR-SV. Each of the n variables in the MAI-AR-SV model is driven by its own lags, capturing in our case country-specific features of inflation, with associated coefficients Γ_ℓ, $\ell = 1, ..., q$; by the lags of r common observable factors ($B_0\ y_{t-\ell}$, the "indexes"), capturing in our case global features of inflation, with associated loading matrices A_ℓ, $\ell = 1, ..., p$; and by variable-specific errors, u_t, whose time-varying covariance matrix Ω_t is expressed as in Cogley and Sargent (2005).

With respect to an unrestricted VAR, the MAI-AR-SV specification leads to a substantial reduction in the number of parameters influencing the conditional means: we go from n^2p

[2]The matrix G can be also made time-varying, but at the cost of a substantial increase in computational complexity when the number of variables is large.

coefficients of the VAR to at most $n \cdot q + n \cdot r \cdot p + r \cdot n$ in the MAI-AR-SV case[3]. In our empirical application, we have $p = q = 4$, $r = 1$ and $n = 20$, so that there are 180 parameters in the MAI-AR-SV while there would be 1600 parameters in an unrestricted VAR.

2.2 An alternative representation of the MAI-AR-SV model

Let us define the obervable factors driving all variables as

$$F_t \equiv B_0 \cdot Y_t, \tag{5}$$

and note that the following decomposition holds:[4]

$$I_n = \Omega_t B_0' \Xi_t^{-1} B_0 + B_{0\perp}' \Xi_{\perp,t}^{-1} B_{0\perp} \Omega_t^{-1}, \tag{6}$$

where $B_{0\perp}$ is the $(n-r) \times n$ orthogonal matrix of B_0, such that the scalar product of any pair of rows of B_0 and $B_{0\perp}$ has zero value[5], $\Xi_t = B_0 \Omega_t B_0'$ and $\Xi_{\perp,t} = B_{0\perp} \Omega_t^{-1} B_{0\perp}'$. Let us also define

$$G_t = B_{0\perp} \Omega_t^{-1} y_t, \tag{7}$$

where G_t are $n-r$ variables that can be interpreted as idiosyncratic components, as there are many of them and, as we will see later on, they are driven by shocks uncorrelated with those driving the common factors F_t.

Using (5)-(7), we can now write the MAI-AR-SV model in (1)-(4) as

$$y_t = \sum_{\ell=1}^{q} \Gamma_\ell [\Omega_t B_0' \Xi_t^{-1} B_0 + B_{0\perp}' \Xi_{\perp,t}^{-1} B_{0\perp} \Omega_t^{-1}] \; y_{t-\ell} + \sum_{\ell=1}^{p} A_\ell \cdot B_0 \; y_{t-\ell} + u_t,$$

or

$$y_t = \sum_{\ell=1}^{q} \Gamma_\ell B_{0\perp}' \Xi_{\perp,t}^{-1} G_{t-\ell} + \sum_{\ell=1}^{\max(p,q)} (\Gamma_\ell \Omega_t B_0' \Xi_t^{-1} + A_\ell) F_{t-\ell} + u_t. \tag{8}$$

[3]Assuming no restrictions in the matrix B_0.

[4]See Carriero et al. (2016b) and the references therein for details.

[5]This is equivalent to state

$$B_0 B_{0\perp}' = \mathbf{0}_{r \times (n-r)}.$$

Next, we derive the model for the factors F_t implied by the MAI-AR-SV model. Starting from (8) and multiplying both sides of it either by B_0 or by $B_{0\perp}\Omega_t^{-1}$, we obtain:

$$F_t = \sum_{\ell=1}^{q} B_0 \Gamma_\ell B'_{0\perp} \Xi_{\perp,t}^{-1} G_{\ t-\ell} + \sum_{\ell=1}^{\max(p,q)} B_0(\Gamma_\ell \Omega_t B'_0 \Xi_t^{-1} + A_\ell) F_{t-\ell} + \omega_t, \tag{9}$$

$$G_t = \sum_{\ell=1}^{q} B_{0\perp}\Omega_t^{-1} \Gamma_\ell B'_{0\perp} \Xi_{\perp,t}^{-1} G_{t-\ell} + \sum_{\ell=1}^{\max(p,q)} B_{0\perp}\Omega_t^{-1}(\Gamma_\ell \Omega_t B'_0 \Xi_t^{-1} + A_\ell) F_{\ t-\ell} + \psi_t,$$

where

$$\begin{bmatrix} \omega_t \\ \psi_t \end{bmatrix} = \begin{bmatrix} B_0 u_t \\ B_{0\perp}\Omega_t^{-1} u_t \end{bmatrix} \overset{i}{\sim} \mathcal{MN}\left(\mathbf{0}, \begin{bmatrix} \Xi_t & 0 \\ 0 & \Xi_{\perp,t} \end{bmatrix}\right), \tag{10}$$

since

$$E(\omega_t \psi_t^{'}) = E(B_0 u_t u_t^{'}\ \Omega_t^{-1} B_{0\perp}^{'}) = B_0 \Omega_t \Omega_t^{-1} B_{0\perp}^{'} = 0.$$

Hence, the r observable factors F_t and the $n-r$ variables G_t jointly evolve as a VAR, with block uncorrelated errors.

The model in (8)-(9) is similar to a factor augmented VAR (FAVAR) model, as for example in Bernanke et al. (2005), or Stock and Watson (2002a) who also allow for variable-specific AR terms. The model in (8)-(9) also features stochastic volatility both in the common (ω_t) and in the idiosyncratic (ψ_t) shocks, which is particularly relevant for modelling inflation, as we will see. Moreover, in the FAVAR model the factors are unobservable, while they are observable in the MAI case, which simplifies model estimation and interpretation of the results. Finally, in general unobserved factors should be modeled with a VARMA rather than a VAR model, as emphasized by Dufour and Stevanović (2013), while in our case we can analytically derive the VAR model followed by the observable factors F_t (jointly with the variables G_t).

2.3 Decomposing the volatilities

We decompose the stochastic volatility of the MAI-AR-SV errors u_t into two orthogonal components, one of them driven by the volatility of the common shocks ω_t, the other by that of the idiosyncratic shocks, ψ_t, orthogonal to ω_t.

Using again the decomposition in (6), we get:

$$u_t = \Omega_t B_0' \Xi_t^{-1} \omega_t + B_{0\perp}' \left(B_{0\perp} \Omega_t^{-1} B_{0\perp}' \right)^{-1} \psi_t,$$

with $\Xi_t = B_0 \Omega_t B_0^{'}$. Hence, due to the orthogonality of ω_t and ψ_t, we can then decompose the total error volatility into the volatility of the common component and that of the idiosyncratic component:

$$\Omega_t = \Omega_t^{com} + \Omega_t^{idio},$$

where

$$\begin{aligned}
\Omega_t^{com} &= \Omega_t B_0' \Xi_t^{-1} B_0 \Omega_t, \\
\Omega_t^{idio} &= B_{0\perp}' \Xi_{\perp,t}^{-1} B_{0\perp}.
\end{aligned}$$

2.4 Decomposing the levels and computing IRFs

It is interesting to decompose the inflation rates in y_t into their common and idiosyncratic components, where the common component is driven by the common shocks ω_t and the idiosyncratic component by the idiosyncratic shocks ψ_t. The decomposition can be also used to compute impulse response functions (IRFs) to common shocks.

We cannot directly use the model in (8), as both F_t and G_t are driven by both ω_t and ψ_t. If the restricted VAR in (2) is stationary, we can write the associated MA representation as

$$\begin{aligned}
y_t &= C(L)^{-1} u_t = B(L) u_t \\
&= \underbrace{B(L) \Omega_t B_0' \Xi_t^{-1} \omega_t}_{\text{Common}} + \underbrace{B(L) B_{0\perp}' \left(B_{0\perp} \Omega_t^{-1} B_{0\perp}' \right)^{-1} \psi_t}_{\text{Idiosyncratic}}.
\end{aligned}$$

As an alternative, we can use the projection approach, proposed for example by Jordà (2005) for IRF computation, to obtain a similar decomposition:

$$y_t = B_1(L) \omega_t + B_2(L) \psi_t.$$

Note that, in both cases, the common and idiosyncratic components are orthogonal at all leads and lags, due to temporal independence and orthogonality of ω_t and ψ_t. Therefore, empirically, we can obtain the common component as the fitted value in a regression of y_t on contemporaneous and lagged values of the (estimated) common shocks ω_t (and the idiosyncratic component as y_t minus the estimated common component), while the IRFs to common shocks are computed from the elements of $B_1(L)$.

In our empirical application on inflation, we have a single factor ($r = 1$), so that ω_t is a scalar, which further simplifies the computation of the common component of inflation rates, and their impulse response functions to global shocks.

The MAI-AR-SV model is estimated by means of Bayesian techniques. The next section describes the specification of prior distributions for model parameters, while section 4 presents the MCMC estimation algorithm, with additional details in the Appendix. Readers not interested in technical details can go directly to the empirical results in Section 5.

3 Estimation of the MAI-AR-SV model

3.1 Specification of the prior distributions

The prior is constructed in various steps, which generally require the use of a training sample $\{-T^*, \ldots, -1, 0\}$.

3.1.1 Prior on B_0 for the Metropolis step

Prior knowledge for the unrestricted elements of B_0 is elicited with a Normal distribution. To define these prior distributions, let us decompose the n variables in r blocks, so to have as many blocks as factors (r).

$$\underbrace{y_t}_{n\times 1} = \begin{bmatrix} \underbrace{y_t^{1\prime}}_{1\times n_1} & \underbrace{y_t^{2\prime}}_{1\times n_2} & \vdots & \underbrace{y_t^{r\prime}}_{1\times n_r} \end{bmatrix}', \qquad n = \sum_{j=1}^{r} n_j.$$

For each $j \in \{1, \ldots, r\}$, we compute the largest eigenvalue score from the Principal components analysis, so to obtain a final set of r score series $\left(S_t^j\right)_{t\in\{1,\ldots,T\}}^{j\in\{1,\ldots,r\}}$. Once obtained the scores, we consider the following $n-r$ univariate regression models:

$$\forall j \in \{1,\dots,r\}\,,\ \forall k \in \{2,\dots,n_j\}\,,\quad S_t^j = B_{0,j,k}\cdot y_{t,k}^j + u_{j,k,t},\quad u_{j,k,t} \overset{iid}{\sim} \mathcal{N}\left(0,\sigma_{j,k}^2\right)$$

To normalize the first element of each $B_{0,j}$, $B_{0,j,1}$ is set at 1. Defining :

$$\forall j \in \{1,\dots,r\}\,,\qquad \underbrace{\widetilde{B}_{0,j}}_{1\times(n_j-1)} \equiv \begin{bmatrix} B_{0,j,2} & \dots & B_{0,j,n_j}\end{bmatrix},$$

for each $\widetilde{B}_{0,j}$, we compute the OLS estimate and its variance.

The prior distribution for $\underbrace{B_0}_{r\times n}$ can be then centered at

$$\underbrace{B_0}_{r\times n} = \begin{bmatrix} 1 & \widetilde{B}_{0,1} & 0 & \mathbf{0}_{1\times(n_2-1)} & \dots & 0 & \mathbf{0}_{1\times(n_r-1)} \\ 0 & \mathbf{0}_{1\times(n_1-1)} & 1 & \widetilde{B}_{0,2} & \dots & 0 & \mathbf{0}_{1\times(n_r-1)} \\ \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \mathbf{0}_{1\times(n_1-1)} & 0 & \mathbf{0}_{1\times(n_2-1)} & \dots & 1 & \widetilde{B}_{0,r}, \end{bmatrix}$$

and the respective variances are coming from each separate regression. Prior covariances among elements are set to zero.

3.1.2 Prior on the loadings A

Defining $A \equiv \begin{bmatrix} A_1 & \dots & A_p\end{bmatrix}$, the prior on $a = vec\left(A'\right)$ is multivariate Normal, centered on $\mathbf{0}$, and with diagonal variance V_a resembling a Minnesota prior.

$$V_a = \begin{bmatrix} \widehat{\sigma}_{y,1}^2 & 0 & \dots & 0 \\ 0 & \widehat{\sigma}_{y,2}^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \widehat{\sigma}_{y,n}^2 \end{bmatrix} \otimes \begin{bmatrix} \Upsilon_1 & 0 & \dots & 0 \\ 0 & \Upsilon_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \Upsilon_p \end{bmatrix},$$

$$\forall \ell \in \{1,\dots,p\}\,,\qquad \Upsilon_\ell = \frac{\lambda_a}{\ell^d}\cdot \begin{bmatrix} \frac{1}{\widehat{\sigma}_{F,1}^2} & 0 & \dots & 0 \\ 0 & \frac{1}{\widehat{\sigma}_{F,2}^2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \frac{1}{\widehat{\sigma}_{F,r}^2} \end{bmatrix},$$

where $\widehat{\sigma}^2_{y,j}$ and $\widehat{\sigma}^2_{F,s}$ are the residual variances of a univariate AR(1) for, respectively, each variable j and each factor s (computed using the prior mean of B_0). λ_a is a tightness parameter.

3.1.3 Prior for the elements of the residual variance

The prior for the elements of G is a multivariate Normal distribution centered at zero, with large diagonal covariance matrix. The prior for σ_0 is a multivariate Normal, centered at $\begin{bmatrix}\widehat{\sigma}^2_{y,1} & \widehat{\sigma}^2_{y,2} & \dots & \widehat{\sigma}^2_{y,n}\end{bmatrix}'$, with identity covariance matrix, as in Primiceri (2005). Prior distributions for the innovation covariance matrix Q_σ is calibrated as in Primiceri (2005).

3.1.4 Prior for the AR coefficients γ

The prior distribution of the AR coefficients in γ is a multivariate Normal distribution. In the spirit of a Minnesota Prior, we choose an a priori unitary mean for the first lag of each variable whose dynamics resemble a random walk, and a zero mean for the higher lags. Regarding the a priori covariance matrix, we assume no correlation across coefficients of different lags and variables, and we set a prior structure for the variances which resembles the Minnesota prior, using the tightness and decay parameters.

$$\bar{\gamma} = \begin{bmatrix} \bar{\gamma}_1 \\ \bar{\gamma}_2 \\ \vdots \\ \bar{\gamma}_q \end{bmatrix} = \begin{bmatrix} \mathbf{1}_{n\times 1} \\ \mathbf{0}_{n\times 1} \\ \vdots \\ \mathbf{0}_{n\times 1} \end{bmatrix}, \qquad V_\gamma = \lambda_\gamma \cdot \begin{bmatrix} 1^{-d} & 0 & \dots & 0 \\ 0 & 2^{-d} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & q^{-d} \end{bmatrix} \otimes I_n.$$

3.2 Gibbs Sampler

This subsection describes each step of the Gibbs Sampler (GS) used to simulate from the joint posterior distribution of both parameters $\{\gamma, A, B_0, G, Q_\sigma\}$ and unobservable states $(\sigma_t)_{t=1}^T$ of the MAI-AR-SV model. Moreover, the Omori et al. (2007) procedure requires drawing the indexes of Normal components of the mixture approximating the $\log\chi^2_1$, contained in the matrix S. This approach is needed as the joint posterior distribution cannot be analytically determined. The steps are the following:

1. Draw the AR coefficients $\gamma \left| A, B_0, G, Q_\sigma \left(\sigma_t\right)_{t=1}^T \right.$,
2. Draw the loadings $A \left| \gamma, B_0, G, Q_\sigma, \left(\sigma_t\right)_{t=1}^T \right.$,
3. Draw the factor weights $B_0 \left| \gamma, A, G, Q_\sigma, \left(\sigma_t\right)_{t=1}^T \right.$,
4. Draw the off-diagonal elements in $G \left| \gamma, A, B_0, Q_\sigma, \left(\sigma_t\right)_{t=1}^T \right.$,
5. Draw the indexes of the mixture in $S \left| \gamma, A, B_0, G, Q_\sigma, \left(\sigma_t\right)_{t=1}^T \right.$,
6. Draw a history of volatilities $\left(\sigma_t\right)_{t=1}^T \left| S, \gamma, A, B_0, G, Q_\sigma \right.$,
7. Draw the covariance of volatilities' innovations $Q_\sigma \left| \gamma, A, B_0, G, \left(\sigma_t\right)_{t=1}^T \right.$.

It is important to note[6] that steps 2 and 3 have $y_t - \mathcal{X}_t \cdot \gamma$ as dependent variable in order to draw A and B_0, while in step 1 we use $y_t - A \cdot Z_t$ to draw the AR coefficients[7] [8].

Each step of the GS for the MAI-AR-SV is described in detail in section A of the Appendix.

4 The global component of inflation volatility

4.1 Data

Following the literature on global inflation (e.g. Ciccarelli and Mojon, 2010 and Borio and Filardo, 2007) we gathered a panel of Consumer Price Indices for a set of 20 OECD countries[9], downloaded from the *OECD main economic indicators* database. The dataset includes 228 observations at quarterly frequency, covering the period from 1960-Q1 to 2016-Q4. We then constructed inflation rates as year on year changes of the indexes[10].

[6]See the Appendix for further details.

[7]$Z_t \equiv (I_p \otimes B_0) \cdot vec\left(\begin{bmatrix} y_{t-1} & \ldots & y_{t-p} \end{bmatrix}\right)$.

[8]$\mathcal{X}_t = \begin{bmatrix} Diag(y_{t-1}) & Diag(y_{t-2}) & \ldots & Diag(y_{t-q}) \end{bmatrix}$.

[9]USA, Australia, Austria, Belgium, Canada, Finland, France, Germany, Greece, Italy, Japan, Luxembourg, Netherlands, New Zealand, Norway, Portugal, Spain, Sweden, Switzerland, UK

[10]Ciccarelli and Mojon (2010) use Year on Year changes of CPI inflation rates for the bulk of their analysis. O'Reilly and Whelan (2005) adopt the same transformation stressing that is cited in the ECB's official inflation mandate. Lodge and Mikolajun (2016) point out that using YoY changes in CPI is preferable since this transformation produces no seasonal pattern by construction.

4.2 MAI-SV and MAI-AR-SV

We start with a MAI-SV specification (that is with $\Gamma_\ell = \mathbf{0},\ \forall \ell$), with $p = 4$ lags and with a single global factor ($r = 1$), similar to the preferred specification of Ciccarelli and Mojon (2010). The resulting model is estimated by a simplified version of the MCMC algorithm presented in Section 4, see Carriero et al. (2018) for details.

Figure 4 reports the inflation rates for each country along with the posterior bands and median of the estimated common global inflation factor. The model is clearly able to capture the substantial co-movement of national inflation rates.

Figure 5 reports the data compared with the in-sample fit of the MAI-SV model for each country, as well as the percentage share of variance explained. On average (across countries) the estimated common component explains roughly 73% of the variance, which is in line with the Principal Component Analysis.

Next, as the residuals of the MAI-SV model are clearly serially correlated at least over parts of the sample, we estimate a MAI-AR-SV model with $p = 4$ lags for the common part, as for the MAI-SV, and $q = 4$ lags for the country-specific AR components. The in-sample fit for the various countries is presented in Figure 6. The fit of the MAI-AR-SV is systematically higher than that of the MAI-SV specification, reaching an average explained variance of about 94%. In particular, the MAI-AR-SV specification is able to capture both the low and the high frequency variation of each inflation series, due to the presence of both common and country-specific autoregressive components.

Notwithstanding the differences mentioned above, the estimated global factor from the MAI-SV and MAI-AR-SV models are very similar, see Figure 7. They are also very similar to the first PC of the inflation rates. The latter is used to form the prior on the B_0 coefficients in the MAI models, but the prior variance is large enough so that results are data driven rather than dictated by the prior. All such measures of common components are also comparable, though with some differences, to an OECD measure of global inflation, also reported in Figure 7. These results are in line with the findings of Ciccarelli and Mojon (2010), even though their sample stops in 2008. As reported also by Ferroni and Mojon (2016), our analysis suggests strong commonality in inflation developments across OECD continues also in the more recent period, and, actually, it has been particularly high during the last financial crisis[11].

[11]Using a more recent sample of inflation rates (1993-2014), Ferroni and Mojon (2016) find that the

Finally, while there can be many drivers of the global inflation factor, Figure 8 shows that after the 90's it is correlated with the Chinese PPI inflation rate and the Oil inflation rate.

4.3 Levels decomposition and persistence

Using the level decomposition discussed in section 2, we are able to disentangle the observed inflation series of each country into orthogonal components driven, respectively, by common and idiosyncratic shocks. Moreover, for each country we measure how much variation is explained by each component.

Figures 9 and 10 report, respectively, the common and idiosyncratic components, compared with the actual series. The common components tend to explain more than 50% of almost all countries' inflation rates, and are particularly important in large economies like the US, UK, Germany and Japan.

Stock and Watson (2007) discuss the persistence of US inflation, using as a measure of persistence the largest autoregressive root of the levels' process. Inference about this measure of persistence is made possible by the Stock (1991) method, which is appropriate when dealing with series displaying high levels of persistence. Stock and Watson (2007) do not find strong evidence of persistence changes in US inflation from the 1970s onwards, reporting the largest AR root of US CPI inflation comprised between 0.85 and 1.05 (as 90% confidence interval). O'Reilly and Whelan (2005) report little evidence of instability for inflation persistence in the Euro Area since the 1970s; they report rolling confidence intervals for the largest AR root of Euro Area CPI inflation that are centered around 0.9 across almost the entire sample.

In light of this literature, using the entire sample, we computed the 90% confidence intervals (CI) for the largest AR roots of all national CPI inflation series, of their common and idiosyncratic components, and of the global factor. Figure 11 compares the CI for the largest AR root of the observed series, their components and the global factor, separately for each country. The picture clearly shows how the common global components tend to preserve the high persistence of the observed series, while the idiosyncratic country-specific components display wider confidence intervals centered on slightly smaller values. The global factor shows a very narrow CI centered on 0.99.

fraction of national inflation rates' variance that is explained by Global Inflation remains dominant.

These results are in line with what reported by Ciccarelli and Mojon (2010), who argue that "*the global component captures the most persistent and possibly nonstationary part of inflation*". Indeed, using a different methodology, they report smaller persistence for the so called "national" components; interpreting such results, they consider the global factor as an attractor and the main driver of persistence coming from the observed data. However, for this specific exercise they use annualized quarter on quarter inflation rates, which is a transformation that tends to display a smaller degree of persistence than the year on year transformation. Performing our analysis using QoQ CPI changes, we measure a degree of persistence in line with Ciccarelli and Mojon (2010) for both global and national inflation components.

4.4 Time-varying residual volatility decomposition

Figure 12 reports the posterior bands of the estimated reduced form conditional inflation volatilities of all countries for the MAI-AR-SV. The estimated volatilities display a relevant degree of commonality. Indeed, the first principal component of the volatilities explains on average about 50% of their variation.

Principal component analysis is however not so suited in this context, due to the time-varying covariance matrix of the errors. Hence, to better understand what is driving the volatilities, we can apply the decomposition discussed in Section 2. Figures 13 and 14 present the decomposition of the estimated volatilities in their common and idiosyncratic components. More specifically, Figure 13 presents results in absolute terms and Figure 14 in relative terms. It turns out that the contribution of the common component is non trivial, reaching values above 50% for some countries and time periods, especially during the last decades, in particular during the Great Recession.

In this multi-country context, it is complex to understand the drivers of the common inflation volatility component. However, for a single country this can be done. Carriero et al. (2018), focusing on the US, find that supply shocks are particularly important, with demand shocks ranked second and monetary/financial shocks third.

Figure 15 shows the posterior bands of the global factor volatility, that is $(\Xi_t)_{t=1}^{T}$. Global inflation volatility was moderate during the 1960s, increased dramatically during the 1970s before the sharp reduction starting in the 1980s associated with the change in monetary policy to fight inflation occurred in several countries. These results are in line with the

US inflation volatility estimated by Stock and Watson (2007). Global inflation volatility has remained very low until mid 2000s, reaching a new spike during the Great Recession, before turning back to the historically low values of the last 3/4 years. Time variation is significant and relatively large throughout the entire estimation sample.

In order to understand which global forces may correlate with global CPI inflation volatility, we estimated two measures of stochastic volatilities from separate univariate AR-SV models for Oil inflation, measured by the WTI price ($/barrel), and for Chinese PPI inflation, available only from the early 1990s. A comparison of median volatilities is reported in Figure 16. From visual inspection, a clear co-movement between Oil and global CPI inflation volatility stands out, showing a correlation around 0.5 from the early 1970s and almost 0.8 from the early 1990s. Also the Chinese PPI inflation volatility displays a positive correlation with global CPI uncertainty: the correlation is around 0.7 from the early 1990s.

4.5 Commonality in core inflations

In light of the correlation (in both levels and volatilities) between the global component of headline CPI inflation and Oil, it is important to detect how much core components of the CPIs remain correlated. To this end, the same exercises of this section have been performed using the non-Food and non-Energy Consumer Prices Indices for the same set of countries, downloaded from the *OECD main economic indicators* database. These data are available only from the late '70s onwards.

Non-Food and non-Energy inflation tends to display a lower degree of commonality, already from a quick graphical inspection. Performing our decompositions, results collected in Appendix C indicate a smaller importance of the common component both in volatilities and in levels: the global core inflation factor explains roughly 25% of the variability of core CPI inflation levels, while the average (across countries) share of stochastic volatility explained by the global component spans from 10% to 20% throughout the sample. The fact that core inflation remains mostly a national phenomenon leaves ample scope for national monetary policies.

5 Forecasting inflation with the MAI-AR-SV model

To provide further evidence on the usefulness of the MAI-AR-SV as a model for multi-country inflation, we now evaluate its out of sample properties, also in comparison with a set of standard competitors.

Using the same inflation series employed in the structural analysis, several models are recursively estimated on a forecasting window of 101 quarterly vintages (forecasting window starts from 1990Q1). The associated out of sample forecasts are produced for six different models and 8 horizons, from 1 to 8 quarters ahead.

The models under evaluation are the following:

- the Multivariate Autoregressive Index model with AR components and Stochastic Volatility (MAI-AR-SV)
- the Multivariate Autoregressive Index model with AR components (MAI-AR)
- the univariate Autoregressive model (AR)
- the univariate Autoregressive model with Stochastic Volatility (AR-SV)
- the Vector Autoregressive model (VAR)
- the Vector Autoregressive model with Stochastic Volatility (VAR-SV)

All models are estimated using Bayesian techniques. AR and VAR priors are constructed using the standard Litterman (1986) a priori assumption of univariate random walk processes. The SV prior in all models is calibrated as in Primiceri (2005). The MAI prior is specified as shown in section 3.

Diagnostics are then computed both in terms of point forecasting and density forecasting, following the evaluation framework of Clark and Ravazzolo (2015).

Specifically, to evaluate the accuracy in terms of point forecasting, we compute the forecasts posterior medians for all vintages, models, variables and horizons. Then, we compute the Root Mean Squared Forecast Error (RMSFE) for each model, variable and horizon, using the variation across vintages. Hence, for each variable $j \in \{1, \ldots, n\}$, each horizon $h \in \{1, \ldots, H\}$ and each model $m \in \{1, \ldots, M\}$ we compute:

$$RMSE_{j,h}^{m} = \sqrt{\frac{1}{T^*} \sum_{t=T+1}^{T+T^*} \left(y_{j,t+h} - \widehat{y}_{j,t+h}^{m} \right)^2},$$

where $\widehat{y}^{m}_{j,t+h}$ is the median of the posterior distribution $\left(\widehat{y}^{m,i}_{j,t+h}\right)_{i=1}^{L_c}$ (L_c is the length of the discretized posterior distribution). To test for significance of the squared forecast errors differences across models, we compute the Diebold and Mariano (1995) t-tests for equality of the average loss.

To evaluate models in terms of density forecasting, we use two measures of accuracy: the average log-predictive score and the average Continuous Ranked Probability Score (CRPS). Even in this case, to test for significantly different performances we employ the Diebold and Mariano test, following Clark and Ravazzolo (2015).

Log Predictive Scores are obtained via non-parametric kernel smoothing density estimators. Adopting a normal kernel $\mathcal{K}_{\mathcal{N}}(\cdot)$ and following an optimal selection strategy of the bandwith parameter $\widehat{\mathcal{H}}$, we can compute for each variable, model, horizon and vintage the empirical density evaluted at the actual observation $y_{j,t+h}$, that is:

$$\widehat{f}_m\left(y_{j,t+h},\widehat{\mathcal{H}}\right)=\frac{1}{\widehat{\mathcal{H}}\cdot L_c}\sum_{i=1}^{L_c}\mathcal{K}_{\mathcal{N}}\left(\frac{y_{j,t+h}-\widehat{y}^{m,i}_{j,t+h}}{\widehat{\mathcal{H}}}\right).$$

Then, applying logarithms and computing the average across forecasting vintages yields the average log score for each variable, model and horizon:

$$\overline{logScore}^{m}_{j,h}=\frac{1}{T^*}\sum_{t=T+1}^{T+T^*}\log\widehat{f}_m\left(y_{j,t+h},\widehat{\mathcal{H}}\right).$$

To compute the average CRPS, following Clark and Ravazzolo (2015), we first compute the CRPS per each variable, model, horizon and vintage, making use of the actual observations, the posterior distribution $\left(\widehat{y}^{m,i}_{j,t+h}\right)_{i=1}^{L_c}$ and a random permutation of the latter $\left(\widehat{y}^{m,i'(i)}_{j,t+h}\right)_{i=1}^{L_c}$ where $i':\{1,\ldots,L_c\}\to\{1,\ldots,L_c\}$ is randomly drawn without replacement. Lastly, we simply compute the average across time vintages:

$$CRPS^{m}_{j,t+h}=\frac{1}{L_c}\sum_{i=1}^{L_c}\left|\widehat{y}^{m,i}_{j,t+h}-y_{j,t+h}\right|-\frac{1}{2\cdot L_c}\sum_{i=1}^{L_c}\left|\widehat{y}^{m,i}_{j,t+h}-\widehat{y}^{m,i'(i)}_{j,t+h}\right|,$$

$$\overline{CRPS}^{m}_{j,h}=\frac{1}{T^*}\sum_{t=T+1}^{T+T^*}CRPS^{m}_{j,t+h}.$$

Figure 17 portrays the relative performance of the competing set of models against the benchmark model MAI-AR-SV, for each country and four selected horizons. Models' point forecasting performance is reported as ratio between their own Root Mean Squared Errors and the benchmark's, so that values larger than one imply that the MAI-AR-SV produces more accurate point forecasts. The MAI-AR-SV model improves significantly upon its counterparts on most variables, especially at short horizons, even though in a smaller number of cases this is reversed. The AR-SV shows competitive point forecasting performance, especially at longer horizons. The highly parametrized VARs generally achieve a lower degree of point forecasting accuracy than the benchmark. Tables 1a and 1b in the appendix report detailed results.

Moving to density forecast evaluation, Figure 18 reports the relative average log predictive scores for the chosen set of models and horizons. Alternative models' performance is reported in terms of log-scores differences with the benchmark MAI-AR-SV, so that negative values favor the MAI-AR-SV. The benchmark model clearly improves upon its competitors: the difference is negative and significant in most cases. Eventually, Figure 19 reports the CRPS reported comparatively as a ratio, where values greater than one indicates a worse density forecasting performance with respect to the MAI-AR-SV. Results are in line with the log-scores, with the benchmark model improving significantly upon its competitors[12].

To conclude, MAI-AR-SV is also a good forecasting model for inflation rates. The introduction of SV is particularly relevant to improve density forecasts. This evidence is in line with findings reported by Clark and Ravazzolo (2015) and D'Agostino et al. (2013). On the other hand, even though the AR-SV shows already good point and density forecasting power for inflation rates, the introduction of a MAI component proves to be quite beneficial. Furthermore, notwithstanding the smaller number of coefficients due to the reduced rank restriction imposed by the MAI structure, the benchmark model attains a higher degree of forecasting accuracy with respect to the standard unrestricted VAR estimated using a Minnesota Prior as shrinkage device.

6 Conclusions

Global developments play an important role in the determination of inflation rates, and indeed earlier literature has found that a substantial amount of the variation in a large

[12] More detailed forecasting results are reported in Appendix B.

set of national inflation rates can be explained by a single global factor. This literature has typically neglected inflation (conditional) volatility, while volatility is clearly relevant both from a policy point of view and for structural analysis and forecasting.

In this paper we study the evolution of inflation rates in many countries, using a novel model that allows for commonality in both levels and volatilities, in addition to country-specific components. We find that allowing for inflation volatility is indeed important, and a large fraction of it can be attributed to a global factor that is also driving the inflation levels.

While other sources can be behind this global factor, it turns out that since the early '90s it is strongly correlated with the Chinese PPI and Oil prices. Moreover, also the global factor stochastic volatility is highly correlated with that of Chinese PPI and Oil prices.

The MAI-AR-SV shows also very good out of sample properties, achieving comparatively better forecasting performances when compared with a set of prominent alternative models, especially in terms of density forecasting.

Bibliography

Abbate, A., Eickmeier, S., Lemke, W., and Marcellino, M. (2016). The changing international transmission of financial shocks: Evidence from a classical time-varying favar. *Journal of Money, Credit and Banking*, 48(4):573–601.

Auer, R. A., Borio, C. E., and Filardo, A. J. (2017). The globalisation of inflation: the growing importance of global value chains. *BIS Working Papers*, (n. 602).

Bai, J. and Ng, S. (2006). Evaluating latent and observed factors in macroeconomics and finance. *Journal of Econometrics*, 131(1):507–537.

Bańbura, M., Giannone, D., and Reichlin, L. (2010). Large bayesian vector auto regressions. *Journal of Applied Econometrics*, 25(1):71–92.

Baumeister, C., Liu, P., and Mumtaz, H. (2010). Changes in the transmission of monetary policy: Evidence from a time-varying factor-augmented var. *Bank of England Working Paper*.

Bauwens, L., Lubrano, M., and Richard, J.-F. (2000). *Bayesian inference in dynamic econometric models*. OUP Oxford.

Bernanke, B. S., Boivin, J., and Eliasz, P. (2005). Measuring the effects of monetary policy: a factor-augmented vector autoregressive (favar) approach. *The Quarterly journal of economics*, 120(1):387–422.

Bianchi, F. and Civelli, A. (2015). Globalization and inflation: Evidence from a time-varying var. *Review of Economic Dynamics*, 18(2):406–433.

Borio, C. E. and Filardo, A. J. (2007). Globalisation and inflation: New cross-country evidence on the global determinants of domestic inflation. *BIS Working Papers 227, Bank for International Settlements*.

Carney, M. (2017). Globalisation and inflation. Speech Given at the 2017 IMF Michel Camdessus Central Banking Lecture.

Carriero, A., Clark, T. E., and Marcellino, M. (2015). Bayesian vars: specification choices and forecast accuracy. *Journal of Applied Econometrics*, 30(1):46–73.

Carriero, A., Clark, T. E., and Marcellino, M. (2016a). Large vector autoregressions with stochastic volatility and flexible priors. *Federal Reserve Bank of Cleveland Working Paper, no. 16-17*.

Carriero, A., Clark, T. E., and Marcellino, M. (2017). Measuring uncertainty and its impact on the economy. *The Review of Economics and Statistics*.

Carriero, A., Corsello, F., and Marcellino, M. (2018). The economic drivers of volatility and uncertainty. *mimeo*.

Carriero, A., Kapetanios, G., and Marcellino, M. (2011). Forecasting large datasets with bayesian reduced rank multivariate models. *Journal of Applied Econometrics*, 26(5):735–761.

Carriero, A., Kapetanios, G., and Marcellino, M. (2016b). Structural analysis with multivariate autoregressive index models. *Journal of Econometrics*, 192(2):332–348.

Carter, C. K. and Kohn, R. (1994). On Gibbs Sampling for State Space Models. *Biometrika*, 81(3):541–553.

Centoni, M. and Cubadda, G. (2003). Measuring the business cycle effects of permanent and transitory shocks in cointegrated time series. *Economics Letters*, 80(1):45–51.

Ciccarelli, M. and Mojon, B. (2010). Global inflation. *The Review of Economics and Statistics*, 92(3):524–535.

Clark, T. E. and Ravazzolo, F. (2015). Macroeconomic forecasting performance under alternative specifications of time-varying volatility. *Journal of Applied Econometrics*, 30(4):551–575.

Cogley, T. and Sargent, T. J. (2005). Drifts and volatilities: monetary policies and outcomes in the post wwii us. *Review of Economic dynamics*, 8(2):262–302.

Corsello, F. (2018). Rank Reduction in BVARs with Time Varying Coefficients and Stochastic Volatility: Specification and Estimation. *Bocconi University, mimeo*.

Cubadda, G. and Guardabascio, B. (2017). Representation, estimation and forecasting of the multivariate index-augmented autoregressive model. *CEIS Research Paper 397, Tor Vergata University, CEIS*.

D'Agostino, A., Gambetti, L., and Giannone, D. (2013). Macroeconomic forecasting and structural change. *Journal of Applied Econometrics*, 28(1):82–101.

Del Negro, M. and Otrok, C. (2008). Dynamic factor models with time-varying parameters: measuring changes in international business cycles. *mimeo*.

Del Negro, M. and Primiceri, G. E. (2015). Time varying structural vector autoregressions and monetary policy: a corrigendum. *The review of economic studies*, 82(4):1342–1345.

Delle Monache, D., Petrella, I., and Venditti, F. (2016). Common faith or parting ways? a time varying parameters factor analysis of euro-area inflation. In *Dynamic Factor Models*, pages 539–565. Emerald Group Publishing Limited.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–263.

Doan, T., Litterman, R., and Sims, C. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric reviews*, 3(1):1–100.

Dufour, J.-M. and Stevanović, D. (2013). Factor-augmented varma models with macroeconomic applications. *Journal of Business & Economic Statistics*, 31(4):491–506.

Eickmeier, S., Lemke, W., and Marcellino, M. (2015). Classical time varying factor-augmented vector auto-regressive models—estimation, forecasting and structural analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(3):493–533.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007.

Ferroni, F. and Mojon, B. (2016). Domestic and global inflation. Technical report, mimeo.

Forni, M. and Gambetti, L. (2010). The dynamic effects of monetary policy: A structural factor model approach. *Journal of Monetary Economics*, 57(2):203–216.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2000). The generalized dynamic-factor model: Identification and estimation. *The review of Economics and Statistics*, 82(4):540–554.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2014). *Bayesian data analysis*, volume 2. CRC press Boca Raton, FL.

Geweke, J. (1977). The dynamic factor analysis of economic time series. *Latent variables in socio-economic models*.

Geweke, J. (1996). Bayesian reduced rank regression in econometrics. *Journal of econometrics*, 75(1):121–146.

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

Jacquier, E., Polson, N. G., and Rossi, P. E. (2004). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics*, 122(1):185–212.

Jacquier, P. and Polson, N. (1994). Rossi,". *J. Business and Economics Statistics*, 124.

Johansen, S. (1995). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford University Press on Demand.

Jordà, Ò. (2005). Estimation and inference of impulse responses local projections. *American economic review*, 95(1):161–182.

Jurado, K., Ludvigson, S. C., and Ng, S. (2015). Measuring uncertainty. *The American Economic Review*, 105(3):1177–1216.

Kadiyala, K. R. and Karlsson, S. (1997). Numerical methods for estimation and inference in bayesian var-models. *Journal of Applied Econometrics*, pages 99–132.

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with arch models. *The review of economic studies*, 65(3):361–393.

Kleibergen, F. and Van Dijk, H. K. (1998). Bayesian simultaneous equations analysis using reduced rank structures. *Econometric Theory*, 14(6):701–743.

Koop, G. M. (2013). Forecasting with medium and large bayesian vars. *Journal of Applied Econometrics*, 28(2):177–203.

Kose, M. A., Otrok, C., and Whiteman, C. H. (2003). International business cycles: World, region, and country-specific factors. *The American Economic Review*, 93(4):1216–1239.

Litterman, R. B. (1986). Forecasting with bayesian vector autoregressions—five years of experience. *Journal of Business & Economic Statistics*, 4(1):25–38.

Lodge, D. and Mikolajun, I. (2016). Advanced economy inflation: The role of global factors. Technical report, ECB, Working Paper Series.

Lütkepohl, H. (1984). Linear transformations of vector arma processes. *Journal of Econometrics*, 26(3):283–293.

Lütkepohl, H. (1990). Asymptotic distributions of impulse response functions and forecast error variance decompositions of vector autoregressive models. *The review of economics and statistics*, pages 116–125.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.

Mumtaz, H. and Surico, P. (2008). Evolving international inflation dynamics: Evidence from a time-varying dynamic factor model. *CEPR Discussion Papers*, (N. 6767).

Omori, Y., Chib, S., Shephard, N., and Nakajima, J. (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics*, 140(2).

O'Reilly, G. and Whelan, K. (2005). Has euro-area inflation persistence changed over time? *The Review of Economics and Statistics*, 87(4):709–720.

Primiceri, G. E. (2005). Time Varying Structural Vector Autoregressions. *The Review of Economic Studies*, 72(3):821–852.

Reinsel, G. (1983). Some results on multivariate autoregressive index models. *Biometrika*, 70(1):145–156.

Reinsel, G. and Velu, R. (1998). Multivariate reduced-rank regression, volume 136 of lectures notes in statistics. New York: Springer-Verlag.

Rogoff, K. (2003). Globalization and global disinflation. In *A symposium sponsored by the Federal Reserve Bank of Kansas City on "Monetary Policy and Uncertainty: Adapting to a Changing Economy"*, volume 88, page 45. Federal Reserve Bank of Kansas City.

Rogoff, K. (2006). Impact of globalization on monetary policy. In *A symposium sponsored by the Federal Reserve Bank of Kansas City on "The new economic geography: effects and policy implications", Jackson Hole, Wyoming*, pages 24–26.

Stock, J. H. (1991). Confidence intervals for the largest autoregressive root in us macroeconomic time series. *Journal of monetary economics*, 28(3):435–459.

Stock, J. H. and Watson, M. W. (1989). New indexes of coincident and leading economic indicators. *NBER macroeconomics annual*, 4:351–394.

Stock, J. H. and Watson, M. W. (2002a). Forecasting using principal components from a large number of predictors. *Journal of the American statistical association*, 97(460):1167–1179.

Stock, J. H. and Watson, M. W. (2002b). Macroeconomic forecasting using diffusion indexes. *Journal of Business & Economic Statistics*, 20(2):147–162.

Stock, J. H. and Watson, M. W. (2007). Why has us inflation become harder to forecast? *Journal of Money, Credit and banking*, 39(s1):3–33.

Stock, J. H. and Watson, M. W. (2010). Modeling inflation after the crisis. Technical report, National Bureau of Economic Research.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*. Wiley.

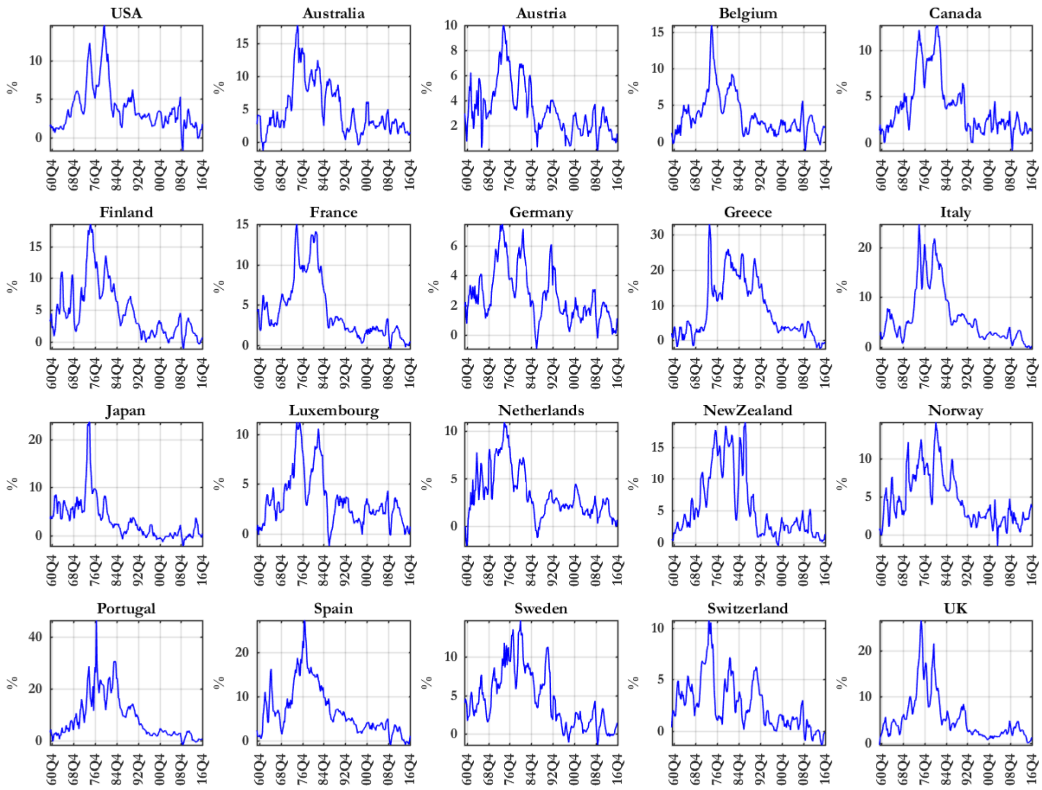


Figure 1: Inflation rates (year on year growth rates in quarterly CPIs)

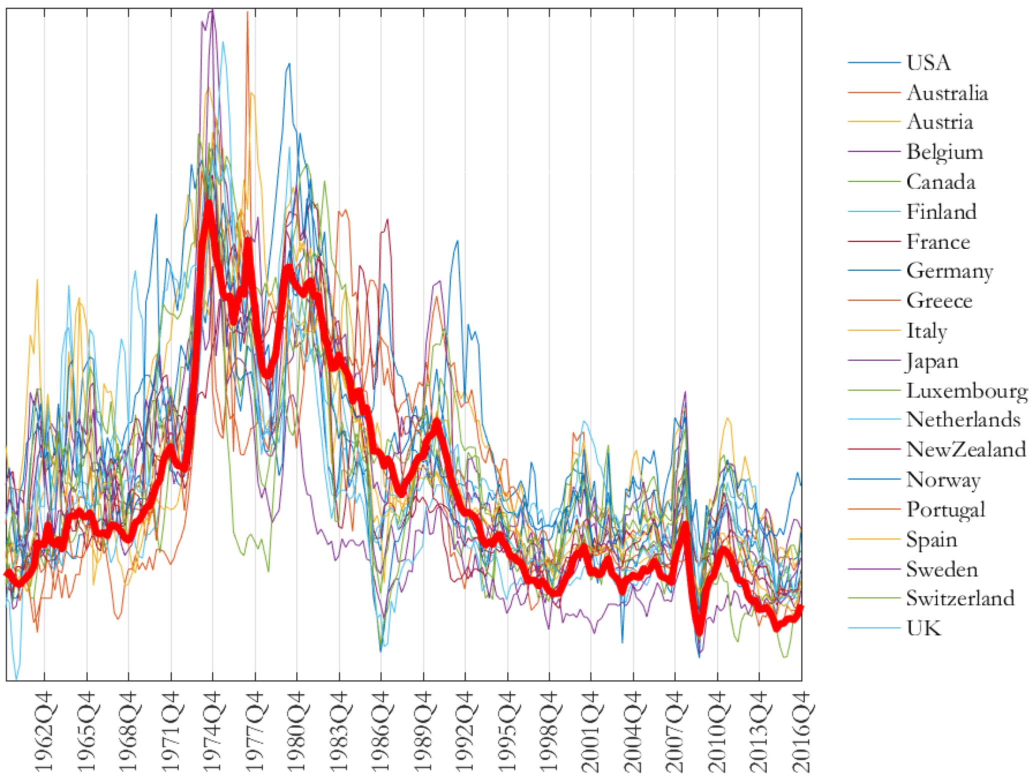


Figure 2: Inflation rates and their first principal component (thick red line)

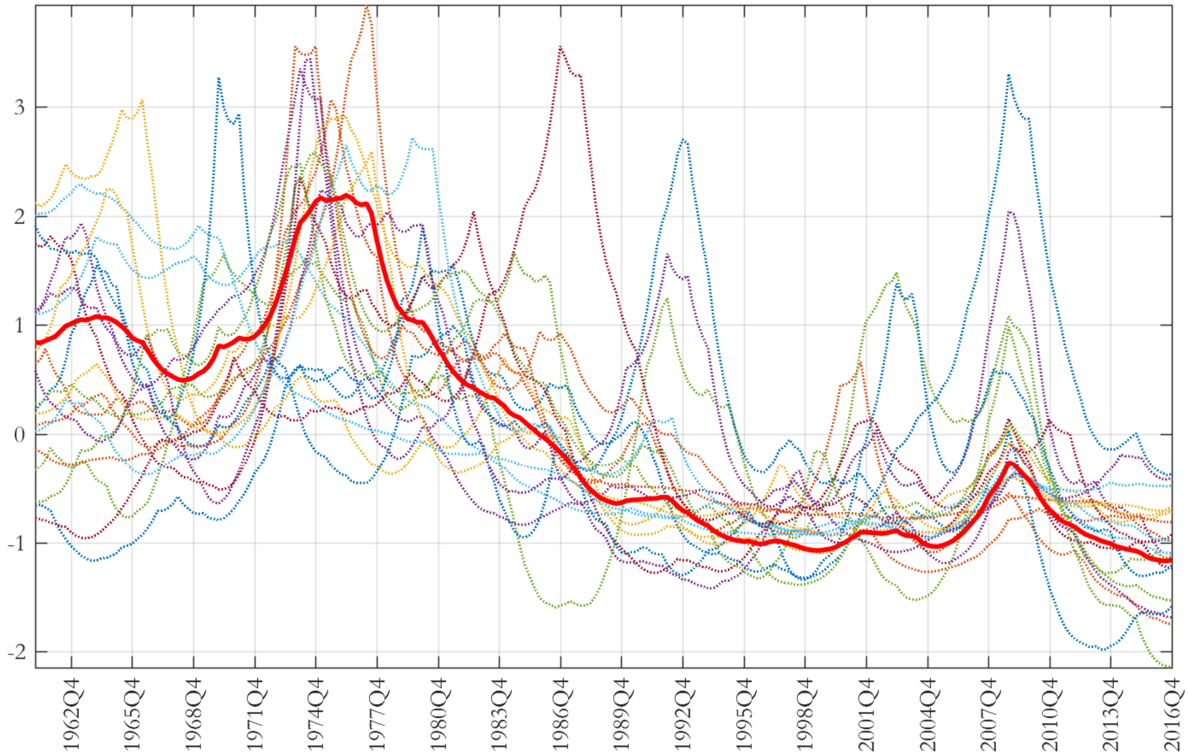


Figure 3: CPI inflation rates Stochastic Volatilities estimated from univariate AR-SV, and their first Principal Component (thick red line)

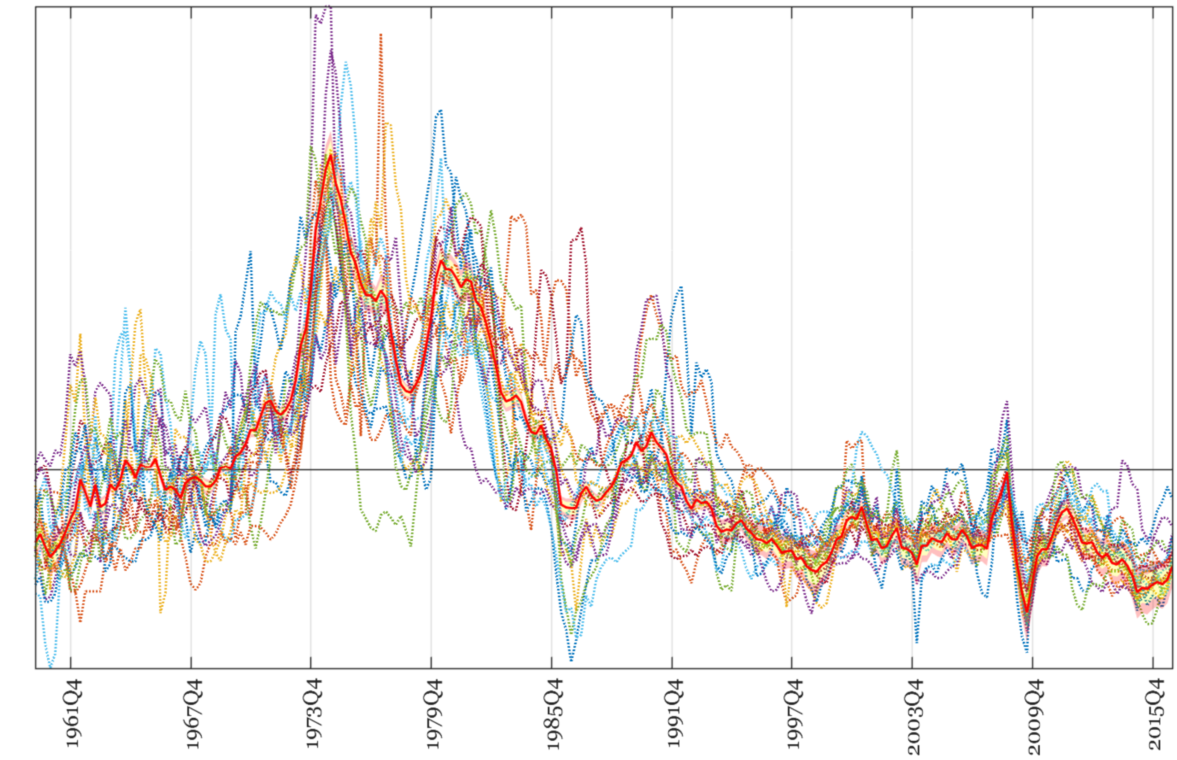


Figure 4: MAI-SV estimated common factor (with posterior bands) Vs Data

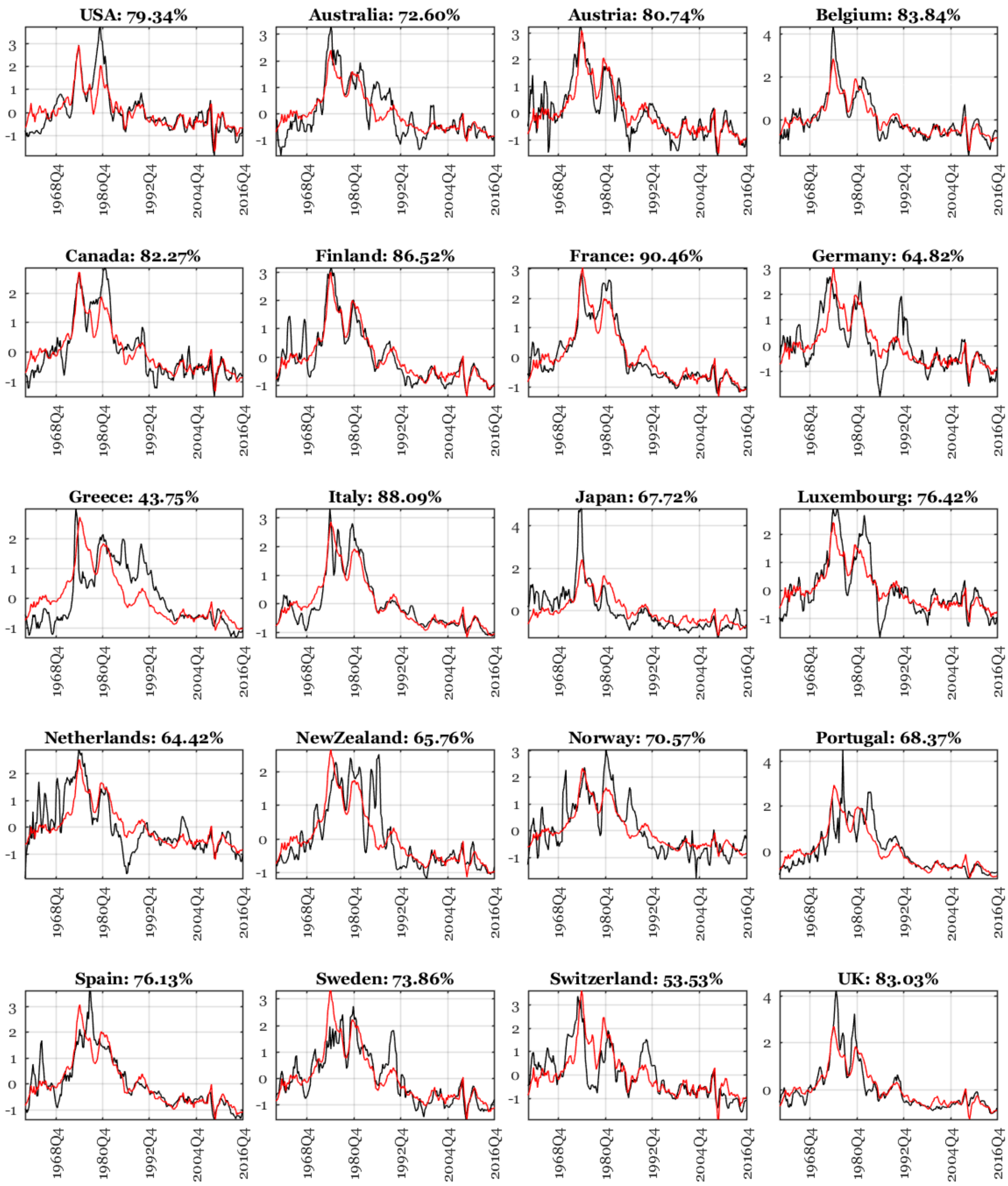


Figure 5: MAI-SV in-sample fit

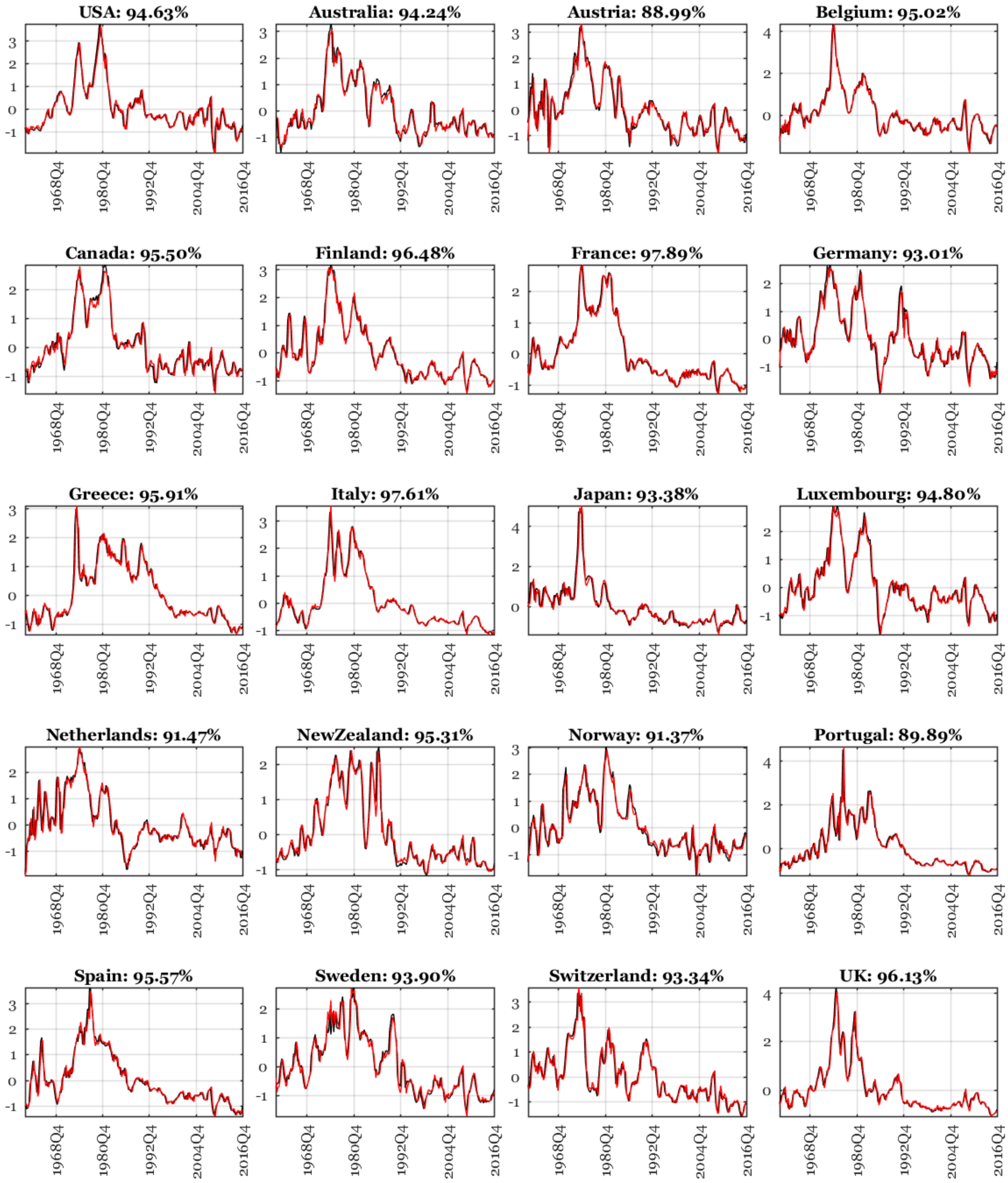


Figure 6: MAI-AR-SV in-sample fit

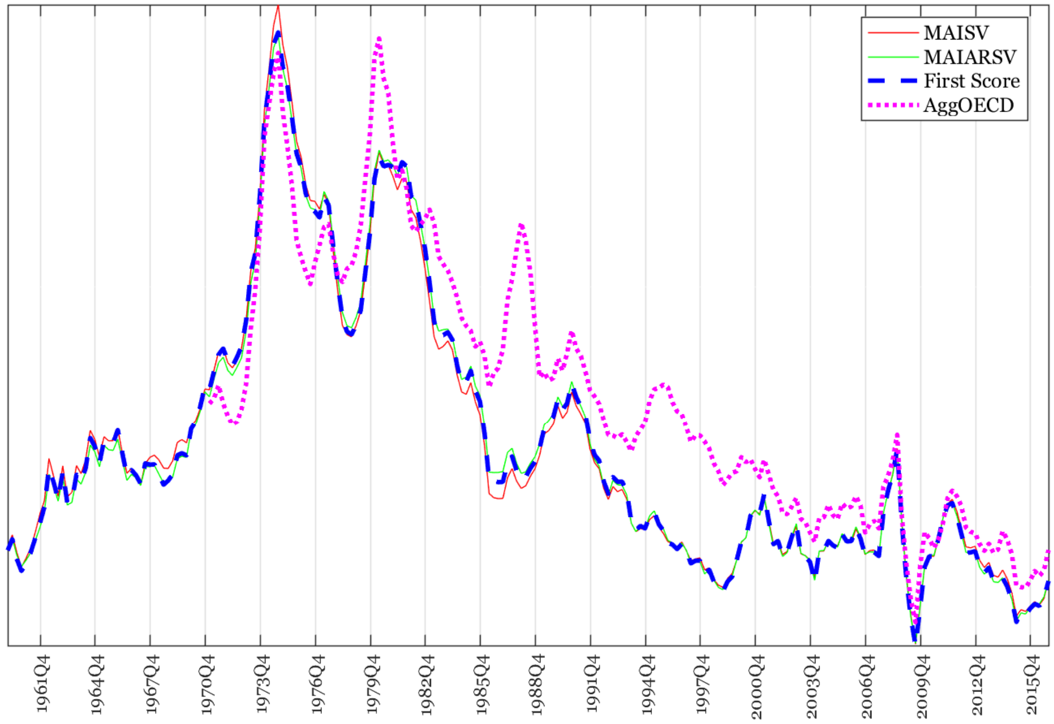


Figure 7: Comparing common factor

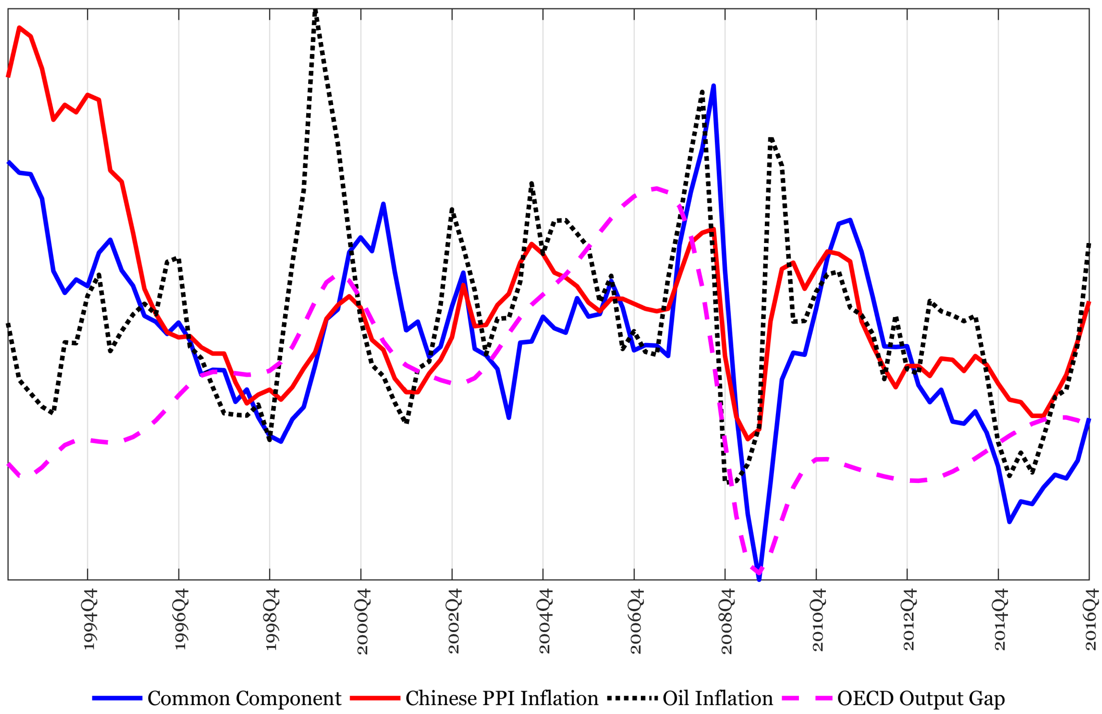


Figure 8: Comparing MAI component with Chinese PPI, Oil Inflation and Global Output Gap

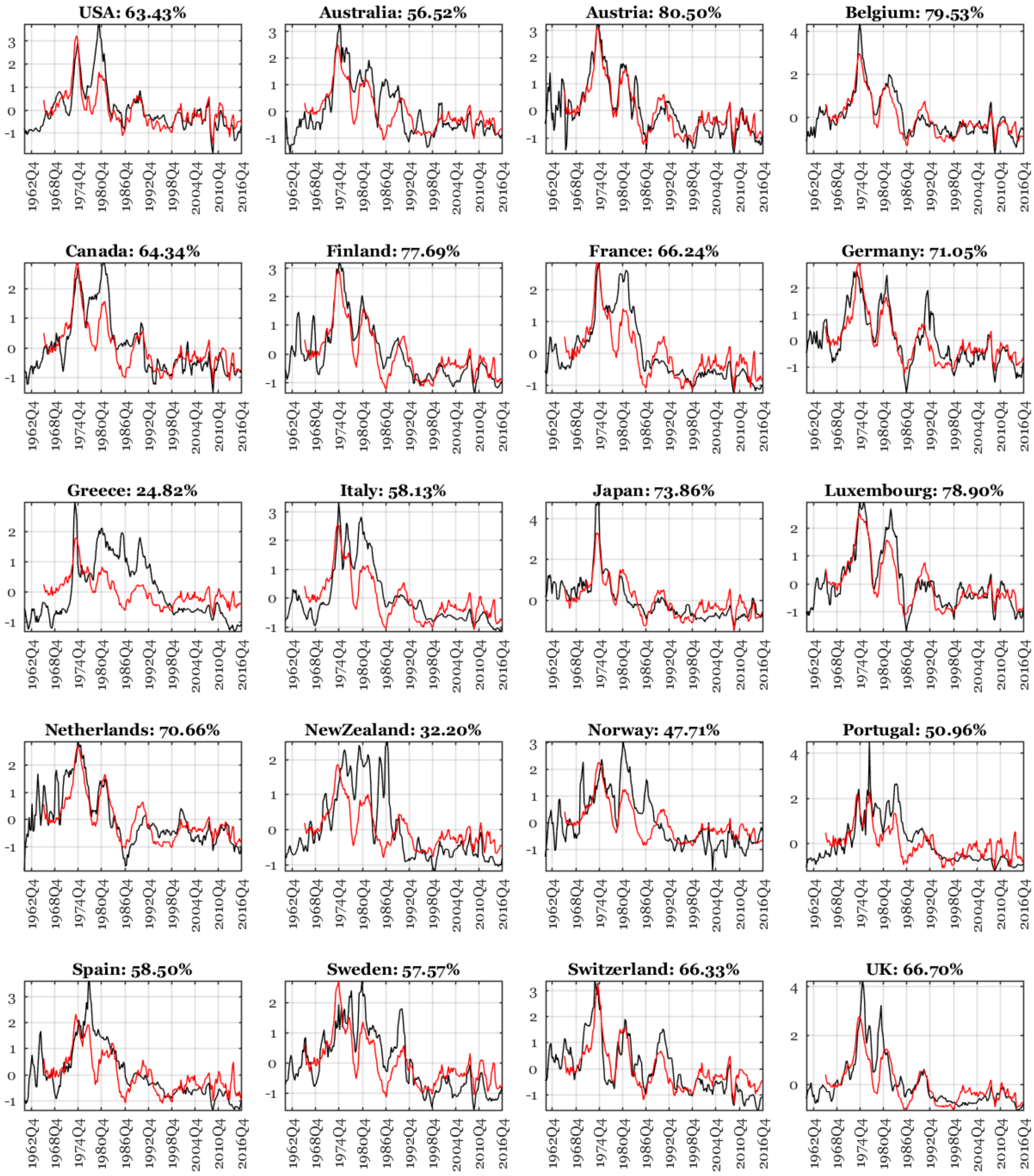


Figure 9: MAI-AR-SV, Actual series and Common component (red)

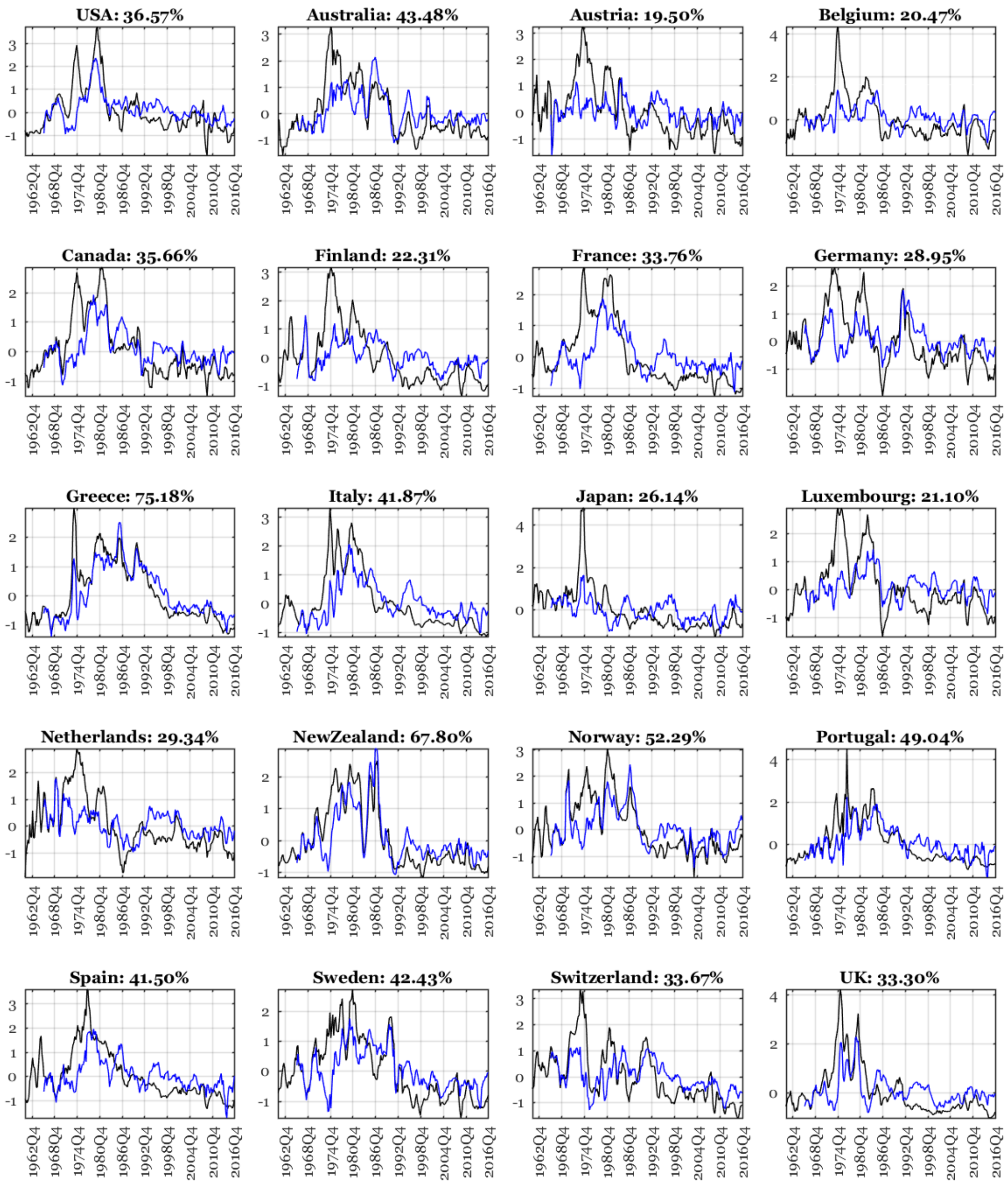


Figure 10: MAI-AR-SV, Actual series and Idiosyncratic component (blue)

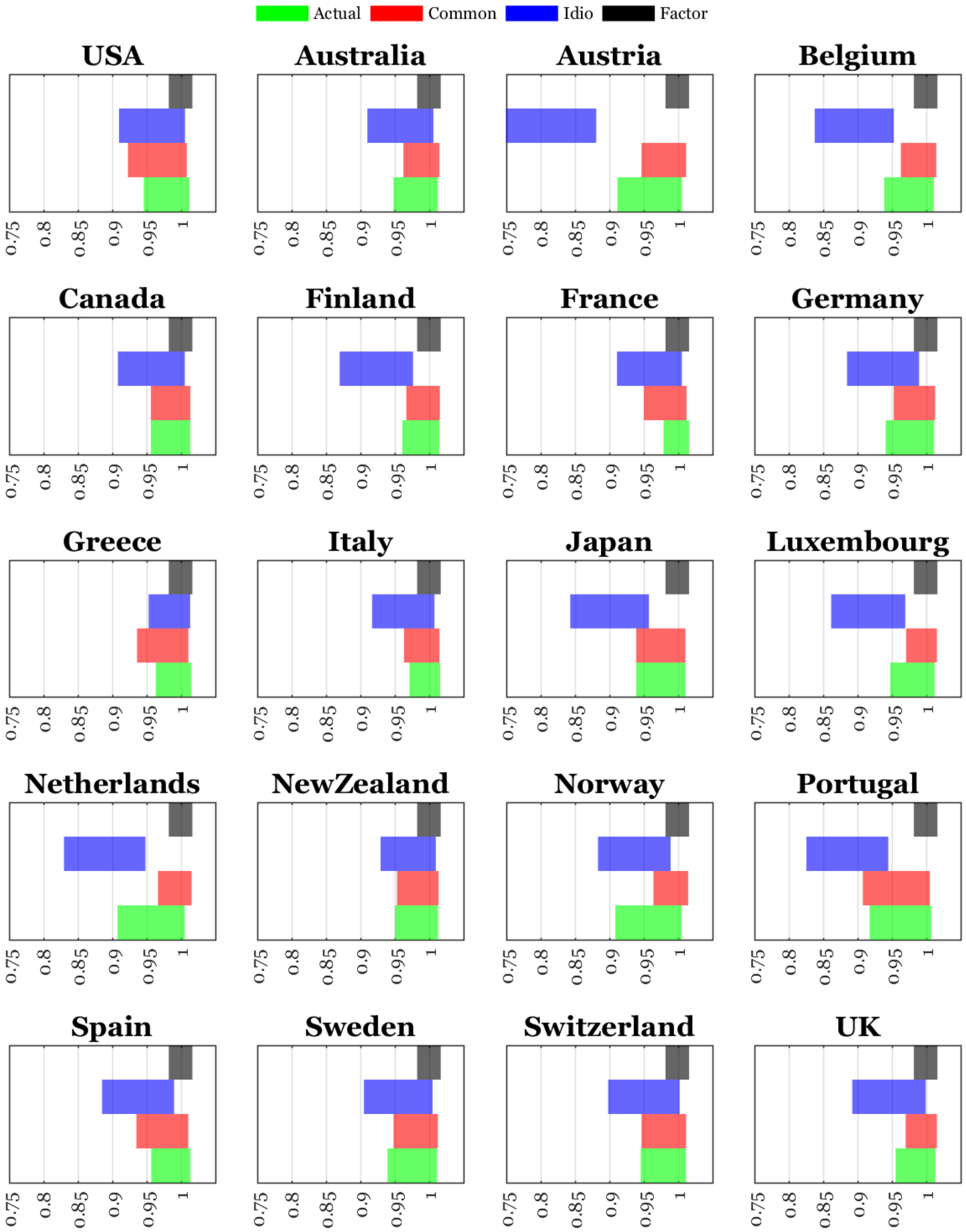


Figure 11: Largest Autoregressive Root (90% confidence intervals) CPI inflation levels, components, and global factor.

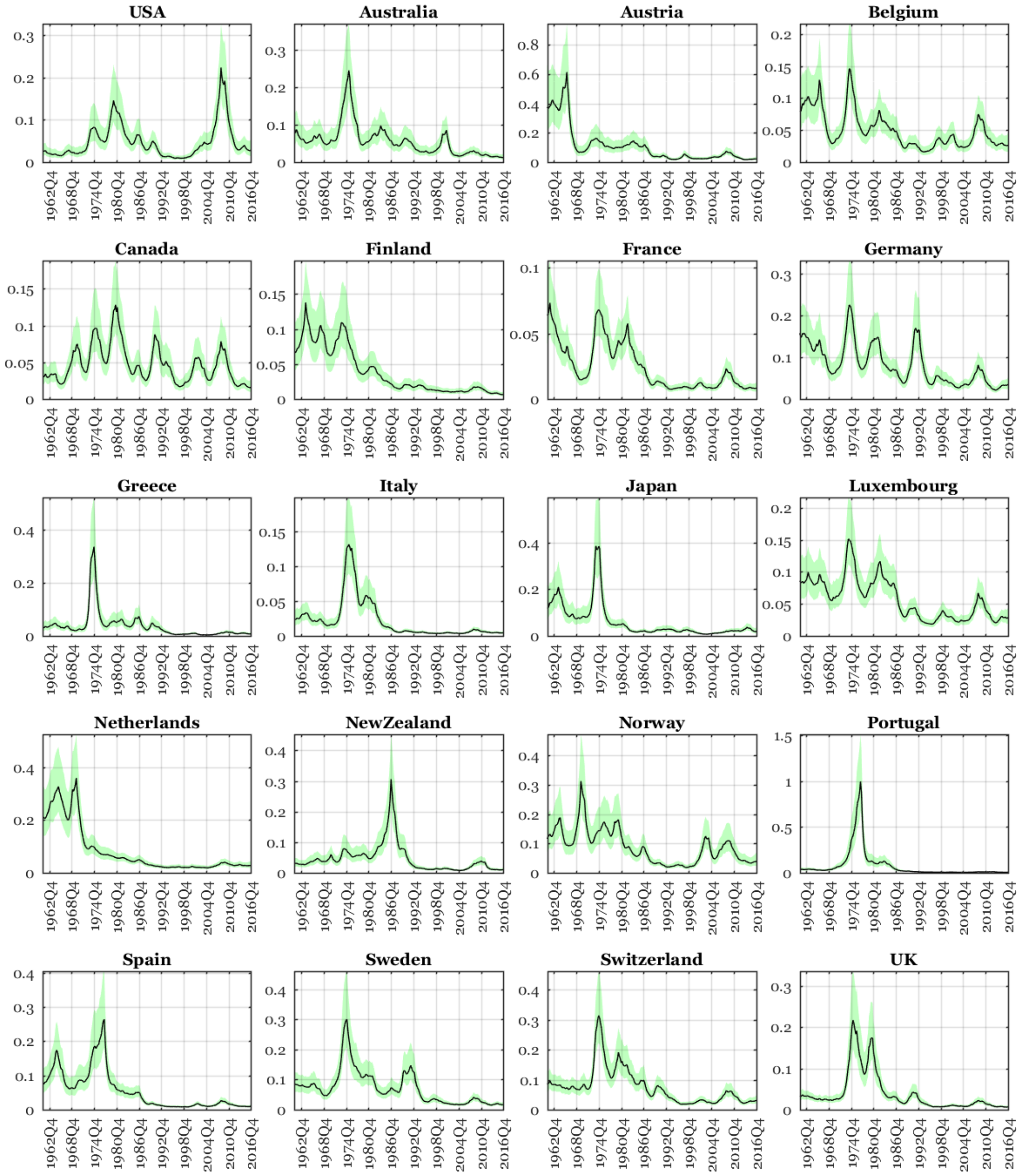


Figure 12: MAI-AR-SV, Residuals' Volatility, posterior bands

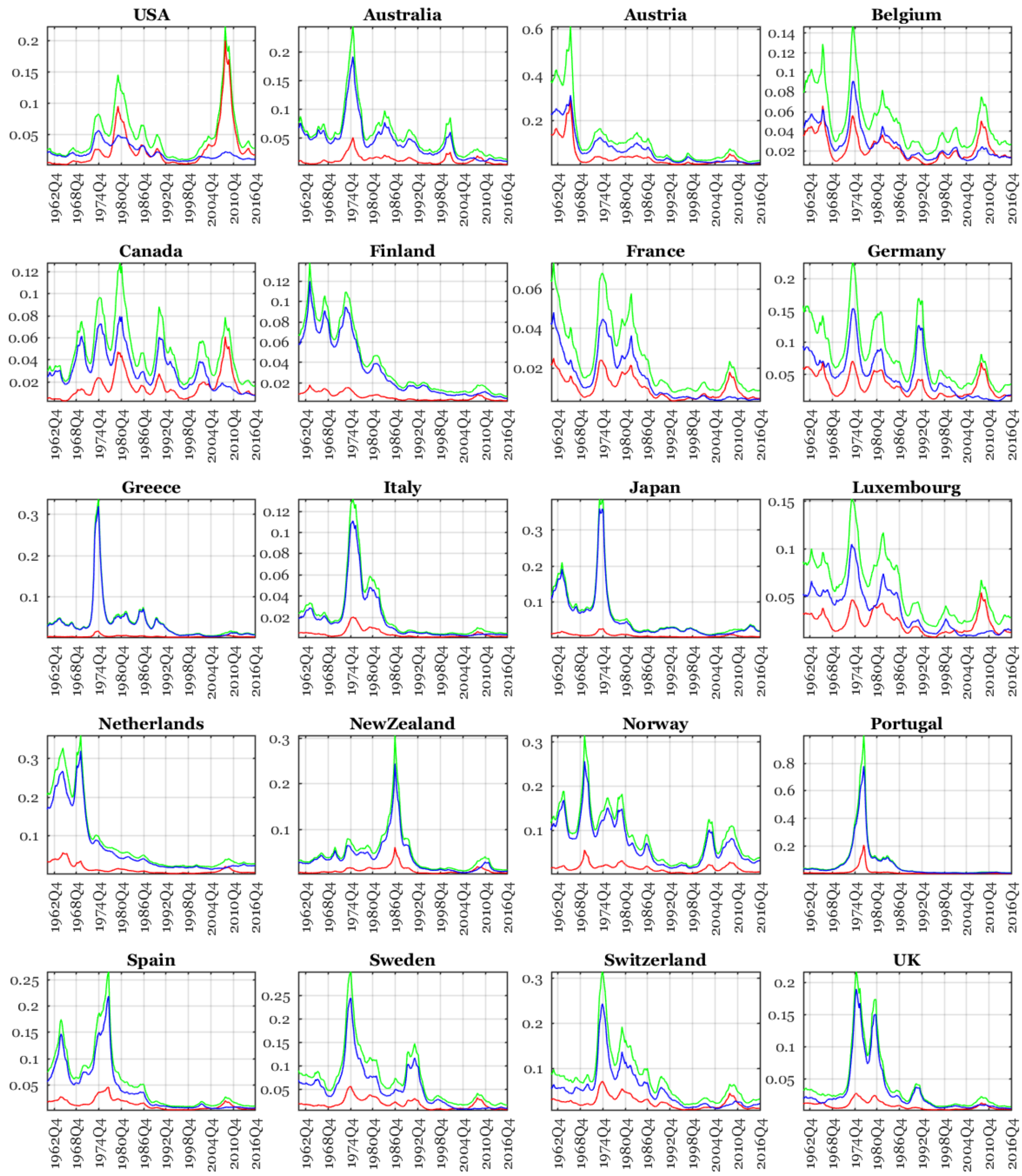


Figure 13: MAI-AR-SV, Residuals' Volatility, TV decomposition, Common (red), Idio (blue), total (green)

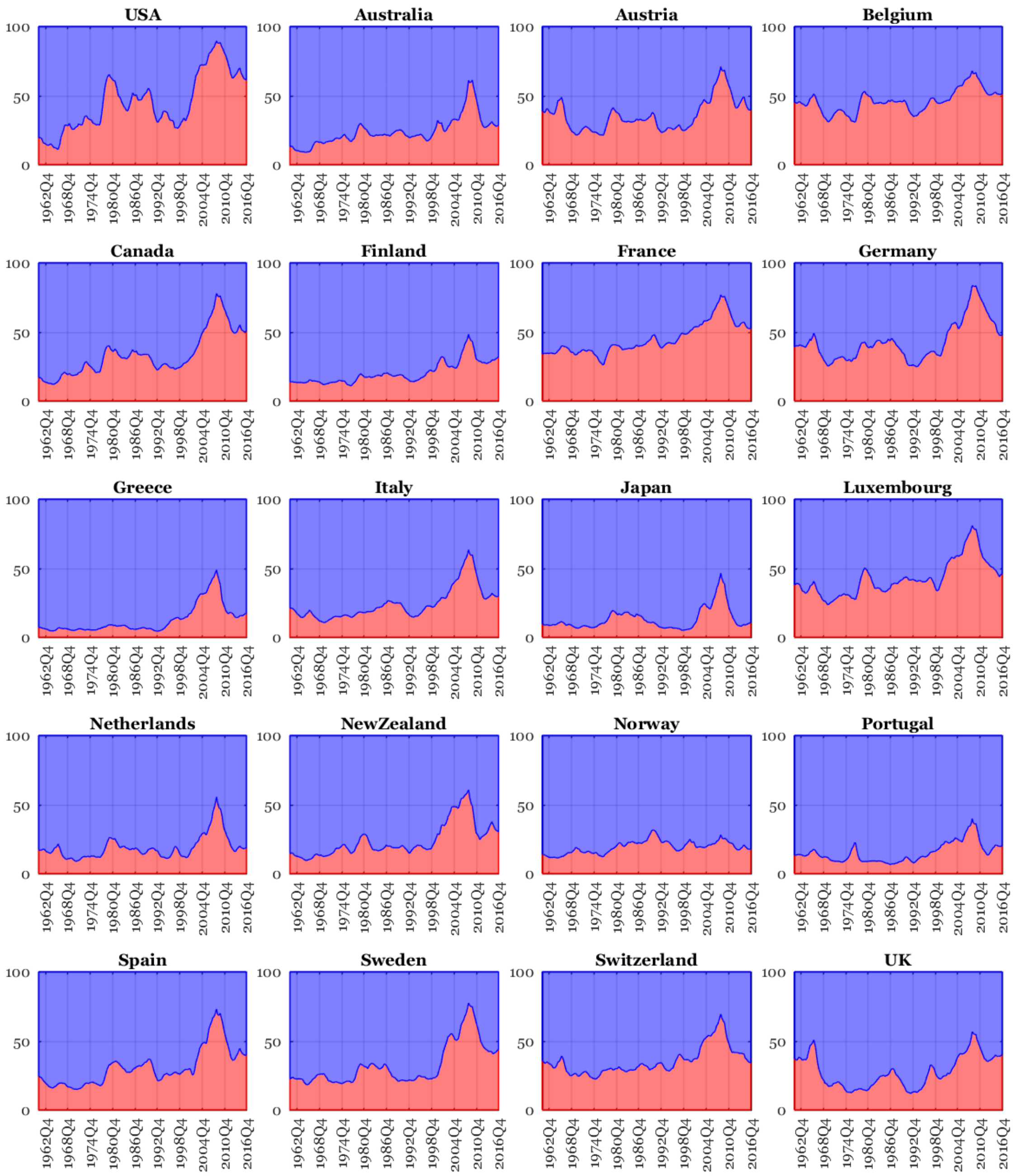


Figure 14: MAI-AR-SV, Residuals' Volatility, TV decomposition shares (%), Common (red), Idio (blue)

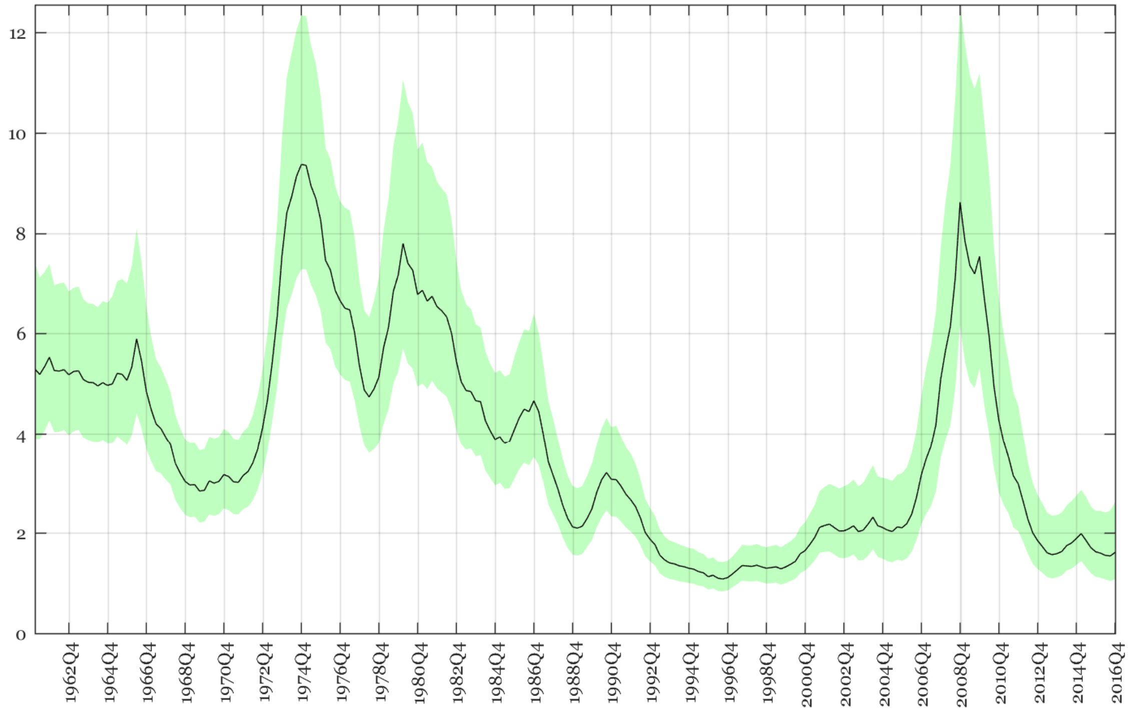


Figure 15: MAI-AR-SV, Global Factor Volatility, Posterior bands.

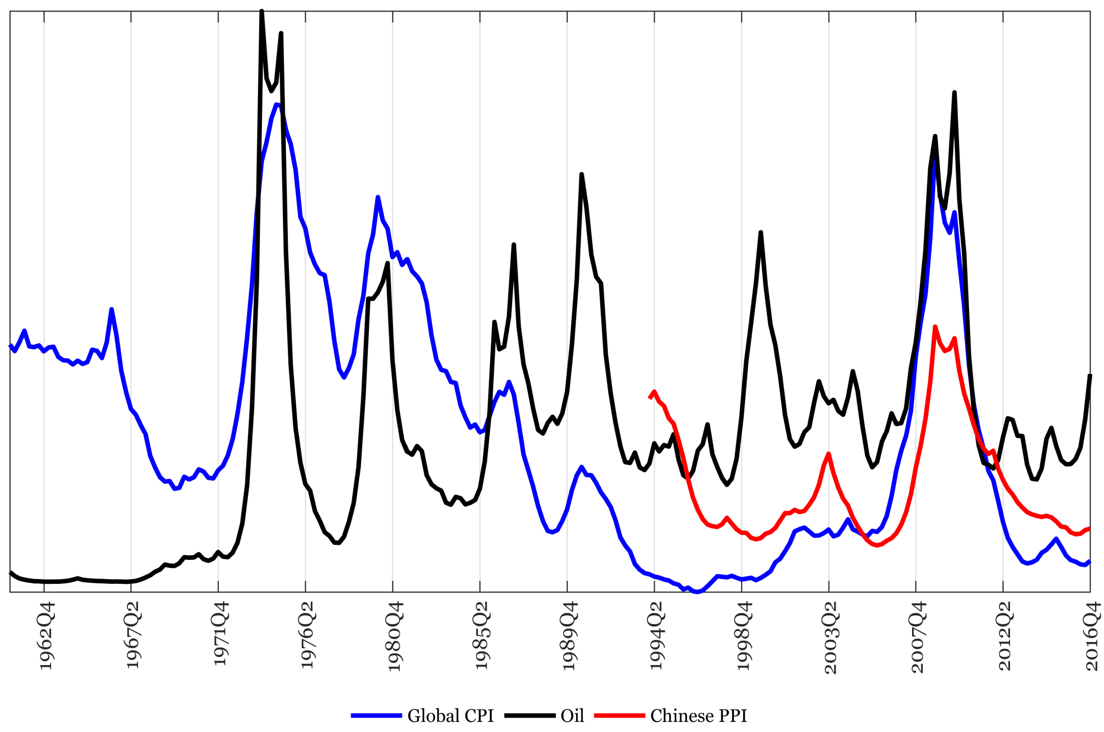


Figure 16: Median Volatilities of Global Factor (MAI-AR-SV), Oil inflation (AR-SV) and Chinese PPI (AR-SV)

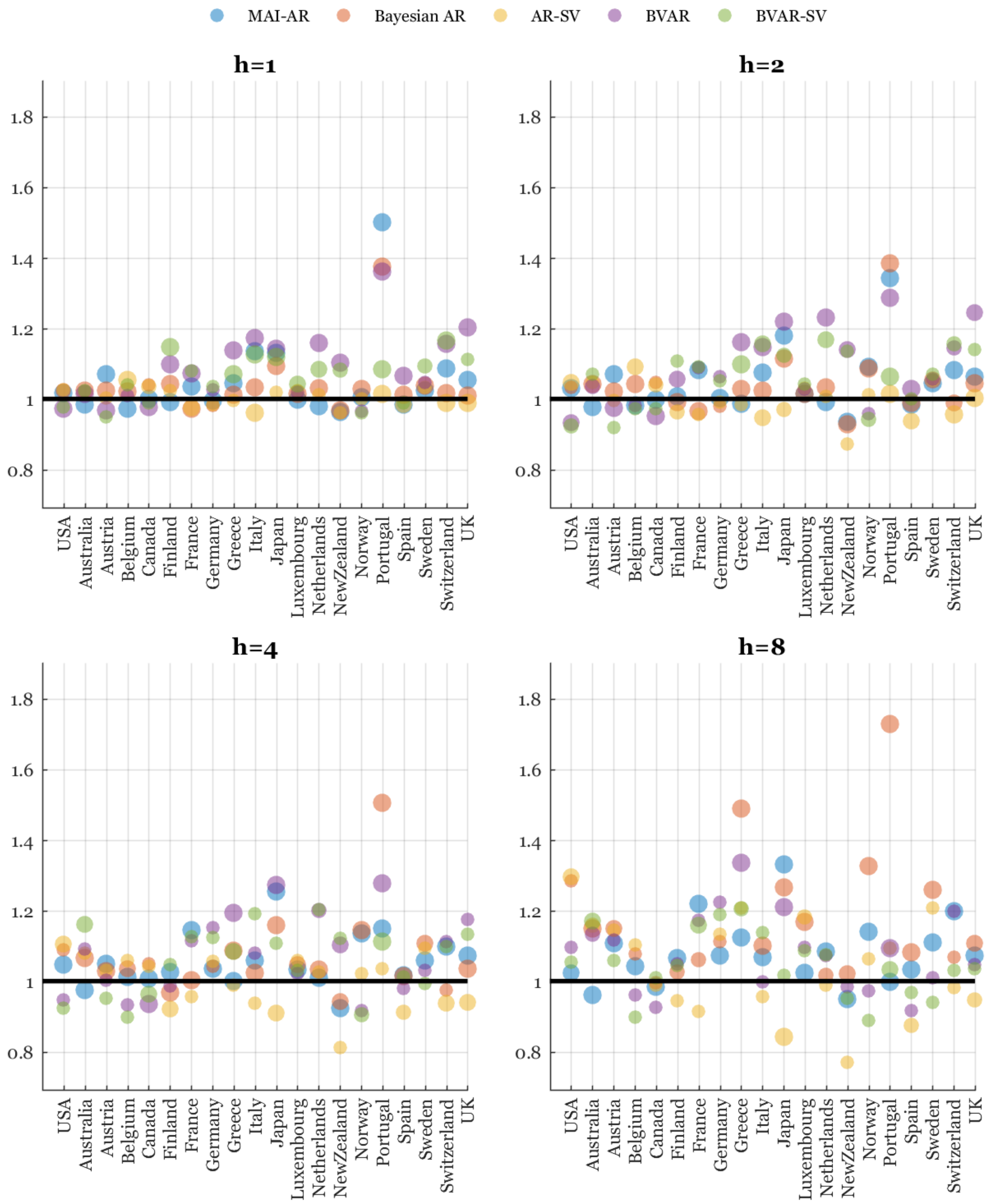
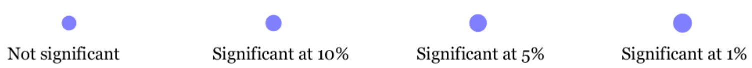


Figure 17: Relative Root Mean Squared Forecast Errors (ratios with MAI-AR-SV)

The round filled marker is larger when the difference is significant according to the Diebold Mariano t-statistic, see legend below.



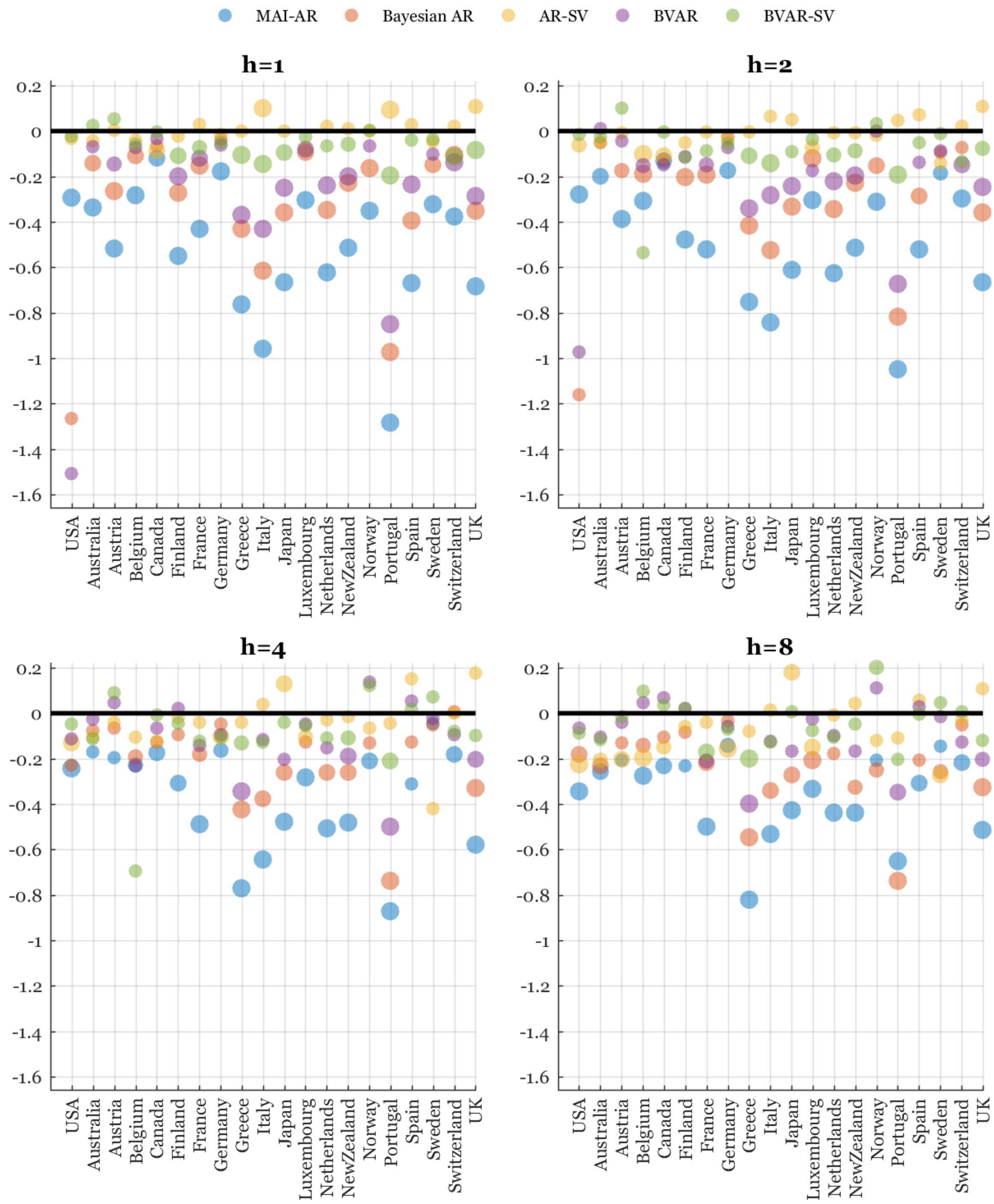
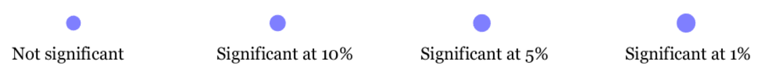


Figure 18: Relative Log Predictive Scores (differences with MAI-AR-SV)

The round filled marker is larger when the difference is significant according to the Diebold Mariano t-statistic, see legend below.



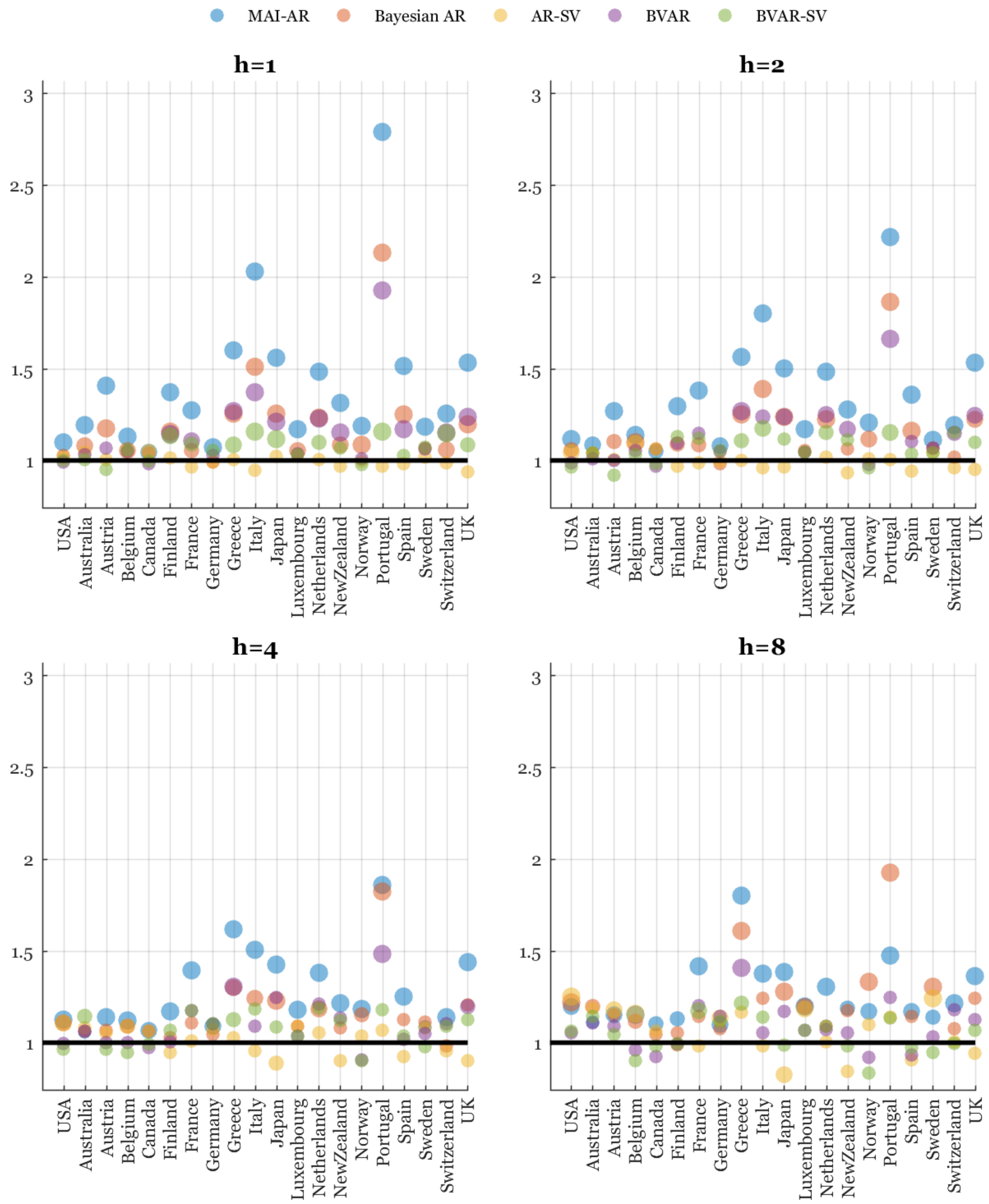
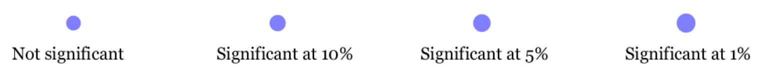


Figure 19: Relative Continuous Rank Probability Scores (ratios with MAI-AR-SV)

The round filled marker is larger when the difference is significant according to the Diebold Mariano t-statistic, see legend below.



A Gibbs Sampler for estimation of the MAI-AR-SV model

A.1 Step 1: draw AR-coefficients γ

The AR components included in the MAI-AR-SV are stacked in the following way:

$$y_t = \sum_{\ell=1}^{q} \Gamma_\ell \cdot y_{t-\ell} + A \cdot Z_t + u_t,$$

$$y_t = \sum_{\ell=1}^{q} \begin{bmatrix} \gamma_{1,\ell} & 0 & \dots & 0 \\ 0 & \gamma_{2,\ell} & \ddots & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \gamma_{n,\ell} \end{bmatrix} \cdot y_{t-\ell} + A \cdot Z_t + u_t,$$

$$y_t = \sum_{\ell=1}^{q} \begin{bmatrix} y_{1,t-\ell} & 0 & \dots & 0 \\ 0 & y_{2,t-\ell} & \ddots & 0 \\ 0 & \ddots & \ddots & 0 \\ 0 & \dots & 0 & y_{n,t-\ell} \end{bmatrix} \cdot \begin{bmatrix} \gamma_{1,\ell} \\ \gamma_{2,\ell} \\ \vdots \\ \gamma_{n,\ell} \end{bmatrix} + A \cdot Z_t + u_t,$$

$$y_t = \sum_{\ell=1}^{q} \mathcal{Y}_{t-\ell} \cdot \gamma_\ell + A \cdot Z_t + u_t,$$

$$y_t = \begin{bmatrix} \mathcal{Y}_{t-1} & \mathcal{Y}_{t-2} & \dots & \mathcal{Y}_{t-q} \end{bmatrix} \cdot \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_q \end{bmatrix} + A \cdot Z_t + u_t,$$

so that we can eventually write a fully stacked form:

$$y_t = \underbrace{\mathcal{X}_t}_{n \times nq} \cdot \underbrace{\gamma}_{nq \times 1} + \underbrace{A}_{n \times rp} \cdot \underbrace{Z_t}_{rp \times 1} + u_t.$$

Given B_0 and A, we can transform the matrix form to have the following linear regression model with common coefficients and variable specific regressors:

$$\begin{aligned}
y_t &= \mathcal{X}_t \cdot \gamma + \ A \cdot \ Z_t + u_t, \\
y_t - A \cdot Z_t &= \mathcal{X}_t \cdot \gamma + u_t, \\
y_t^\circ &= \mathcal{X}_t \cdot \gamma + u_t,
\end{aligned}$$

with

$$u_t \overset{i}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{\Omega_t}_{n\times n}\right), \qquad u_t = G^{-1}\Sigma_t\varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} \mathcal{MN}\left(\ \mathbf{0},\ I_n\ \right).$$

Considering separate equations to estimate the AR coefficients contained in γ would ignore the cross-correlations of the innovations in u_t. Considering that within the GS we draw directly the elements g in the matrix G and the stochastic volatilities σ_t in Σ_t, for efficiency purposes we can compute the following transformation of the equation:

$$\begin{aligned}
y_t^\circ &= \mathcal{X}_t \cdot \gamma + u_t, \\
y_t^\circ &= \mathcal{X}_t \cdot \gamma + G^{-1}\Sigma_t\varepsilon_t, \\
\Sigma_t^{-1}G \cdot y_t^\circ &= \Sigma_t^{-1}G \cdot \mathcal{X}_t \cdot \gamma + \underbrace{\Sigma_t^{-1}G \cdot G^{-1}\Sigma_t}_{I_n} \cdot \varepsilon_t, \\
\Sigma_t^{-1}G \cdot y_t^\circ &= \Sigma_t^{-1}G \cdot \mathcal{X}_t \cdot \gamma + \varepsilon_t, \\
\widetilde{y_t^\circ} &= \widetilde{\mathcal{X}_t} \cdot \gamma + \varepsilon_t.
\end{aligned}$$

Finally obtaining a multivariate linear regression with homoskedastic residuals and unitary diagonal covariance matrix:

$$\widetilde{y_t^\circ} = \widetilde{\mathcal{X}_t} \cdot \gamma + \varepsilon_t, \qquad \varepsilon_t \overset{iid}{\sim} \mathcal{MN}\left(\ \mathbf{0},\ I_n\ \right).$$

The transformed model can be stacked in columns:

$$\begin{bmatrix} \widetilde{y}_1^{\circ} \\ \vdots \\ \widetilde{y}_T^{\circ} \end{bmatrix} = \begin{bmatrix} \widetilde{\mathcal{X}_1} \\ \vdots \\ \widetilde{\mathcal{X}_T} \end{bmatrix} \cdot \gamma + \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_T \end{bmatrix},$$

$$\underbrace{\widetilde{Y}^{\circ}}_{nT\times 1} = \underbrace{\widetilde{\mathcal{X}}}_{nT\times nq} \cdot \underbrace{\gamma}_{nq\times 1} + \underbrace{\varepsilon^{\circ}}_{nT\times 1}, \qquad \varepsilon^{\circ} \sim \mathcal{MN}\left(\,\mathbf{0},\ I_T \otimes I_n\,\right).$$

With the stacked version of the model, adopting the Normal conjugate prior for coefficients γ:

$$\gamma \sim \mathcal{MN}\left(\,\bar{\gamma},\ V_\gamma\,\right),$$

we can eventually draw from the posterior of γ:

$$\gamma \sim \mathcal{MN}\left(\,\widetilde{\gamma},\ \widetilde{V}_\gamma\,\right),$$

where

$$\widetilde{\gamma} = \widetilde{V}_\gamma \cdot \left(\widetilde{\mathcal{X}}' \cdot \widetilde{Y}^{\circ} + V_\gamma^{-1} \cdot \bar{\gamma}\right), \qquad \widetilde{V}_\gamma = \left(\widetilde{\mathcal{X}}' \cdot \widetilde{\mathcal{X}} + V_\gamma^{-1}\right)^{-1}.$$

A.2 Step 2: Draw loadings A

The second step of the GS aims at drawing the loadings contained in A. Recalling the following:

$$A \equiv \begin{bmatrix} A_1 & \dots & A_p \end{bmatrix}, \qquad x_t \equiv vec\left(x_t^{\bullet}\right), \qquad \underbrace{x_t^{\bullet}}_{n\times p} \equiv \begin{bmatrix} y_{t-1} & \dots & y_{t-p} \end{bmatrix},$$

$$Z_t \equiv \begin{bmatrix} B_0 y_{t-1} \\ \vdots \\ B_0 y_{t-p} \end{bmatrix} = \left(I_p \otimes B_0\right) \cdot x_t = vec\left(B_0 \cdot x_t^{\bullet}\right),$$

the model can be restated as:

$$y_t - \mathcal{X}_t \cdot \gamma = \begin{bmatrix} A_1 & \dots & A_p \end{bmatrix} \begin{bmatrix} B_0 y_{t-1} \\ \vdots \\ B_0 y_{t-p} \end{bmatrix} + u_t,$$

$$y_t^\bullet = \underbrace{A}_{n \times rp} \cdot \underbrace{Z_t}_{rp \times 1} + u_t,$$

and can be stacked as:

$$\begin{bmatrix} y_1^{\bullet\prime} \\ y_2^{\bullet\prime} \\ \vdots \\ y_T^{\bullet\prime} \end{bmatrix} = \begin{bmatrix} Z_1' \\ Z_2' \\ \vdots \\ Z_T' \end{bmatrix} A' + \begin{bmatrix} u_1' \\ u_2' \\ \vdots \\ u_T' \end{bmatrix},$$

$$\underbrace{y^\bullet}_{T \times n} = \underbrace{Z}_{T \times rp} \cdot \underbrace{A'}_{rp \times n} + u.$$

Defining $a \equiv vec\,(A')$, and exploiting the Kronecker product's properties, this form can be vectorized and transformed in:

$$vec\,(y^\bullet) = vec\,(Z \cdot A' \cdot I_n) + vec\,(u)\,,$$

$$\underbrace{Y^\bullet}_{nT \times 1} = \underbrace{(I_n \otimes Z)}_{n \times nrp} \cdot \underbrace{a}_{nrp \times 1} + U,$$

where $\underbrace{U}_{nT \times 1}$ has the following distribution:

$$U \sim \mathcal{MN}\left(\mathbf{0},\, \underbrace{V_u}_{n \times n} \right),$$

and

$$
\begin{aligned}
V_u \equiv & \begin{bmatrix}
\Omega_1^{(1,1)} & 0 & \cdots & 0 & \cdots & \cdots & \Omega_1^{(1,n)} & 0 & \cdots & 0 \\
0 & \Omega_2^{(1,1)} & \ddots & \vdots & \cdots & \cdots & 0 & \Omega_2^{(1,n)} & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 & \cdots & \cdots & \vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \Omega_T^{(1,1)} & \cdots & \cdots & 0 & \cdots & 0 & \Omega_T^{(1,n)} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\
\Omega_1^{(n,1)} & 0 & \cdots & 0 & \cdots & \cdots & \Omega_1^{(n,n)} & 0 & \cdots & 0 \\
0 & \Omega_2^{(n,1)} & \ddots & \vdots & \cdots & \cdots & 0 & \Omega_2^{(n,n)} & \ddots & \vdots \\
\vdots & \ddots & \ddots & 0 & \cdots & \cdots & \vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \Omega_T^{(n,1)} & \cdots & \cdots & 0 & \cdots & 0 & \Omega_T^{(nn)}
\end{bmatrix} \\
= & \sum_{t=1}^{T} \left[\Omega_t \otimes (e_t \cdot e_t') \right].
\end{aligned}
$$

To use an informative prior on a we follow the approach by Gelman et al. (2014). The strategy incorporates the prior as observations. Considering a multivariate Normal prior with the following moments:

$$a \sim \mathcal{MN} \left(\; \bar{a}, \; V_a \; \right),$$

it is possible to augment the model with nrp observations that express the prior information:

$$
\begin{aligned}
\begin{bmatrix} Y^{\bullet} \\ \bar{a} \end{bmatrix} &= \begin{bmatrix} I_n \otimes Z \\ I_{nrp} \end{bmatrix} a + \begin{bmatrix} U \\ U_a \end{bmatrix}, \\
Y^{\diamond} &= Z^{\diamond} a + U^{\diamond}, \quad U^{\diamond} \sim \mathcal{MN} \left(\; \mathbf{0}_{nT+nrp}, \; V^{\diamond} \; \right), \\
V^{\diamond} &= \begin{bmatrix} V_u & \mathbf{0}_{nT \times nrp} \\ \mathbf{0}_{nrp \times nT} & V_a \end{bmatrix}.
\end{aligned}
$$

A draw for a then comes from the following posterior:

$$
\begin{aligned}
a &\sim \mathcal{MN} \left(\; \widetilde{a}, \; \left(Z^{\diamond\prime} V^{\diamond -1} Z^{\diamond} \right)^{-1} \; \right), \\
\widetilde{a} &= \left(Z^{\diamond\prime} V^{\diamond -1} Z^{\diamond} \right)^{-1} Z^{\diamond\prime} V^{\diamond -1} Y^{\diamond}.
\end{aligned}
$$

In order to decrease the computational burden of this step throughout the sampling, the strategy proposed by Carriero et al. (2016a) is adopted, as generalized in Carriero et al. (2018): the triangular structure of the error is exploited, and coefficients are drawn equation by equation.

A.3 Step 3: Draw the factor weights elements in B_0

Given the restrictions and the nonlinear role of B_0, a Random walk Metropolis step on the kernel of the posterior of each element of B_0 is implemented, nested into the GS In order to do this, we first write the likelihood of the model. Given the reduced form VAR written as:

$$y_t = \mathcal{X}_t \cdot \gamma + A \cdot Z_t + u_t \ , \qquad u_t \overset{i}{\sim} \mathcal{MN}\left(\mathbf{0}, \ \Omega_t \right),$$

conditioning on all the elements, using the chain rule, we can write the likelihood kernel as:

$$f\left(\left(y_t\right)_{t=1}^{T} \middle| \gamma, A, \left(\Omega_t\right)_{t=1}^{T}, \ B_0 \right) \propto \left(\prod_{t=1}^{T} \left|\Omega_t\right|^{-\frac{1}{2}} \right) \exp\left\{ -\frac{1}{2} \sum_{t=1}^{T} \widehat{y}_t' \cdot \Omega_t^{-1} \cdot \widehat{y}_t \right\},$$

where

$$\widehat{y}_t \equiv y_t - A \cdot Z_t - \mathcal{X}_t \cdot \gamma.$$

Now we consider the $r^* \equiv n - r$ scalar unrestricted elements of B_0, i.e. $\left(b_{0,j}\right)_{j=1}^{r^*}$. Then, $\forall j \in \{1, \ldots, r^*\}$ we can define the set $b_{0,j-} \equiv \left(b_{0,s}\right)_{s \neq j}$.

For a given prior $f\left(b_{0,j}\right)$ on each element $b_{0,j}$, we can write the kernel of the conditional posterior of $b_{0,j}$ as:

$$f_{post}\left(b_{0,j} \middle| \left(y_t, \Omega_t\right)_{t=1}^{T}, \ A, \ b_{0,j-} \right) \propto f\left(\left(y_t\right)_{t=1}^{T} \middle| A, \ B_0, \ \left(\Omega_t\right)_{t=1}^{T} \right) \cdot f\left(b_{0,j}\right).$$

We are now ready to design the Metropolis step, separately for each j. Given the last step B_0^{i-1}, a random walk candidate is computed as:

$$b_{0,j}^{*} = b_{0,j}^{i-1} + c_j \cdot \eta_t,$$

where c_j is a scaling factor calibrated to have an acceptance rate of approximately 30%-35% and $\eta_t \overset{iid}{\sim} \mathcal{N}\left(0, v_j\right)$, with v_j being the variance of prior $f\left(b_{0,j}\right)$. The candidate draw is accepted with probability:

$$\alpha_j = \min\left\{1, \frac{f_{post}\left(b^*_{0,j}\middle|\left(y_t, \Omega_t^{i-1}\right)_{t=1}^T,\ A,\ b_{0,j-}^{i-1}\right)}{f_{post}\left(b^{i-1}_{0,j}\middle|\left(y_t, \Omega_t^{i-1}\right)_{t=1}^T,\ A,\ b_{0,j-}^{i-1}\right)}\right\}.$$

When the candidate is accepted, then $b^i_{0,j-} = b^*_{0,j}$, otherwise $b^i_{0,j-} = b^{i-1}_{0,j-}$. Repeating this procedure $\forall j \in \{1,\ldots,r^*\}$, we build a draw B_0^i from the distribution of interest.

A.4 Step 4: draw the off-diagonal elements in G

To draw the off-diagonal elements, we restate the reduced form in the following way:

$$\begin{aligned} y_t &= \underbrace{\mathcal{X}_t}_{n\times nq} \cdot \underbrace{\gamma}_{nq\times 1} + \underbrace{A}_{n\times rp} \cdot \underbrace{Z_t}_{rp\times 1} + u_t, \\ y_t - \mathcal{X}_t \cdot \gamma - A \cdot Z_t &= G^{-1}\Sigma_t\varepsilon_t, \\ \widehat{y}_t &= G^{-1}\Sigma_t\varepsilon_t, \\ G \cdot \widehat{y}_t &= \Sigma_t\varepsilon_t. \end{aligned}$$

Removing ones from the diagonal of G, and bringing off diagonal elements on the right hand side, produces:

$$G = I_n + G^*.$$

This can be combined in the model to obtain:

$$\begin{aligned} \left(I_n + G^*\right)\ \widehat{y}_t &= \Sigma_t\ \varepsilon_t, \\ \widehat{y}_t &= -G^*\ \widehat{y}_t + \Sigma_t\ \varepsilon_t. \end{aligned}$$

Exploiting the Kronecker product's properties, we get:

$$-\ I_n \underbrace{G^*}_{n\times n} \underbrace{\widehat{y}_t}_{n\times 1} = -\underbrace{\left(I_n \otimes \widehat{y}_t'\right)}_{n\times n^2} \underbrace{vec\left(G^{*\prime}\right)}_{n^2\times 1}$$

where $vec\left(G^{*\prime}\right)$ has zeros in positions $\left[(i-1)\,n + j\right]_{j\in\{1,\ldots,n\}}^{i\in\{1,\ldots,n\}}$. By removing the zeros, we obtain exactly the elements below the main diagonal of G gathered in the m-dimensional vector g. Removing the corresponding columns in $-(I_n \otimes \widehat{y}_t')$ we construct the matrix W_t, which has the following form:

$$\underbrace{W_t}_{n\times m} = -1 \cdot \begin{bmatrix} 0 & \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & & & & & & & & & & 0 \\ \widehat{y}_{1,t} & 0 & \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & & & & & & & & & \vdots \\ 0 & \widehat{y}_{1,t} & \widehat{y}_{2,t} & 0 & \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots & & & & & & & \vdots \\ 0 & 0 & 0 & \widehat{y}_{1,t} & \widehat{y}_{2,t} & \widehat{y}_{3,t} & 0 & \dots\dots\dots\dots\dots\dots & & & & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & \dots\dots\dots\dots\dots\dots\dots\dots & & & & & & 0 & \widehat{y}_{1,t} & \widehat{y}_{2,t} & \widehat{y}_{3,t} & \dots & \widehat{y}_{n-1,t} \end{bmatrix}.$$

We can then rewrite the model as:

$$\begin{aligned} \widehat{y}_t &= -G^* \widehat{y}_t + \Sigma_t \varepsilon_t, \\ \widehat{y}_t &= -(I_n \otimes \widehat{y}_t')\, vec\,(G^{*\prime}) + \Sigma_t \varepsilon_t, \\ \widehat{y}_t &= W_t\ g + \varepsilon_t^*, \qquad \varepsilon_t^* \sim \mathcal{MN}\left(\ \mathbf{0}_{n\times 1},\ \Sigma_t^2\ \right). \end{aligned}$$

Next, we stack the model as:

$$\begin{bmatrix} \widehat{y}_1 \\ \widehat{y}_2 \\ \vdots \\ \widehat{y}_T \end{bmatrix} = \begin{bmatrix} W_1 \\ W_2 \\ \vdots \\ W_T \end{bmatrix} g + \begin{bmatrix} \varepsilon_1^* \\ \varepsilon_2^* \\ \vdots \\ \varepsilon_T^* \end{bmatrix},$$

$$\underbrace{\widehat{y}}_{nT\times 1} = \underbrace{W}_{Tn\times m} \cdot \underbrace{g}_{m\times 1} + \varepsilon^*, \qquad \varepsilon^* \sim \mathcal{MN}\left(\ \mathbf{0}_{nT\times 1},\ \Sigma^2\ \right) m$$

where Σ is the diagonal matrix containing all the stacked stochastic volatilities vectors in the main diagonal:

$$\Sigma = Diag\left(\begin{bmatrix} \sigma_1' & \sigma_2' & \dots & \sigma_T' \end{bmatrix}'\right).$$

We can then use a similar approach as the one implemented for a, following Gelman et al. (2014). Given the prior :

$$g \sim \mathcal{MN}\left(\ \bar{g},\ V_g\ \right),$$

we augment the model with r observations that express the prior information:

$$\begin{bmatrix} \widehat{y} \\ \bar{g} \end{bmatrix} = \begin{bmatrix} W \\ I_m \end{bmatrix} g + \begin{bmatrix} \varepsilon^* \\ \varepsilon_g \end{bmatrix},$$

$$\widehat{Y}^\diamond = W^\diamond g + \varepsilon^\diamond, \quad \varepsilon^\diamond \sim \mathcal{MN}\left(\mathbf{0}_{nT+m},\ V_\varepsilon^\diamond \right),$$

$$V_\varepsilon^\diamond = \begin{bmatrix} \Sigma^2 & \mathbf{0}_{nT\times m} \\ \mathbf{0}_{m\times nT} & V_g \end{bmatrix}.$$

A draw for g is finally obtained through the following posterior:

$$g \sim \mathcal{MN}\left(\widetilde{g},\ \left(W^{\diamond\prime} V_\varepsilon^{\diamond-1} W^\diamond\right)^{-1} \right),$$

$$\widetilde{g} = \left(W^{\diamond\prime} V_\varepsilon^{\diamond-1} W^\diamond\right)^{-1} W^{\diamond\prime} V_\varepsilon^{\diamond-1} \widehat{Y}^\diamond$$

A.5 Step 5-6: Draw the indexes of the mixture in S and then a history of volatilities $(\sigma_t)_{t=1}^T$

An important final step concerns the draw of stochastic volatilities. However, before drawing the (unobservable) stochastic volatilities is necessary to draw the matrix S containing the indexes of Normal components of the mixture, as suggested by Del Negro and Primiceri (2015).

To start building the necessary form, recall the model formulation used previously and transform it as:

$$y_t = \mathcal{X}_t \cdot \gamma + A \cdot Z_t + G^{-1}\Sigma_t \varepsilon_t,$$

$$\underbrace{G\left(y_t - \mathcal{X}_t \cdot \gamma - A \cdot Z_t\right)}_{\widetilde{y}_t} = \Sigma_t \varepsilon_t,$$

$$\widetilde{y}_t = \Sigma_t \varepsilon_t.$$

Having this formulation, we adopt the same procedure as in the MAI-SV Gibbs Sampler illustrated in the Appendix of Carriero et al. (2018), which implements the Omori et al. (2007) procedure to approximate the $\log \chi_1^2$ innovations as mixture of Normal components.

A.6 Step 7: Draw a covariance matrix Q_σ

Conditioning on the new $(\sigma_t^i)_{t=0}^T$, we can draw the covariance matrix Q_σ. Indeed, recall that:

$$\log \sigma_t = \log \sigma_{t-1} + \nu_{\sigma,t}, \qquad \nu_{\sigma,t} \overset{iid}{\sim} \mathcal{MN}\left(\mathbf{0},\ \underbrace{Q_\sigma}_{n\times n}\right).$$

But then, having a complete history of the sigmas, given the random walk law of motion, is equivalent to having a complete histories of innovations $\nu_{\sigma,t}$. Stacking the $\nu_{\sigma,t}$ across time, we get:

$$\underbrace{\nu_\sigma^*}_{n\times T} = \begin{bmatrix} \nu_{\sigma,1} & \nu_{\sigma,2} & \dots & \nu_{\sigma,T} \end{bmatrix},$$

and we can easily compute the innovations sum of squares matrix:

$$\underbrace{S_\sigma}_{n\times n} = \underbrace{\nu_\sigma^*}_{n\times T} \underbrace{\nu_\sigma^{*\prime}}_{T\times n}.$$

If the prior on the matrix Q_σ is a $n \times n$ Inverse Wishart with scale matrix $\bar{Q}_\sigma$ and degrees of freedom $\tau_{\sigma,0}$:

$$Q_\sigma \sim \mathcal{IW}_n\left(\bar{Q}_\sigma,\ \tau_{\sigma,0}\right),$$

then the posterior is conjugate and given by:

$$Q_\sigma \big| \left(\sigma_t^i\right)_{t=0}^T \sim \mathcal{IW}_n\left(S_\sigma + \bar{Q}_\sigma,\ \tau_{\sigma,0} + T\right).$$

B Forecasting Evaluation Tables

The following tables contain the Root Mean Squares Errors, Predictive Log Scores and Continous Rank Probability Scores relative to the forecasting evaluation section.

Table 1a: Root Mean Squared Forecast Errors (RMSE for MAI-AR-SV, RMSE ratios in all others)

USA

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.254	1.017***	1.024*	1.026	0.972***	0.976
$h=2$	0.417	1.029***	1.032	1.045**	0.933**	0.923*
$h=3$	0.525	1.035***	1.061	1.073**	0.933	0.918
$h=4$	0.602	1.046***	1.088	1.105**	0.947	0.922
$h=5$	0.583	1.053***	1.127	1.148**	0.987	0.952
$h=6$	0.553	1.048**	1.183	1.199**	1.046	1.004
$h=7$	0.543	1.040**	1.242	1.255**	1.087	1.038
$h=8$	0.538	1.023**	1.283	1.296**	1.096	1.054

Australia

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.203	0.983***	1.022***	1.020	1.015***	1.023
$h=2$	0.309	0.976***	1.042***	1.046	1.035*	1.069
$h=3$	0.401	0.971***	1.060***	1.067	1.081*	1.137*
$h=4$	0.485	0.974***	1.065***	1.078	1.091	1.162**
$h=5$	0.518	0.969***	1.087***	1.101	1.120*	1.198**
$h=6$	0.547	0.966***	1.105***	1.119*	1.132*	1.209**
$h=7$	0.562	0.963***	1.125***	1.137*	1.134*	1.195**
$h=8$	0.576	0.961***	1.149***	1.159*	1.133*	1.169**

Austria

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.209	1.069***	1.024***	1.001	0.968***	0.948
$h=2$	0.343	1.069***	1.020***	0.995	0.974***	0.919
$h=3$	0.450	1.047***	1.022***	1.014	0.980*	0.920
$h=4$	0.535	1.049***	1.027***	1.029	1.001	0.951
$h=5$	0.575	1.062***	1.056***	1.055	1.033	0.979
$h=6$	0.595	1.078***	1.099***	1.094	1.066	1.009
$h=7$	0.618	1.094***	1.129***	1.125	1.097	1.039
$h=8$	0.637	1.107***	1.148**	1.143	1.117	1.057

Belgium

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.200	0.972***	1.020***	1.054***	1.010	1.040
$h=2$	0.331	0.980***	1.042***	1.091**	0.982	0.973
$h=3$	0.438	0.990***	1.045***	1.078*	0.959	0.927
$h=4$	0.544	1.012***	1.034**	1.058	0.932	0.897
$h=5$	0.594	1.027***	1.033*	1.057*	0.925	0.891
$h=6$	0.609	1.037***	1.039	1.066*	0.942	0.900
$h=7$	0.616	1.041***	1.053	1.078	0.955	0.902
$h=8$	0.611	1.041***	1.077	1.103	0.960	0.898

Canada

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.227	1.000***	1.040	1.036*	0.975***	0.988
$h=2$	0.363	0.997***	1.047	1.040*	0.950***	0.971
$h=3$	0.456	0.993***	1.053	1.043	0.941***	0.964*
$h=4$	0.530	1.006***	1.048	1.041	0.936***	0.963**
$h=5$	0.551	1.002***	1.017	1.011	0.929**	0.978**
$h=6$	0.560	0.989***	0.989	0.983	0.921**	0.992**
$h=7$	0.574	0.986***	0.986	0.980	0.920*	0.999
$h=8$	0.588	0.984***	0.995	0.987	0.925	1.008

Finland

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.124	0.991***	1.041***	1.022	1.098***	1.147***
$h=2$	0.233	1.006***	0.991***	0.963*	1.055**	1.108
$h=3$	0.335	1.019***	0.973***	0.931**	1.013	1.073
$h=4$	0.436	1.026***	0.966***	0.922**	0.985	1.046
$h=5$	0.513	1.042***	0.961***	0.907*	0.990	1.041
$h=6$	0.567	1.049***	0.973**	0.908	1.007	1.044
$h=7$	0.606	1.056***	0.999*	0.925	1.032	1.047
$h=8$	0.634	1.066***	1.026*	0.944	1.049	1.044

France

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.108	1.034***	0.972***	0.974**	1.072***	1.078
$h=2$	0.180	1.080***	0.964***	0.956	1.089	1.091
$h=3$	0.234	1.114***	0.991***	0.947	1.125	1.125
$h=4$	0.292	1.144***	1.003***	0.955	1.113	1.125
$h=5$	0.323	1.172***	1.016**	0.946	1.131	1.143
$h=6$	0.341	1.196***	1.036**	0.937	1.169	1.173
$h=7$	0.360	1.212***	1.050**	0.926	1.182	1.175**
$h=8$	0.382	1.22***	1.059*	0.914	1.173	1.158**

Germany

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.274	0.996***	0.982	0.984	1.026	1.035
$h=2$	0.418	1.003***	0.978	0.991	1.063	1.051
$h=3$	0.505	1.022***	1.018	1.030	1.136	1.104
$h=4$	0.594	1.035***	1.042	1.057	1.152	1.124
$h=5$	0.624	1.040***	1.063	1.078	1.186	1.171
$h=6$	0.655	1.050***	1.084	1.102	1.202	1.188
$h=7$	0.689	1.061***	1.101	1.120	1.210	1.186
$h=8$	0.717	1.072***	1.111	1.132	1.224	1.188

Greece

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.111	1.044***	1.011***	0.996	1.138***	1.069***
$h=2$	0.193	0.985***	1.027***	0.984	1.161***	1.097***
$h=3$	0.256	0.999***	1.057***	0.985	1.169***	1.082***
$h=4$	0.297	0.999***	1.087***	0.988	1.193***	1.083**
$h=5$	0.306	1.004***	1.151***	1.005	1.239***	1.110**
$h=6$	0.313	1.023***	1.239***	1.045	1.277***	1.151**
$h=7$	0.312	1.065***	1.359***	1.116	1.320***	1.190**
$h=8$	0.311	1.124***	1.490***	1.206	1.336***	1.205*

Italy

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.065	1.135***	1.032***	0.960***	1.172***	1.126***
$h=2$	0.128	1.074***	1.024***	0.946**	1.146***	1.155**
$h=3$	0.187	1.057***	1.040***	0.951*	1.116*	1.190
$h=4$	0.250	1.057***	1.023***	0.937	1.080	1.190
$h=5$	0.296	1.059***	1.030***	0.933	1.061	1.191
$h=6$	0.331	1.058***	1.054***	0.939	1.049	1.185
$h=7$	0.363	1.064***	1.071***	0.940	1.027	1.160
$h=8$	0.387	1.066***	1.101***	0.956	0.998	1.137

Statistically significant differences according to the Diebold-Mariano t-statistic are indicated by asterisks, where

*,** and ** correspond respectively to 10%,5% and 1% significance levels

Table 1b: Root Mean Squared Forecast Errors (RMSE for MAI-AR-SV, RMSE ratios in all others)

Japan

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h = 1$	0.142	1.131***	1.092***	1.018	1.143***	1.119***
$h = 2$	0.230	1.180***	1.113***	0.969*	1.218***	1.123*
$h = 3$	0.313	1.230***	1.131***	0.927**	1.274***	1.134
$h = 4$	0.400	1.253***	1.159***	0.908**	1.273***	1.107
$h = 5$	0.464	1.289***	1.178***	0.867***	1.278***	1.085
$h = 6$	0.517	1.312***	1.21***	0.851***	1.262***	1.063
$h = 7$	0.563	1.322***	1.243***	0.848***	1.234***	1.033
$h = 8$	0.602	1.330***	1.264***	0.842***	1.208***	1.016

Luxembourg

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h = 1$	0.194	0.998***	1.013***	1.024	1.020	1.042**
$h = 2$	0.307	1.015***	1.014***	1.030	1.027	1.042
$h = 3$	0.377	1.027***	1.029***	1.042	1.038	1.048
$h = 4$	0.448	1.033***	1.046**	1.059	1.021	1.029
$h = 5$	0.473	1.041***	1.066**	1.083	1.027	1.033
$h = 6$	0.474	1.038***	1.101**	1.118	1.046	1.046
$h = 7$	0.485	1.030***	1.140**	1.155*	1.069	1.062
$h = 8$	0.492	1.024***	1.168***	1.182*	1.096	1.086

Netherlands

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h = 1$	0.156	0.980***	1.030***	1.012	1.157***	1.085**
$h = 2$	0.231	0.991***	1.032***	1.007	1.231***	1.168**
$h = 3$	0.283	1.000***	1.065***	1.040	1.267***	1.234*
$h = 4$	0.359	1.009***	1.033***	1.025	1.200*	1.201
$h = 5$	0.400	1.032***	1.025***	1.006	1.168	1.186
$h = 6$	0.436	1.051***	1.033**	1.005	1.130	1.158
$h = 7$	0.480	1.068***	1.019**	0.996	1.090	1.110
$h = 8$	0.510	1.083***	1.015*	0.987	1.072	1.075

New Zealand

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h = 1$	0.153	0.962***	0.966***	0.959	1.101***	1.082*
$h = 2$	0.259	0.935***	0.928***	0.873	1.140**	1.134
$h = 3$	0.354	0.919***	0.925**	0.813	1.133*	1.141
$h = 4$	0.437	0.922***	0.942**	0.811	1.102**	1.120
$h = 5$	0.491	0.921***	0.948**	0.782	1.064**	1.085**
$h = 6$	0.532	0.921***	0.964**	0.765	1.036*	1.045*
$h = 7$	0.561	0.931***	0.986**	0.760	1.008	0.998
$h = 8$	0.575	0.947***	1.021**	0.768	0.984	0.951

Norway

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h = 1$	0.248	1.004***	1.028***	0.999	0.965	0.961
$h = 2$	0.350	1.090***	1.086***	1.012	0.957	0.940*
$h = 3$	0.443	1.121***	1.121***	1.014	0.936	0.929*
$h = 4$	0.538	1.134***	1.145***	1.020	0.917	0.905*
$h = 5$	0.578	1.162***	1.171***	1.001	0.938	0.903
$h = 6$	0.601	1.165***	1.224***	1.014	0.959	0.901
$h = 7$	0.614	1.157***	1.282***	1.036	0.973	0.898
$h = 8$	0.630	1.140***	1.327***	1.063	0.972	0.889

Portugal

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h = 1$	0.079	1.500***	1.375***	1.013***	1.361***	1.085***
$h = 2$	0.136	1.341***	1.384***	1.014***	1.287***	1.064***
$h = 3$	0.182	1.261***	1.446***	1.023*	1.274***	1.094***
$h = 4$	0.227	1.148***	1.504***	1.035	1.276***	1.112***
$h = 5$	0.257	1.080***	1.580***	1.048	1.267***	1.099***
$h = 6$	0.286	1.032***	1.637***	1.058	1.220***	1.076***
$h = 7$	0.312	0.995***	1.690***	1.079	1.154***	1.050**
$h = 8$	0.335	0.998***	1.729***	1.093	1.094***	1.032**

Spain

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h = 1$	0.116	0.982***	1.012***	0.984**	1.065***	0.989
$h = 2$	0.204	0.983***	0.987***	0.936**	1.028***	0.999
$h = 3$	0.277	0.994***	1.001***	0.922**	1.006	1.011
$h = 4$	0.341	1.017***	1.013***	0.911*	0.979	1.012
$h = 5$	0.385	1.026***	1.027***	0.894*	0.968	1.018
$h = 6$	0.421	1.026***	1.048***	0.882**	0.955	1.002
$h = 7$	0.457	1.028***	1.063***	0.875*	0.934	0.977
$h = 8$	0.488	1.033***	1.081***	0.874*	0.916	0.968

Sweden

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h = 1$	0.235	1.024***	1.038***	1.033	1.047	1.093*
$h = 2$	0.383	1.044***	1.053**	1.052	1.056	1.070
$h = 3$	0.499	1.050***	1.089**	1.077	1.075	1.055
$h = 4$	0.609	1.059***	1.108**	1.093	1.031	0.992
$h = 5$	0.668	1.069***	1.126**	1.105	1.014	0.966
$h = 6$	0.698	1.081***	1.170***	1.138	1.012	0.956
$h = 7$	0.721	1.097***	1.215***	1.172	1.007	0.942
$h = 8$	0.740	1.109***	1.257***	1.206	1.01	0.938

Switzerland

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h = 1$	0.194	1.085***	1.017***	0.988***	1.155***	1.165***
$h = 2$	0.355	1.081***	0.988**	0.956***	1.143*	1.157*
$h = 3$	0.496	1.083***	0.976	0.941***	1.125	1.128
$h = 4$	0.609	1.097***	0.974	0.937**	1.111	1.098
$h = 5$	0.673	1.121***	0.985	0.941*	1.114	1.071
$h = 6$	0.705	1.152***	1.012	0.952*	1.137	1.052
$h = 7$	0.733	1.179***	1.041	0.968	1.166	1.037
$h = 8$	0.762	1.198***	1.067	0.982	1.199	1.030

United Kingdom

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h = 1$	0.114	1.054***	1.008***	0.989***	1.201***	1.112
$h = 2$	0.184	1.062***	1.044***	1.003***	1.245**	1.138
$h = 3$	0.245	1.069***	1.057***	0.980**	1.237	1.154
$h = 4$	0.316	1.073***	1.035***	0.939**	1.173	1.132
$h = 5$	0.359	1.077***	1.043***	0.921*	1.122	1.104
$h = 6$	0.396	1.077***	1.056***	0.911*	1.088	1.080
$h = 7$	0.433	1.074***	1.073**	0.918*	1.063	1.054
$h = 8$	0.459	1.072***	1.107**	0.947*	1.047	1.035

Statistically significant differences according to the Diebold-Mariano t-statistic are indicated by asterisks, where

*,** and ** correspond respectively to 10%,5% and 1% significance levels

Table 2a: Average Log Predictive Scores (scores for MAI-AR-SV, score differences in all others)

USA

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.166	−0.292***	−1.263	−0.032	−1.506	−0.022
$h=2$	−0.343	−0.277***	−1.161	−0.063*	−0.971	−0.016
$h=3$	−0.600	−0.267***	−0.532	−0.094*	−0.410	−0.044
$h=4$	−0.805	−0.241***	−0.229	−0.130**	−0.110	−0.048
$h=5$	−0.842	−0.276***	−0.180*	−0.154**	−0.053	−0.056
$h=6$	−0.867	−0.308***	−0.162**	−0.180***	−0.052	−0.073
$h=7$	−0.892	−0.337***	−0.180**	−0.208***	−0.067	−0.086
$h=8$	−0.927	−0.343***	−0.182**	−0.223***	−0.064	−0.087

Australia

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.206	−0.334***	−0.139**	−0.044	−0.069	0.025
$h=2$	−0.291	−0.197**	−0.052	−0.050	0.012	−0.024
$h=3$	−0.552	−0.163*	−0.053	−0.083	−0.013	−0.080
$h=4$	−0.722	−0.169	−0.075	−0.104	−0.026	−0.112
$h=5$	−0.770	−0.215*	−0.136	−0.139	−0.067	−0.139**
$h=6$	−0.809	−0.238**	−0.183*	−0.164*	−0.091	−0.150*
$h=7$	−0.833	−0.260**	−0.216**	−0.182*	−0.105	−0.134*
$h=8$	−0.871	−0.256**	−0.233**	−0.204*	−0.105	−0.114

Austria

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.126	−0.516***	−0.265***	0.004	−0.146*	0.052
$h=2$	−0.377	−0.386***	−0.172*	−0.007	−0.044	0.102
$h=3$	−0.696	−0.241*	−0.078	0.022	0.047	0.143
$h=4$	−0.863	−0.195	−0.064	−0.040	0.045	0.088
$h=5$	−0.919	−0.213**	−0.102	−0.081	0.002	0.045
$h=6$	−0.957	−0.227**	−0.132	−0.128	−0.023	0.021
$h=7$	−1.016	−0.212*	−0.131	−0.170*	−0.032	0.000
$h=8$	−1.059	−0.206	−0.129	−0.204**	−0.039	−0.014

Belgium

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.239	−0.283***	−0.110**	−0.040	−0.074	−0.058
$h=2$	−0.191	−0.307***	−0.189***	−0.101***	−0.151*	−0.535
$h=3$	−0.453	−0.288**	−0.247**	−0.111*	−0.376*	−0.650
$h=4$	−0.723	−0.226*	−0.191*	−0.106	−0.232	−0.694
$h=5$	−0.822	−0.240**	−0.151*	−0.125*	−0.051	0.050
$h=6$	−0.863	−0.267**	−0.154*	−0.152**	−0.038	0.060
$h=7$	−0.912	−0.270***	−0.140*	−0.172***	0.007	0.077
$h=8$	−0.942	−0.273***	−0.141*	−0.195***	0.046	0.096

Canada

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.075	−0.119**	−0.064	−0.086**	−0.032	−0.003
$h=2$	−0.352	−0.130**	−0.127	−0.109**	−0.147	−0.004
$h=3$	−0.556	−0.160**	−0.127	−0.137*	−0.070	−0.020
$h=4$	−0.705	−0.173**	−0.126	−0.122	−0.066	−0.005
$h=5$	−0.758	−0.193**	−0.096	−0.109	0.030	0.019
$h=6$	−0.789	−0.211**	−0.086	−0.104	0.063	0.036
$h=7$	−0.826	−0.227**	−0.095	−0.124	0.069	0.047
$h=8$	−0.862	−0.232**	−0.105	−0.149*	0.067	0.034

Finland

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.651	−0.550***	−0.271***	−0.022	−0.199***	−0.108**
$h=2$	0.055	−0.478***	−0.204***	−0.049	−0.111	−0.114
$h=3$	−0.301	−0.402***	−0.156*	0.014	−0.045	−0.091
$h=4$	−0.581	−0.307**	−0.093	−0.018	0.020	−0.042
$h=5$	−0.759	−0.259*	−0.056	0.018	0.048	−0.010
$h=6$	−0.850	−0.249*	−0.068	−0.007	0.040	−0.004
$h=7$	−0.924	−0.235*	−0.074	−0.031	0.032	0.014
$h=8$	−0.978	−0.232	−0.082	−0.056	0.020	0.022

France

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.806	−0.429***	−0.153***	0.028	−0.119**	−0.072*
$h=2$	0.340	−0.519***	−0.190***	−0.003	−0.149*	−0.088
$h=3$	0.078	−0.550***	−0.219***	−0.032	−0.172	−0.129
$h=4$	−0.163	−0.488***	−0.182*	−0.040	−0.142	−0.124
$h=5$	−0.263	−0.502***	−0.207**	−0.043	−0.175	−0.136
$h=6$	−0.333	−0.514***	−0.223**	−0.046	−0.206*	−0.163*
$h=7$	−0.410	−0.503***	−0.221**	−0.043	−0.211*	−0.167**
$h=8$	−0.479	−0.498***	−0.216**	−0.039	−0.211*	−0.171**

Germany

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	−0.072	−0.178***	−0.035	−0.014	−0.062	−0.052
$h=2$	−0.504	−0.172**	−0.023	−0.037	−0.073	−0.055
$h=3$	−0.706	−0.188**	−0.054	−0.081	−0.127	−0.099
$h=4$	−0.904	−0.162*	−0.047	−0.103**	−0.092	−0.105
$h=5$	−0.994	−0.165*	−0.039	−0.127**	−0.079	−0.092
$h=6$	−1.068	−0.162*	−0.042	−0.139***	−0.071	−0.091
$h=7$	−1.140	−0.151*	−0.036	−0.153***	−0.064	−0.073
$h=8$	−1.194	−0.141*	−0.031	−0.155***	−0.059	−0.067

Greece

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.839	−0.761***	−0.428***	−0.002	−0.369***	−0.104***
$h=2$	0.314	−0.752***	−0.414***	−0.003	−0.339***	−0.109**
$h=3$	0.035	−0.777***	−0.423***	−0.022	−0.340***	−0.122*
$h=4$	−0.139	−0.768***	−0.424***	−0.039	−0.342***	−0.131*
$h=5$	−0.216	−0.793***	−0.458***	−0.047	−0.372***	−0.145**
$h=6$	−0.269	−0.806***	−0.491***	−0.056	−0.391***	−0.170**
$h=7$	−0.306	−0.821***	−0.527***	−0.067	−0.405***	−0.187***
$h=8$	−0.350	−0.820***	−0.545***	−0.081	−0.399***	−0.198***

Italy

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	1.262	−0.956***	−0.616***	0.101***	−0.431***	−0.146***
$h=2$	0.638	−0.843***	−0.522***	0.063	−0.283***	−0.142***
$h=3$	0.254	−0.752***	−0.451***	0.043	−0.179	−0.139*
$h=4$	−0.026	−0.644***	−0.375**	0.039	−0.117	−0.125
$h=5$	−0.190	−0.596***	−0.352**	0.030	−0.111	−0.127
$h=6$	−0.298	−0.574***	−0.350**	0.029	−0.128	−0.131
$h=7$	−0.394	−0.547***	−0.338**	0.029	−0.128	−0.126
$h=8$	−0.468	−0.529***	−0.338**	0.012	−0.124	−0.126

Statistically significant differences according to the Diebold-Mariano t-statistic are indicated by asterisks, where

*,** and ** correspond respectively to 10%,5% and 1% significance levels

Table 2b: Average Log Predictive Scores (scores for MAI-AR-SV, score differences in all others)

Japan

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.505	−0.666***	−0.356***	0.000	−0.251***	−0.094**
$h=2$	0.033	−0.609***	−0.331***	0.051	−0.243***	−0.091
$h=3$	−0.270	−0.553***	−0.306***	0.093*	−0.241**	−0.082
$h=4$	−0.537	−0.475***	−0.260**	0.128**	−0.201	−0.040
$h=5$	−0.671	−0.474***	−0.274**	0.151**	−0.214	−0.037
$h=6$	−0.788	−0.453***	−0.267**	0.167**	−0.197	−0.016
$h=7$	−0.870	−0.445***	−0.276**	0.172**	−0.190	−0.007
$h=8$	−0.946	−0.426***	−0.270**	0.181**	−0.167	0.008

Luxembourg

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.245	−0.302***	−0.094**	−0.075	−0.079**	−0.027
$h=2$	−0.199	−0.304***	−0.120***	−0.072*	−0.174	−0.038
$h=3$	−0.426	−0.307***	−0.130**	−0.070	−0.19	−0.055
$h=4$	−0.619	−0.282***	−0.127	−0.105*	−0.048	−0.054
$h=5$	−0.687	−0.312***	−0.149*	−0.117**	−0.042	−0.060
$h=6$	−0.718	−0.332***	−0.182**	−0.133**	−0.038	−0.054
$h=7$	−0.766	−0.332***	−0.193***	−0.145**	−0.027	−0.058
$h=8$	−0.803	−0.333***	−0.207***	−0.148**	−0.025	−0.076

Netherlands

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.388	−0.623***	−0.347***	0.020	−0.237***	−0.064
$h=2$	−0.002	−0.625***	−0.342***	−0.006	−0.219***	−0.106*
$h=3$	−0.213	−0.604***	−0.338***	−0.037	−0.211**	−0.130*
$h=4$	−0.441	−0.508***	−0.261**	−0.030	−0.151	−0.108
$h=5$	−0.540	−0.500***	−0.255**	−0.032	−0.141	−0.111
$h=6$	−0.634	−0.479***	−0.232**	−0.022	−0.123	−0.109
$h=7$	−0.729	−0.451***	−0.197*	−0.016	−0.108	−0.102
$h=8$	−0.792	−0.438***	−0.176	−0.006	−0.099	−0.105

New Zealand

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.501	−0.511***	−0.227***	0.009	−0.198***	−0.056*
$h=2$	0.008	−0.512***	−0.226***	−0.008	−0.195***	−0.087*
$h=3$	−0.249	−0.524***	−0.264***	−0.015	−0.214***	−0.112*
$h=4$	−0.442	−0.481***	−0.259**	−0.014	−0.189**	−0.109*
$h=5$	−0.540	−0.465***	−0.275*	0.012	−0.171*	−0.083
$h=6$	−0.591	−0.463***	−0.308*	0.024	−0.186	−0.079
$h=7$	−0.645	−0.450***	−0.319*	0.037	−0.183	−0.062
$h=8$	−0.691	−0.436***	−0.323*	0.041	−0.167	−0.048

Norway

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.040	−0.350***	−0.162***	0.000	−0.064	0.003
$h=2$	−0.361	−0.311***	−0.152**	−0.019	0.001	0.032
$h=3$	−0.624	−0.265***	−0.144	−0.047	0.068	0.058
$h=4$	−0.836	−0.210**	−0.129	−0.065	0.136	0.121
$h=5$	−0.940	−0.184	−0.128	−0.052	0.152	0.165
$h=6$	−0.984	−0.188	−0.167	−0.070	0.144	0.198*
$h=7$	−0.995	−0.204	−0.218	−0.095	0.115	0.201*
$h=8$	−1.022	−0.204	−0.249*	−0.120	0.110	0.201*

Portugal

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.998	−1.281***	−0.973***	0.094***	−0.848***	−0.194***
$h=2$	0.525	−1.049***	−0.818***	0.045	−0.673***	−0.193***
$h=3$	0.258	−0.949***	−0.765***	−0.010	−0.581***	−0.210***
$h=4$	0.051	−0.870***	−0.737***	−0.044	−0.497***	−0.208**
$h=5$	−0.067	−0.803***	−0.738***	−0.064	−0.464***	−0.212*
$h=6$	−0.155	−0.753***	−0.744***	−0.092	−0.429***	−0.217*
$h=7$	−0.232	−0.698***	−0.750***	−0.106	−0.390**	−0.213
$h=8$	−0.304	−0.650***	−0.738***	−0.107	−0.345**	−0.202

Spain

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.734	−0.670***	−0.393***	0.029	−0.236***	−0.040
$h=2$	0.189	−0.520***	−0.286**	0.071	−0.136	−0.050
$h=3$	−0.196	−0.376	−0.173	0.159	−0.005	0.022
$h=4$	−0.429	−0.310	−0.127	0.152	0.052	0.019
$h=5$	−0.485	−0.361**	−0.201	0.089	−0.001	−0.036
$h=6$	−0.577	−0.340*	−0.205	0.094	0.002	−0.033
$h=7$	−0.653	−0.328**	−0.212	0.065	0.011	−0.020
$h=8$	−0.729	−0.308**	−0.204	0.058	0.027	−0.005

Sweden

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.108	−0.320***	−0.146**	−0.046	−0.102	−0.036
$h=2$	−0.425	−0.184*	−0.092	−0.140	−0.086	−0.012
$h=3$	−0.784	−0.052	−0.025	−0.567	−0.047	0.063
$h=4$	−0.981	−0.044	−0.052	−0.420	−0.022	0.073
$h=5$	−1.081	−0.040	−0.060	−0.095	0.020	0.105
$h=6$	−1.098	−0.083	−0.134	−0.157	0.010	0.089
$h=7$	−1.103	−0.125	−0.208	−0.209	−0.006	0.071
$h=8$	−1.118	−0.145	−0.256*	−0.272**	−0.014	0.047

Switzerland

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.199	−0.374***	−0.103***	0.020	−0.135***	−0.111***
$h=2$	−0.362	−0.297***	−0.072	0.021	−0.148**	−0.134
$h=3$	−0.715	−0.215**	−0.021	0.000	−0.118	−0.104
$h=4$	−0.935	−0.179**	0.008	−0.001	−0.094	−0.078
$h=5$	−1.048	−0.183**	0.007	−0.010	−0.076	−0.036
$h=6$	−1.102	−0.201**	−0.019	−0.016	−0.088	−0.009
$h=7$	−1.153	−0.213***	−0.029	−0.013	−0.102	0.009
$h=8$	−1.197	−0.215**	−0.050	−0.027	−0.126	0.007

United Kingdom

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.738	−0.681***	−0.351***	0.106*	−0.284***	−0.082***
$h=2$	0.227	−0.663***	−0.356***	0.108	−0.245***	−0.074*
$h=3$	−0.061	−0.645***	−0.364***	0.141	−0.229***	−0.091*
$h=4$	−0.287	−0.579***	−0.330***	0.176	−0.202**	−0.096
$h=5$	−0.414	−0.553***	−0.319**	0.191	−0.187**	−0.101
$h=6$	−0.500	−0.536***	−0.321**	0.182	−0.188**	−0.110
$h=7$	−0.580	−0.520***	−0.316**	0.163	−0.189**	−0.108
$h=8$	−0.636	−0.511***	−0.324***	0.107	−0.204*	−0.121

Statistically significant differences according to the Diebold-Mariano t-statistic are indicated by asterisks, where

*,** and ** correspond respectively to 10%,5% and 1% significance levels

Table 3a: Average Continuous Rank Probability Scores (CRPS for MAI-AR-SV, CRPS ratios in all others)

USA

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.123	1.098***	1.028	1.02	0.993	1.000
$h=2$	0.206	1.115***	1.060*	1.042**	0.987	0.966
$h=3$	0.264	1.126***	1.097**	1.071**	0.998	0.970
$h=4$	0.315	1.126***	1.109**	1.106**	0.993	0.966
$h=5$	0.318	1.155***	1.125**	1.141***	1.006	0.983
$h=6$	0.316	1.172***	1.153***	1.185***	1.028	1.019
$h=7$	0.319	1.194***	1.195***	1.227***	1.056	1.055
$h=8$	0.327	1.197**	1.218***	1.249***	1.052	1.063

Australia

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.107	1.191***	1.081**	1.045	1.033	1.004
$h=2$	0.176	1.086**	1.041	1.040	1.008	1.039
$h=3$	0.231	1.051	1.055	1.071	1.043	1.107
$h=4$	0.277	1.056	1.063	1.081	1.063	1.141*
$h=5$	0.293	1.075	1.103	1.117	1.095	1.175**
$h=6$	0.306	1.090	1.137	1.143	1.107	1.184**
$h=7$	0.312	1.107	1.172*	1.165	1.109	1.165*
$h=8$	0.320	1.112	1.196*	1.186	1.107	1.141

Austria

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.114	1.404***	1.172***	1.006	1.065	0.951
$h=2$	0.191	1.268***	1.102*	1.002	1.001	0.918
$h=3$	0.248	1.192***	1.087	1.031	0.998	0.929
$h=4$	0.301	1.138***	1.066	1.053	1.001	0.963
$h=5$	0.327	1.142***	1.092	1.081	1.027	0.987
$h=6$	0.343	1.148***	1.124	1.121*	1.052	1.008
$h=7$	0.361	1.149***	1.142	1.158**	1.074	1.030
$h=8$	0.375	1.153**	1.152	1.179**	1.088	1.043

Belgium

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.107	1.128***	1.053**	1.047	1.050	1.060
$h=2$	0.169	1.139***	1.099***	1.096**	1.055	1.029
$h=3$	0.215	1.143***	1.117**	1.103**	1.042	0.989
$h=4$	0.274	1.122***	1.087*	1.089*	0.999	0.944
$h=5$	0.309	1.127***	1.083	1.099*	0.975	0.921
$h=6$	0.325	1.144***	1.095	1.121**	0.973	0.913
$h=7$	0.339	1.147***	1.101	1.135**	0.972	0.910
$h=8$	0.346	1.150***	1.117*	1.158***	0.959	0.900

Canada

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.126	1.044**	1.039	1.049*	0.983	0.996
$h=2$	0.196	1.048*	1.060	1.060	0.969	0.984
$h=3$	0.246	1.053*	1.073	1.073	0.976	0.993
$h=4$	0.285	1.065**	1.060	1.063	0.971	0.990
$h=5$	0.297	1.081**	1.039	1.044	0.968	1.001
$h=6$	0.305	1.087**	1.029	1.034	0.945	0.994
$h=7$	0.314	1.093**	1.034	1.048	0.931	0.983
$h=8$	0.324	1.102*	1.046	1.063	0.922	0.981

Finland

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.069	1.368***	1.155***	1.014	1.141***	1.133**
$h=2$	0.128	1.295***	1.091*	0.970	1.092	1.129
$h=3$	0.184	1.232***	1.052	0.948	1.046	1.107
$h=4$	0.241	1.172***	1.027	0.945	1.003	1.068
$h=5$	0.286	1.149**	1.011	0.934	0.987	1.044
$h=6$	0.319	1.138**	1.017	0.942	0.981	1.026
$h=7$	0.345	1.129*	1.035	0.966	0.984	1.009
$h=8$	0.365	1.13*	1.053	0.986	0.989	0.997

France

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.059	1.271***	1.053*	0.966	1.108**	1.090*
$h=2$	0.095	1.382***	1.083*	0.984	1.146	1.120
$h=3$	0.122	1.436***	1.124*	0.998	1.199	1.180
$h=4$	0.156	1.392***	1.105	1.007	1.172	1.174
$h=5$	0.174	1.407***	1.121	1.002	1.191	1.185
$h=6$	0.188	1.421***	1.137	0.995	1.214	1.201*
$h=7$	0.202	1.420***	1.140	0.989	1.215	1.193**
$h=8$	0.217	1.416***	1.144	0.980	1.200	1.175**

Germany

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.147	1.072***	0.989	0.992	1.028	1.053
$h=2$	0.225	1.079**	0.981	0.995	1.050	1.055
$h=3$	0.274	1.093**	1.027	1.042	1.102	1.097
$h=4$	0.329	1.089**	1.046	1.070	1.100	1.101
$h=5$	0.354	1.100**	1.059	1.096*	1.107	1.116
$h=6$	0.379	1.099**	1.065	1.114*	1.122	1.124
$h=7$	0.405	1.1**	1.073	1.128**	1.132	1.12
$h=8$	0.427	1.098**	1.077	1.134**	1.142	1.115

Greece

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.059	1.597***	1.252***	1.003	1.270***	1.085**
$h=2$	0.101	1.563***	1.250***	1.000	1.267***	1.107*
$h=3$	0.133	1.616***	1.286***	1.016	1.286***	1.120*
$h=4$	0.158	1.615***	1.300**	1.025	1.305***	1.124*
$h=5$	0.167	1.667***	1.366**	1.043	1.351***	1.150*
$h=6$	0.174	1.711***	1.440***	1.076	1.385***	1.184*
$h=7$	0.178	1.761***	1.533***	1.117	1.415***	1.209*
$h=8$	0.182	1.800***	1.609***	1.164	1.408***	1.215*

Italy

	MAI-AR-SV	*MAI-AR*	*AR*	*AR-SV*	*BVAR*	*BVAR-SV*
$h=1$	0.036	2.026***	1.509***	0.948	1.370***	1.158***
$h=2$	0.068	1.797***	1.388***	0.959	1.238*	1.174**
$h=3$	0.101	1.651***	1.321***	0.968	1.150	1.193
$h=4$	0.136	1.502***	1.243**	0.957	1.089	1.182
$h=5$	0.163	1.442***	1.221	0.954	1.074	1.178
$h=6$	0.183	1.408***	1.224	0.960	1.078	1.173
$h=7$	0.201	1.384***	1.225	0.964	1.07	1.155
$h=8$	0.216	1.376***	1.239	0.983	1.054	1.14

Statistically significant differences according to the Diebold-Mariano t-statistic are indicated by asterisks, where

*,** and ** correspond respectively to 10%,5% and 1% significance levels

Table 3b: Average Continuous Rank Probability Scores (CRPS for MAI-AR-SV, CRPS ratios in all others)

Japan

	MAI-AR-SV	MAI-AR	AR	AR-SV	BVAR	BVAR-SV
$h=1$	0.079	1.558***	1.253***	1.021	1.209***	1.117***
$h=2$	0.129	1.502***	1.236***	0.962	1.234**	1.117
$h=3$	0.177	1.456***	1.219***	0.911*	1.252*	1.115
$h=4$	0.226	1.423***	1.227***	0.889*	1.245	1.083
$h=5$	0.264	1.418***	1.235***	0.851**	1.247	1.053
$h=6$	0.299	1.406***	1.244***	0.835**	1.223	1.025
$h=7$	0.328	1.397***	1.266***	0.831**	1.199	1.000
$h=8$	0.353	1.384***	1.276***	0.826**	1.169	0.985

Luxembourg

	MAI-AR-SV	MAI-AR	AR	AR-SV	BVAR	BVAR-SV
$h=1$	0.103	1.171***	1.054**	1.031	1.036	1.030
$h=2$	0.164	1.168***	1.051	1.043	1.047	1.047
$h=3$	0.201	1.191***	1.076	1.061	1.060	1.059
$h=4$	0.244	1.177***	1.089	1.087	1.034	1.037
$h=5$	0.263	1.190***	1.107	1.106*	1.035	1.036
$h=6$	0.269	1.206***	1.150**	1.141**	1.038	1.034
$h=7$	0.280	1.201***	1.179**	1.171**	1.053	1.051
$h=8$	0.289	1.197***	1.196**	1.185**	1.066	1.066

Netherlands

	MAI-AR-SV	MAI-AR	AR	AR-SV	BVAR	BVAR-SV
$h=1$	0.087	1.483***	1.233***	1.006	1.230***	1.096*
$h=2$	0.130	1.484***	1.224***	1.017	1.250**	1.151*
$h=3$	0.160	1.465***	1.239***	1.053	1.271*	1.212*
$h=4$	0.199	1.379***	1.184**	1.054	1.208	1.189
$h=5$	0.224	1.361***	1.163	1.038	1.172	1.169
$h=6$	0.246	1.347***	1.152	1.029	1.133	1.147
$h=7$	0.275	1.311***	1.112	1.018	1.09	1.112
$h=8$	0.295	1.303***	1.089	1.005	1.071	1.090

New Zealand

	MAI-AR-SV	MAI-AR	AR	AR-SV	BVAR	BVAR-SV
$h=1$	0.082	1.313***	1.085**	0.969	1.152***	1.067
$h=2$	0.139	1.276***	1.061	0.933	1.170**	1.114
$h=3$	0.186	1.265***	1.080	0.904	1.179*	1.137
$h=4$	0.229	1.213***	1.081	0.899	1.135	1.120
$h=5$	0.256	1.193**	1.093	0.872	1.105	1.089
$h=6$	0.274	1.181*	1.118	0.849	1.089	1.061
$h=7$	0.289	1.180*	1.144	0.841	1.068	1.020
$h=8$	0.298	1.187*	1.174	0.844	1.052	0.983

Norway

	MAI-AR-SV	MAI-AR	AR	AR-SV	BVAR	BVAR-SV
$h=1$	0.132	1.189***	1.084***	0.996	1.007	0.979
$h=2$	0.194	1.206***	1.115**	1.010	0.978	0.959
$h=3$	0.250	1.194***	1.137**	1.018	0.940	0.935
$h=4$	0.303	1.183***	1.151*	1.036	0.908	0.903
$h=5$	0.333	1.184***	1.171*	1.024	0.913	0.881
$h=6$	0.349	1.186***	1.225**	1.045	0.919	0.863
$h=7$	0.356	1.184***	1.291**	1.071	0.926	0.848
$h=8$	0.368	1.168**	1.331***	1.098	0.919	0.836

Portugal

	MAI-AR-SV	MAI-AR	AR	AR-SV	BVAR	BVAR-SV
$h=1$	0.046	2.787***	2.130***	0.971	1.922***	1.158***
$h=2$	0.076	2.214***	1.864***	1.004	1.661***	1.150**
$h=3$	0.101	2.006***	1.821***	1.037	1.551***	1.170*
$h=4$	0.124	1.857***	1.821***	1.069	1.482***	1.179
$h=5$	0.140	1.739***	1.870***	1.095	1.447***	1.181
$h=6$	0.155	1.639***	1.902***	1.118	1.394***	1.175
$h=7$	0.170	1.547***	1.923***	1.134	1.323**	1.158
$h=8$	0.184	1.475***	1.923***	1.137	1.247	1.135

Spain

	MAI-AR-SV	MAI-AR	AR	AR-SV	BVAR	BVAR-SV
$h=1$	0.063	1.514***	1.248***	0.981	1.171***	1.027
$h=2$	0.109	1.358***	1.163***	0.941	1.102	1.036
$h=3$	0.148	1.298***	1.147**	0.932	1.069	1.041
$h=4$	0.186	1.251***	1.127	0.923	1.018	1.033
$h=5$	0.212	1.234***	1.132	0.909	1.001	1.033
$h=6$	0.236	1.208***	1.138	0.901	0.980	1.014
$h=7$	0.260	1.180**	1.137	0.900	0.951	0.985
$h=8$	0.280	1.17**	1.142	0.905	0.931	0.974

Sweden

	MAI-AR-SV	MAI-AR	AR	AR-SV	BVAR	BVAR-SV
$h=1$	0.124	1.182***	1.067	1.016	1.064	1.070
$h=2$	0.209	1.113***	1.065	1.044	1.069	1.041
$h=3$	0.275	1.093*	1.099	1.071	1.095	1.032
$h=4$	0.342	1.083	1.112	1.082	1.047	0.977
$h=5$	0.381	1.082	1.125	1.090	1.024	0.947
$h=6$	0.399	1.097	1.179*	1.131	1.023	0.941
$h=7$	0.410	1.120	1.246**	1.186*	1.026	0.936
$h=8$	0.420	1.137*	1.305***	1.240***	1.033	0.944

Switzerland

	MAI-AR-SV	MAI-AR	AR	AR-SV	BVAR	BVAR-SV
$h=1$	0.108	1.256***	1.057**	0.985	1.151***	1.145***
$h=2$	0.194	1.192***	1.017	0.960	1.149*	1.151
$h=3$	0.274	1.151***	0.996	0.954	1.133	1.128
$h=4$	0.343	1.138***	0.984	0.953	1.103	1.088
$h=5$	0.384	1.153***	0.994	0.957	1.096	1.050
$h=6$	0.406	1.175***	1.019	0.969	1.115	1.021
$h=7$	0.426	1.199***	1.047	0.987	1.144	1.001
$h=8$	0.446	1.215***	1.074	1.003	1.179	0.995

United Kingdom

	MAI-AR-SV	MAI-AR	AR	AR-SV	BVAR	BVAR-SV
$h=1$	0.063	1.533***	1.195***	0.936	1.238***	1.086*
$h=2$	0.104	1.531***	1.224**	0.952	1.243**	1.100
$h=3$	0.138	1.517***	1.245**	0.933	1.241**	1.130*
$h=4$	0.176	1.436***	1.205	0.903	1.195*	1.123
$h=5$	0.202	1.406***	1.195	0.883	1.158	1.105
$h=6$	0.222	1.382***	1.197	0.878	1.135	1.087
$h=7$	0.241	1.368***	1.214	0.898	1.126	1.074
$h=8$	0.255	1.361***	1.241	0.940	1.126	1.068

Statistically significant differences according to the Diebold-Mariano t-statistic are indicated by asterisks, where

*,** and ** correspond respectively to 10%,5% and 1% significance levels

C Additional Figures

C.1 Core Inflation, Data and Decompositions

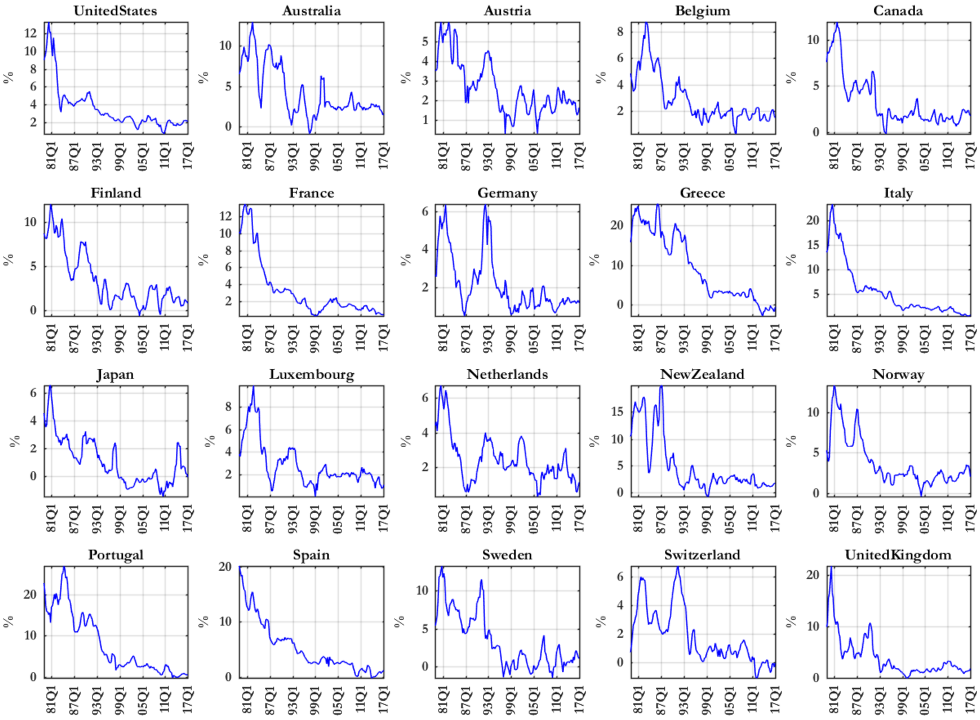


Figure 20: Non-Food & non-Energy inflation rates (year on year growth rates in quarterly CPIs)

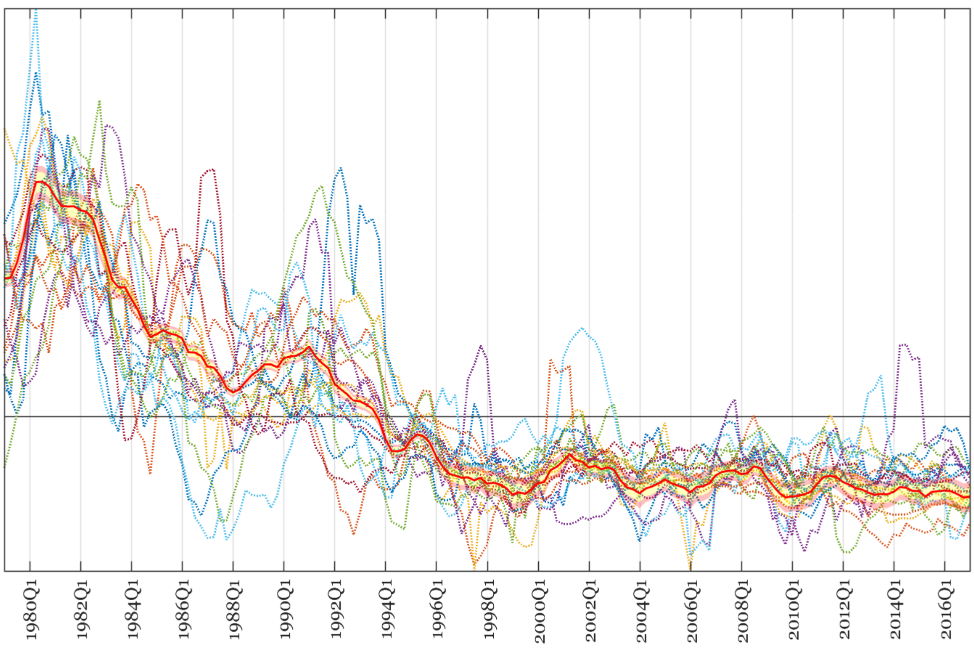


Figure 21: MAI-AR-SV estimated common factor (with posterior bands) Vs Data. Core inflation

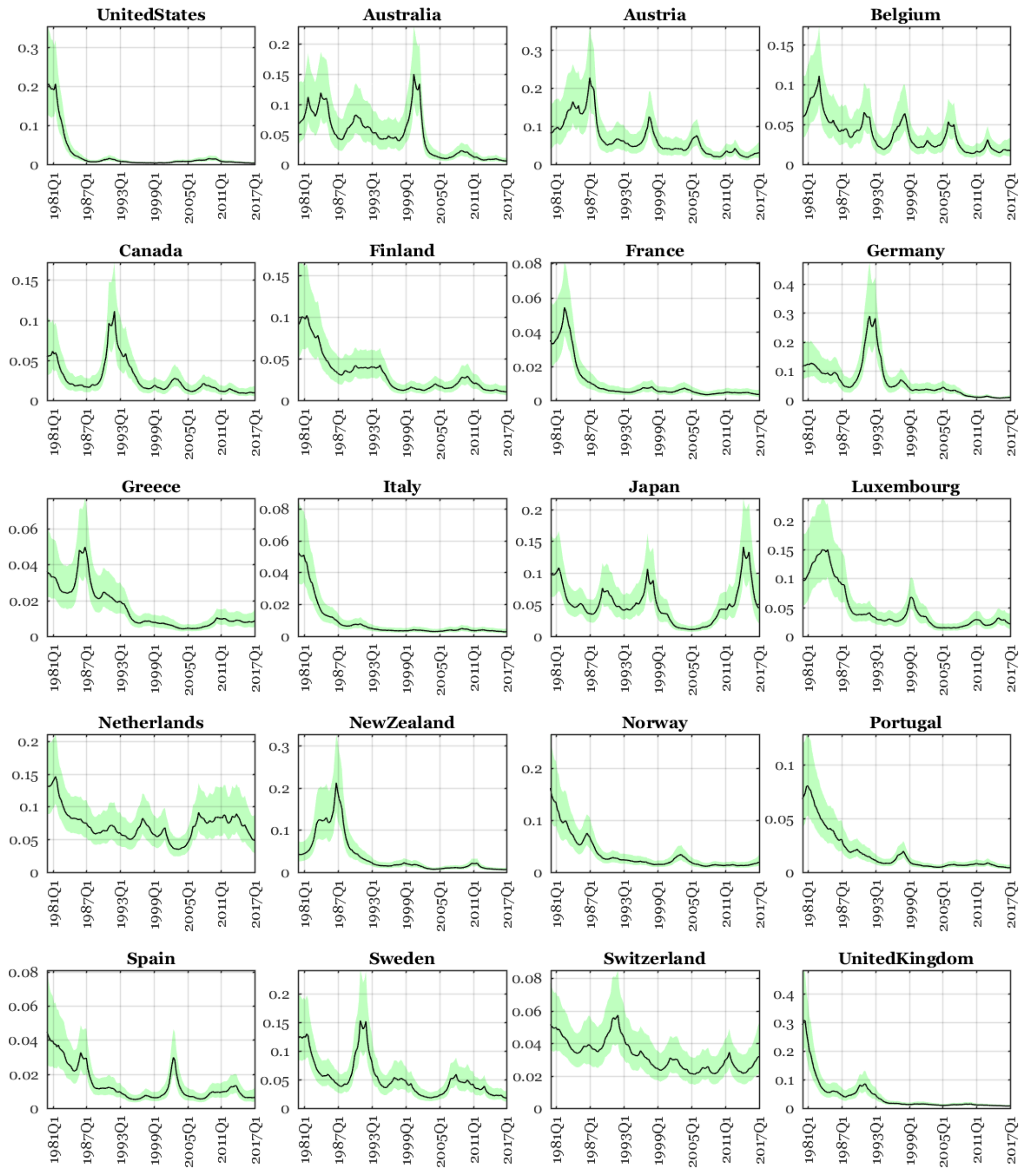


Figure 22: MAI-AR-SV, Residuals' Volatility, posterior bands. Core inflation

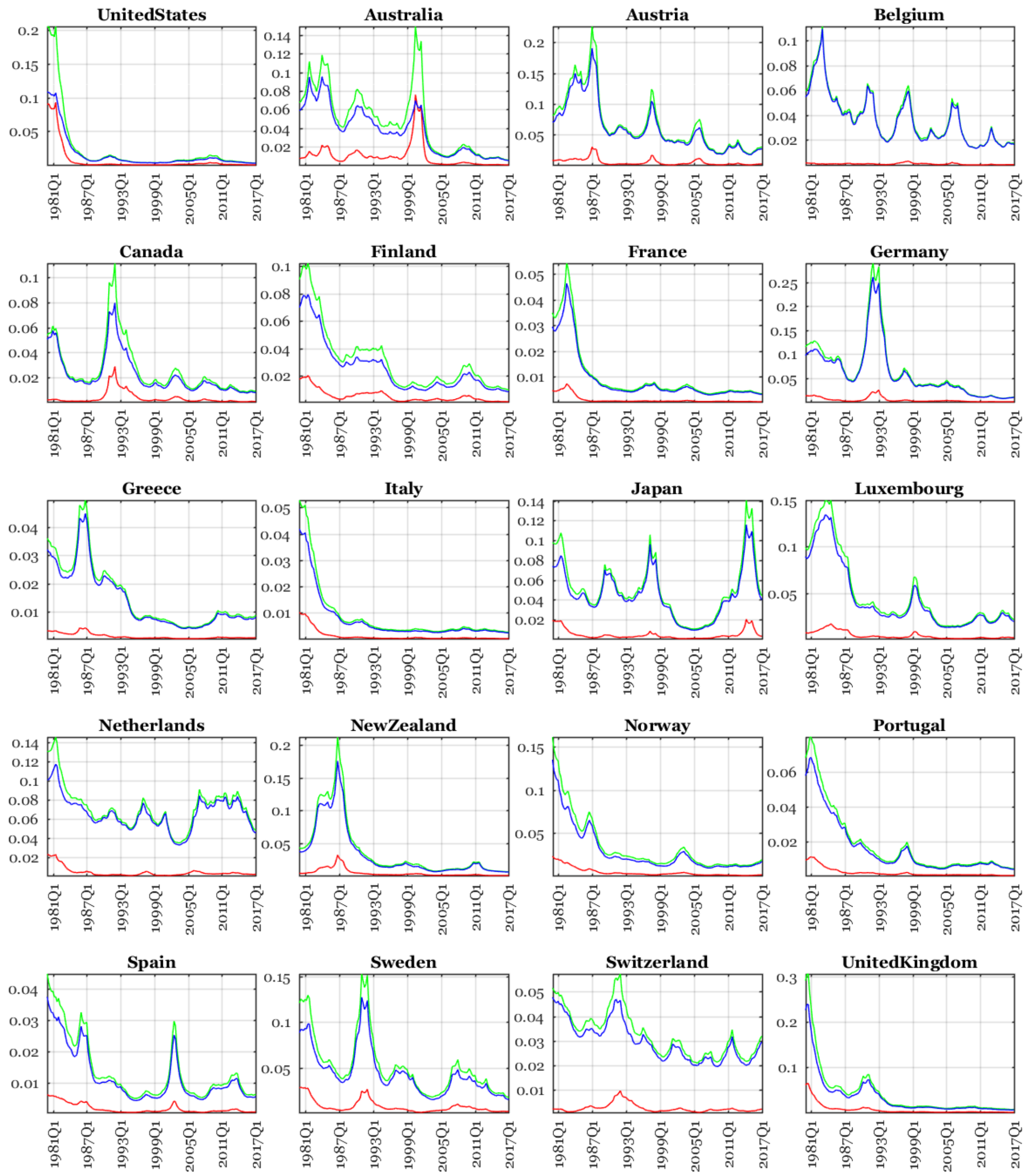


Figure 23: MAI-AR-SV, Residuals' Volatility, TV decomposition, Common (red), Idio (green), total (blue). Core inflation

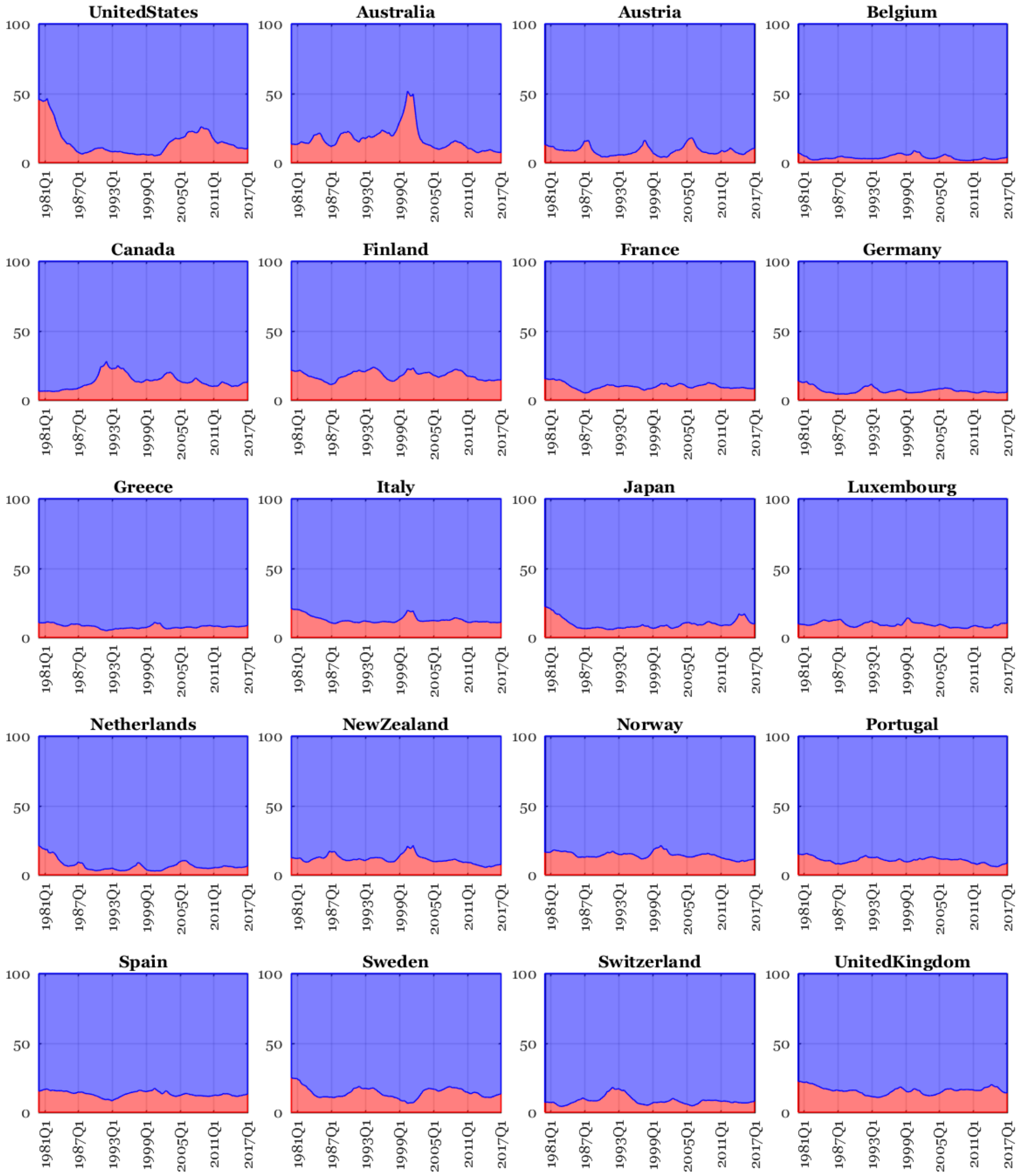


Figure 24: MAI-AR-SV, Residuals' Volatility, TV decomposition shares (%), Common (red), Idio (blue). Core inflation

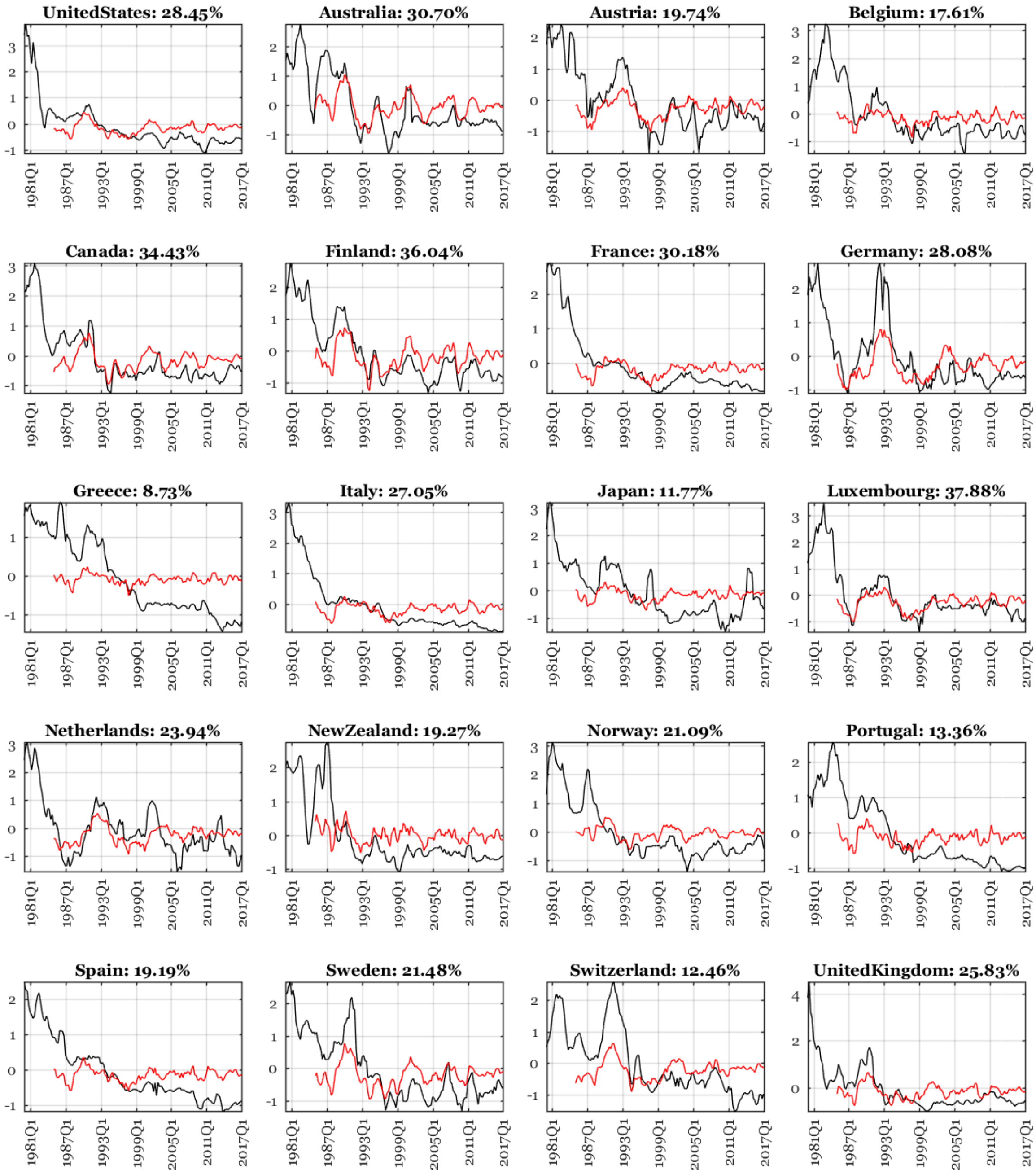


Figure 25: MAI-AR-SV, Actual series and Common component (red). Core inflation

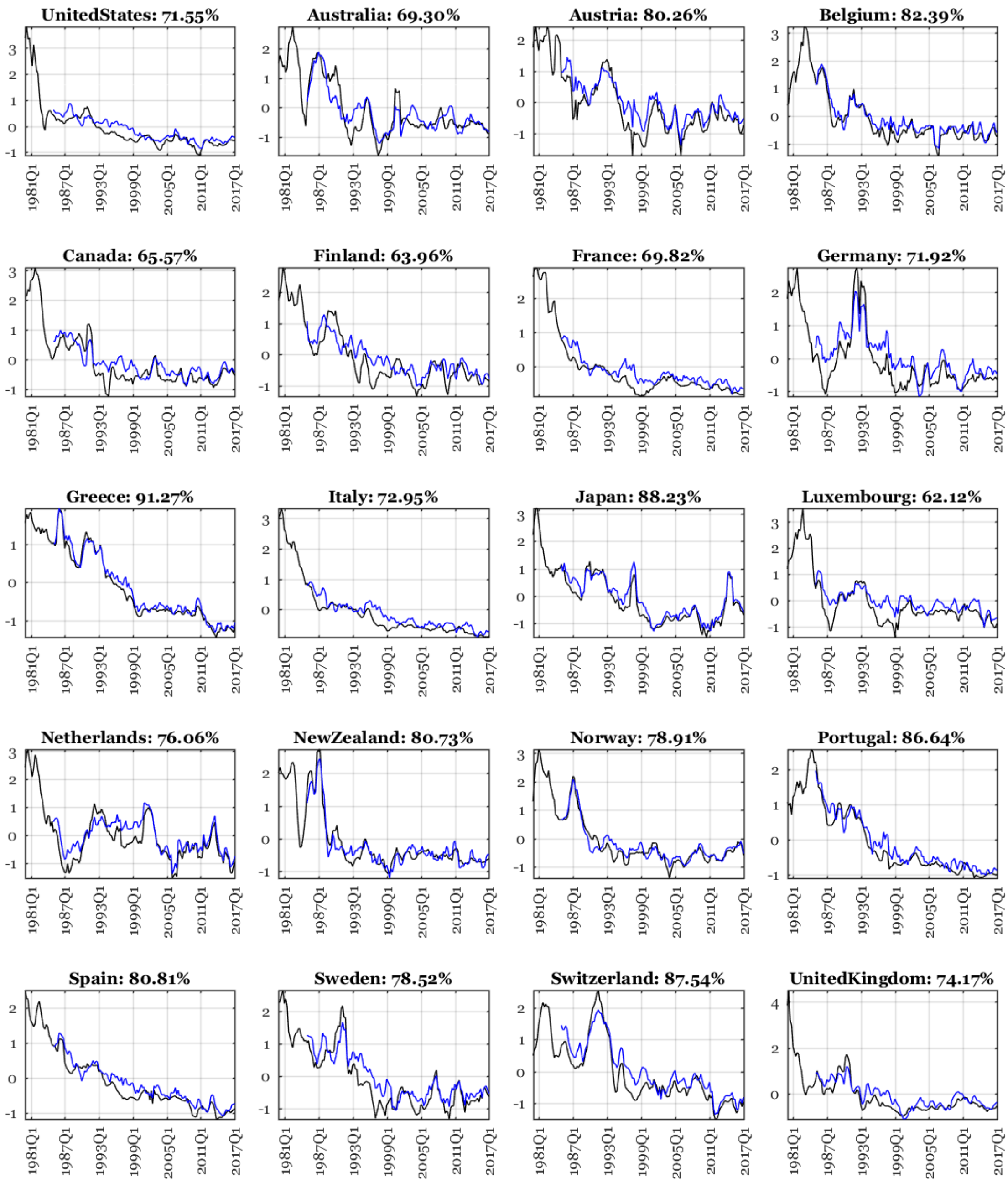


Figure 26: MAI-AR-SV, Actual series and Idiosyncratic component (blue). Core inflation