ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

ECONOMICS

Ciclo 35

Settore Concorsuale: 13/A5 - ECONOMETRIA

Settore Scientifico Disciplinare: SECS-P/05 - ECONOMETRIA

ESSAYS IN BAYESIAN MACROECONOMETRICS

Presentata da: Andrea Renzetti

Coordinatore Dottorato

Andrea Mattozzi

Supervisore

Andrea Carriero

Esame finale anno 2024

Contents

Introduction **1**

Chapter 1: Theory coherent shrinkage of time varying parameters in VARs **3**

Chapter 2: Modelling and Forecasting Macroeconomic Risk with Time Varying Skewness Stochastic Volatility Models **63**

Chapter 3: Labour-at-risk **104**

Introduction

This doctoral dissertation introduces three novel chapters, each contributing to the enhancement of Vector Autoregressive (VARs) models for forecasting and policy analysis in macroeconomics.

The first chapter introduces a "theory-coherent" shrinkage prior for Time-Varying Parameters Vector Autoregressive (TVP-VAR) models to address the limitations of the traditional TVP-VAR models, which often suffer from overfitting, leading to imprecise estimates of the time-varying parameters and inaccurate forecasts. This approach leverages a theoretical framework derived from underlying economic theory to form a prior for time-varying parameters. The resulting TVP-VAR, which encodes the restrictions implied by the economic theory as a prior - hence labeled Theory Coherent TVP-VAR (TC-TVP-VAR) - significantly improves both inference precision and forecast accuracy over standard TVP-VAR models. An application demonstrates enhanced forecast accuracy of GDP growth and the inflation rate from a TC-TVP-VAR that exploits the standard 3-equations New Keynesian block to form a prior for the time-varying parameters. Then, a second application demonstrates the usefulness of the proposed shrinkage prior for impulse response analysis. In particular, the application shows that the proposed prior allows estimating more precisely the effects of macroeconomic shocks inside and outside the Zero Lower Bound (ZLB) period, allowing to address the inferential challenges faced by the standard TVP-VAR model. On US data, the application finds convincing evidence of a different propagation of risk premium shocks inside and outside the ZLB.

The second chapter focuses on the crucial task of monitoring macroeconomic risks, proposing a VAR model with stochastic volatility and time-varying skewness for modeling and forecasting macroeconomic risk. The chapter develops efficient posterior simulation samplers for Bayesian estimation of stochastic volatility VAR models with time-varying skewness featuring both *Skew-Normal* and *Skew-t* shocks. By applying these models to predict downside risks to GDP growth in the US, the chapter demonstrates their competitiveness compared to semi-parametric approaches such as quantile regression. Furthermore, the estimation of a medium-scale model on US data highlights the relevance of time-varying skewness in macroeconomic and financial shocks.

Finally, the third chapter exploits a Bayesian VAR model with stochastic volatility and time-varying skewness to estimate the degree of labor at risk in the euro area and the United States. Examining the asymmetry of shocks to changes in the unemployment rate as a function of real activity and financial risk factors, the chapter uncovers time-varying volatility and skewness in the conditional distribution of the changes in the unemployment rate. The multivariate nature of the model also allows for the measurement of stagflation risk, revealing an increasing risk for the euro area in 2022. This chapter underscores the importance of labor at risk in understanding the inflation-unemployment trade-off.

Collectively, these three chapters contribute novel insights and methodologies to the field of macroeconomic analysis, providing advanced tools for policymakers and researchers to navigate the complexities of dynamic economic systems.

Theory coherent shrinkage of Time-Varying Parameters in VARs*

Andrea Renzetti‡

Abstract

This paper introduces a novel "theory coherent" shrinkage prior for time varying parameters VARs. The proposed prior can be used to sharpen inference about the time varying parameters by leveraging on prior information from an underlying economic theory about the macroeconomic variables in the model. The paper reveals that exploiting prior information from conventional economic theory to form a prior for the time varying parameters significantly improves inference precision and forecast accuracy over the standard TVP-VAR. More specifically, using the classical 3-equation New Keynesian block to form a prior for the TVP-VAR substantially enhances forecast accuracy of output growth and of the inflation rate in a standard model of monetary policy. Additionally, prior information from economic theory can be used to address the inferential challenges faced by the standard TVP-VAR during the zero lower bound period.

J.E.L Classification Code: C32, C53

Keywords: Time Varying Parameters VARs, Bayesian Econometrics, DSGE-VARs

---

*I would like to thank Frank Schorfheide and Andrea Carriero for invaluable guidance and support. I also thank participants at the NBER-NSF SBIES Conference at the Federal Reserve of Philadelphia and at the Workshop in Empirical Macroeconomics at King's College London. I thank Massimiliano Marcellino, Luca Fanelli, Florian Huber, Agostino Consolo, Michele Piffer, Daniel J. Lewis, Tommaso Tornese for helpful comments.

†Department of Economics University of Bologna, Piazza Scaravilli 2, 40126 Bologna, Italy

‡Department of Economics, Bocconi University, Via Roentgen 1, 20136, Milano, Italy

1 Introduction

Over the past four decades vector autoregressive models have become the leading tool for description, forecasting, structural inference and policy analysis of macroeconomic data (Sims, 1980; Stock and Watson, 2001). A natural progression in the literature was to allow for time-varying parameters to capture changes in the complex dynamic interrelationship among the variables in the system (Cogley and Sargent, 2002; Primiceri, 2005; Cogley and Sargent, 2005). On one side, this class of models known as Time Varying Parameters VARs (TVP-VARs) can be flexible enough to fit many different forms of structural instabilities and evolving nonlinear relationships among the macroeconomic variables. On the other side, due to the growing number of parameters, they can easily become too flexible with adverse consequences on the precision of inference and on the reliability of the forecasts.

In this paper I propose to use economic theory to sharpen inference in TVP-VARs. The approach consists in exploiting prior information coming from a more tightly parameterized model derived by an underlying economic theory about the macroeconomic variables in the system. This prior information provides a set of economically grounded (*fuzzy*) restrictions which are incorporated into a shrinkage prior for the TVP-VAR. The resulting model, that I label Theory Coherent TVP-VAR (TC-TVP-VAR), is a flexible statistical model for the data that leverages on economic theory to enhance inference about the time varying parameters. In the TC-TVP-VAR, two crucial hyper-parameters govern the behavior of the time varying coefficients: the first one determines their intrinsic time variation, while the other one determines their degree of theory coherence stemming from the amount of shrinkage towards the cross-equation restrictions implied by the economic theory. Importantly, the TC-TVP-VAR can incorporate both constant and time varying restriction functions for the time varying parameters, thereby extending the DSGE-VAR toolkit developed in Del Negro and Schorfheide (2004). Both the optimal amount of intrinsic persistence and the degree of theory coherence of the time varying parameters can be optimally tuned

by maximizing the marginal data density of the TC-TVP-VAR which is available in closed form. Moreover, the TC-TVP-VAR can also be used as a tool for learning about the deep parameters from the underlying economic theory. As a matter of fact, the deep parameters from the economic theory are another set of hyper-parameters of the model that are indirectly estimated by projecting the TVP-VAR estimates onto the restrictions implied by the model from the economic theory. Thus, in this approach, learning about the deep parameters from the economic theory happens indirectly through learning about the TVP-VAR parameters.

In the paper I show that incorporating the restrictions implied by the economic theory into a prior for the coefficients of TVP-VARs can be beneficial to improve forecast accuracy and to obtain more precise estimates of typical objects of interest such as the impulse response functions. In particular, I find that encoding the restrictions from a conventional three equations New Keynesian model into a prior for the parameters of a trivariate TVP-VAR for output growth, inflation rate and the interest rate improves both point and density forecast accuracy of both output growth and the inflation rate at all the horizons considered (one quarter ahead, two quarters ahead and one year ahead). Then, I exploit the TC-TVP-VAR to investigate whether the US economy's performance was affected by a binding zero lower bound (ZLB) constraint as predicted by a standard New Keynesian model. Indeed, according to a standard New-Keynesian model, the economy is expected to exhibit a different response to demand and supply shocks when the ZLB constraint is in effect. However, and more importantly, the short length of the ZLB period in the US makes the standard TVP-VAR unfit to detect the change in the responses predicted by the NK model (Benati and Lubik, 2023). In other words, whether or not there was a change in the response of the economy during the ZLB as predicted by a standard NK model, cannot be directly inferred by using a standard TVP-VAR. Based on a simulation study, I show that the TC-TVP-VAR can in principle be used to solve this inferential problem. In particular, I exploit the time varying restriction functions

implied by a medium scale NK model that accounts for forward guidance and the ZLB period to parametrize the shrinkage prior in the TC-TVP-VAR. I show that this approach allows to estimate more precisely the response of the economy to macroeconomic shocks inside and outside the ZLB period, solving the inferential problems of the standard TVP-VAR. Finally, estimating the model on US data, I find that there are convincing evidences supporting a change in the response of the economy to risk premium shocks inside the ZLB period similar to the one predicted by a standard NK model. This finding has clearly important policy implications for the conduct of fiscal and macroprudential policies at the ZLB.

Related Literature This paper shows how to exploit prior information grounded on the basis of an economic theory to impose parsimony on the coefficients of TVP-VARs. In this aspect, the contribution conceptually borrows from ideas from the seminal work of Ingram and Whiteman (1994) and operationally from the insights in Del Negro et al. (2004) which show how to exploit the non-linear cross equation restrictions implied by a DSGE to form a prior for the parameters of a constant parameters VAR model. Extending the framework of Del Negro et al. (2004) to TVP-VARs is important at least for two reasons. First, because in macroeconomic applications the assumption of constant coefficients is often restrictive. Indeed, instabilities in the autoregressive coefficients of VARs used to model the dynamics of key macroeconomic indicators such as output and inflation have been widely documented in the literature (Cogley et al., 2002; Primiceri, 2005; D'Agostino, Gambetti, and Giannone, 2013). Especially in this setting, as the model becomes more flexible, additional shrinkage can be particularly beneficial to reduce overfitting. Second, because economic theories themselves might imply time-varying restriction functions for the coefficients. For example, macroeconomic theories assuming rational expectations extended so as to allow some parameters to vary according to Markov process with given transition probabilities, lead to state space representation with time varying coefficients

(Farmer, Waggoner, and Zha, 2009). At the same time, solutions for linear stochastic rational expectations models in the face of a finite sequence of anticipated structural changes lead to state space representation with time varying coefficients (Cagliarini and Kulish, 2013).[1] Likewise, outside the rational expectation framework, macroeconomic theories that assume learning can lead to state space representation with time varying coefficients (Milani, 2007). In all those cases, the proposed TC-TVP-VAR can be used both to incorporate the implied time varying restrictions into a prior for the time varying coefficients of the VAR and to estimate the deep parameters of the underlying economic model. In this sense, the paper is also related to the strand of literature that exploits an auxiliary flexible statistical model for the data to make indirect inference on the deep parameters of a structural model from the economic theory (see for example Gallant and McCulloch (2009) Fessler and Kasy (2019)).[2]

The paper is also related to the increasing number of studies that has recently focused on the issue of mitigating complexity and over-parametrization in TVP-VARs. One strand of literature has focused on identifying fixed versus time varying coefficients, by concentrating on the variance selection problem in the generic state equations of each of the TVP-VARs' coefficients. (Frühwirth-Schnatter and Wagner, 2010; Belmonte, Koop, and Korobilis, 2014; Kalli and Griffin, 2014; Bitto and Frühwirth-Schnatter, 2019; Huber, Koop, and Onorante, 2021). This literature has produced shrinkage priors for variances aimed at "automatically" reducing time-varying coefficients to static ones if the model is overfitting.[3] While treating the coefficients of the model as independent stochastic processes and just focusing on the problem of

[1] Another notable case within the rational expectations framework are models log-linearized around time varying trend for inflation (Cogley and Sbordone, 2008; Ascari and Sbordone, 2014; Ascari, Bonomolo, and Haque, 2023).

[2] In the TC-TVP-VAR the structural parameters from the underlying economic theory are estimated by implicitly minimizing the weighted discrepancy between the unrestricted TVP-VAR estimates and the restriction functions. This approach can be thought as a Bayesian version of Smith Jr. (1993) as in Del Negro et al. (2004).

[3] Similarly, from a frequentist perspective, Coulombe (2021) showed that time varying parameters can be framed as ridge regressions problems and used cross validation to tune the optimal amount of time variation in each of the state equations of the coefficients of the TVP-VAR.

tuning the optimal amount of time variation of the single coefficients, this strand of literature typically entirely neglects co-movement and correlation among the coefficients. However, in macroeconomic applications, the high degree of co-movement in the parameters is an empirical regularity. This fact was already found and stressed by Cogley et al. (2005) in one of the papers that introduced TVP-VARs in the field. In the same paper, the authors envisaged that the reduced-form parameters should move in a highly structured way because of the cross-equation restrictions suggesting that *"a formal treatment of cross-equation restrictions with parameter drift is a priority for future work"* (p. 274). More in line with these considerations, a smaller number of studies (Wind and Gambetti, 2014; Stevanovic, 2016; Chan, Eisenstat, and Strachan, 2020) proposed to use a factor structure to model the time variation of the parameters. Despite being compatible with the idea of the coefficients varying in a highly structured way because of the cross-equation restrictions associated with macroeconomic equations, this approach is purely statistical and abstracts from any macroeconomic theory disciplining the behavior of the coefficients. This paper fills this gap in the literature by showing how to exploit prior information grounded on the basis of an economic theory to state a priori a plausible correlation structure among the time varying parameters of the model. While focusing on economic theory as a source of potentially useful information on the time varying parameters, the paper methodologically contributes to literature on priors for TVP-VARs, by specifying a joint shrinkage prior for the whole history of the time varying parameters. This is done by writing the TVP-VAR in static compact form and exploiting band matrices to specify a joint prior for the time varying parameters.

Finally, the paper is more broadly related to the econometric literature showing that moment conditions from economic theory can successfully be exploited for forecasting macroeconomic and financial variables (Giacomini and Ragusa (2014) and Carriero, Clark, and Marcellino (2021) most notably).

Outline The paper is organized as follows. In section 2 I introduce the TC-TVP-VAR, presenting an analytical derivation of the proposed theory coherent prior and a simulation from the prior to showcase its main properties. In addition, I discuss the fit-complexity trade-off linked to the calibration of the hyper-parameters determining the intrinsic persistence of the time varying-coefficients and the degree of shrinkage towards the restrictions from the theory. I conclude section 2 by presenting an MCMC sampler used for estimation of the TC-TVP-VAR. Afterwards, in section 3 I exploit the well known 3-equations New Keynesian (NK) block to form a prior for the parameters of a TVP-VAR for GDP growth, inflation and the interest rate for the US economy. I compare the forecasts from a TC-TVP-VAR that encodes the restrictions from the NK model as a prior for the time varying coefficients to the forecasts from a standard TVP-VAR, showing that the former outperforms the latter both in terms of point and density forecast accuracy. Then, in section 4 I conduct a simulation study and show that a standard TVP-VAR struggles to detect the change in the response of the economy to macroeconomic shocks during the ZLB period as predicted by a standard NK model. I show that the TC-TVP-VAR can be used to solve this inferential problem. Finally, I use the TC-TVP-VAR to investigate whether the US economy was affected by a binding ZLB. Section 5 concludes.

2 Theory coherent TVP-VAR

The construction of a theory coherent prior builds on the idea that an economic theory implies restrictions on the parameters of the TVP-VAR. Intuitively we can find out these restrictions by specifying a prior distribution on the deep parameters from the theory, simulating the data from the theory and then estimating a TVP-VAR on the simulated data. Imposing a prior on the deep parameters from the theory will then induce a prior on the parameters of the TVP-VAR encoding the restrictions implied by the economic theory. Based on this idea, in this section I derive analytically

a prior for a TVP-VAR which is theory coherent, in the sense that it centers the time varying coefficients on the cross equation restrictions implied by an underlying economic theory about the variables in the system. I consider the case in which the population moments implied by the model from the economic theory are available in closed form as a function of the deep structural parameters. De facto, this allows to avoid stochastic simulation of the artificial observations from the economic theory and write the prior for the time varying parameters of the VAR directly as a function of the deep structural parameters from the economic theory.

Notation Before moving on, I introduce some notations conventions used in the paper. Scalars are in lowercase and normal weight. Vectors are in lowercase and in bold. Matrices are in uppercase and bold.

2.1 Construction of the prior

A TVP-VAR is given by:

$$\underbrace{\boldsymbol{y}_t'}_{1\times N} = \underbrace{\boldsymbol{x}_t'}_{1\times k}\underbrace{\boldsymbol{\Phi}_t}_{k\times N} + \underbrace{\boldsymbol{u}_t'}_{1\times N} \qquad \boldsymbol{u_t} \sim \mathcal{N}(\mathbf{0}_{N\times 1}, \boldsymbol{\Sigma_u}) \tag{1}$$

$$vec(\boldsymbol{\Phi_t}) = vec(\boldsymbol{\Phi_{t-1}}) + \boldsymbol{\eta}_t \qquad \boldsymbol{\eta}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}) \tag{2}$$

where $\boldsymbol{y_t}$ is an N dimensional random vector observed for $t = 1, \ldots, T$ periods, $\boldsymbol{x_t} = [1, \boldsymbol{y}_{t-1}', \ldots, \boldsymbol{y}_{t-p}']'$, p is the lag order and $k = 1 + Np$ is the number of coefficients in each equation of the VAR. For convenience, we can rewrite the TVP-VAR in (1) in "static" compact form as

$$\underbrace{\boldsymbol{Y}}_{T\times N} = \underbrace{\boldsymbol{X}}_{T\times Tk}\underbrace{\boldsymbol{\Phi}}_{Tk\times N} + \underbrace{\boldsymbol{U}}_{T\times N} \qquad \boldsymbol{U} \sim MVN(\mathbf{0}, \boldsymbol{\Sigma_u}, \boldsymbol{I_T}) \tag{3}$$

and the state equations in (2) as

$$\boldsymbol{H}_{Tk}\boldsymbol{\Phi} = \boldsymbol{\Phi_{00}} + \boldsymbol{\eta} \tag{4}$$

which graphically is

$$\underbrace{\begin{bmatrix} \boldsymbol{I}_k & \mathbf{0} & \dots & \mathbf{0} \\ -\boldsymbol{I}_k & \boldsymbol{I}_k & \dots & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{I}_k & \ddots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & -\boldsymbol{I}_k & \boldsymbol{I}_k \end{bmatrix}}_{\boldsymbol{H_{Tk}}} \underbrace{\begin{bmatrix} \boldsymbol{\Phi_1} \\ \boldsymbol{\Phi_2} \\ \boldsymbol{\Phi_3} \\ \vdots \\ \boldsymbol{\Phi_T} \end{bmatrix}}_{\boldsymbol{\Phi}} = \underbrace{\begin{bmatrix} \boldsymbol{\Phi_0} \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}}_{\boldsymbol{\Phi_{00}}} + \underbrace{\begin{bmatrix} \boldsymbol{\eta}_1 \\ \boldsymbol{\eta}_2 \\ \boldsymbol{\eta}_3 \\ \vdots \\ \boldsymbol{\eta}_T \end{bmatrix}}_{\boldsymbol{\eta}}$$

In this representation, the matrix $\boldsymbol{\Phi}$ stores the matrices with the time varying coefficients one on the top of the other, for all the time periods $t = 1, \dots, T$. I assume that an hyperparameter λ governs the serial correlation of the coefficients stored in the matrices $\boldsymbol{\Phi_1}, \dots, \boldsymbol{\Phi_T}$. In particular, I assume that the variance covariance matrix $\boldsymbol{\Omega}$ has the Kronecker structure $\boldsymbol{\Omega} = \boldsymbol{\Sigma_u} \otimes (\lambda^2 \boldsymbol{I_k})^{-1}$. This assumption implies that the variance in the state equation of the coefficients is proportional to the variance of the innovations of the equation to which the coefficients appertain σ_{ii}^2. In practice this means that the generic state equation of the coefficient attached to the j^{th} regressor in the i^{th} equation reads as follows

$$\phi_{jt}^{(i)} = \phi_{jt-1}^{(i)} + \eta_{jt}^{(i)} \qquad\qquad \eta_{jt}^{(i)} \sim \mathcal{N}\left(0, \frac{\sigma_{ii}^2}{\lambda^2}\right) \tag{5}$$

with $i = 1, \dots, N$ and $j = 1, \dots, k$. As $\lambda \to 0$ the variance of the *Normal* prior centering a coefficient at time t on the realization at time $t-1$ increases, meaning that this prior becomes more and more diffuse and the coefficients of two consecutive time periods are not forced a priori to be close to each other. Conversely as $\lambda \to \infty$ the prior centering the coefficient at time t on the realization at time $t-1$ becomes

more and more tight. Clearly the calibration of λ involves a fit complexity trade off since by letting $\lambda \to \infty$ we allow the coefficient at time t to be arbitrary distant from the coefficient at $t-1$ up to the point that we can almost fit the data perfectly in sample.[4] Exploiting representation (4) and the Kronecker structure of $\boldsymbol{\Omega}$, we can conclude that equation (4) implies a joint normal prior for $\boldsymbol{\Phi}$, that is

$$vec(\boldsymbol{\Phi}|\boldsymbol{\Sigma}_u, \lambda) \sim \mathcal{N}(vec(\underbrace{\boldsymbol{H}_{\boldsymbol{Tk}}^{-1}\boldsymbol{\Phi}_{\mathbf{00}}}_{\underline{\Phi}_0}), \boldsymbol{\Sigma}_u \otimes \underbrace{(\lambda^2 \boldsymbol{H}'_{Tk}\boldsymbol{H}_{Tk})^{-1}}_{\underline{\Psi}(\lambda)}) \tag{6}$$

since $|\boldsymbol{H}_{Tk}| = 1$, $\boldsymbol{H}_{Tk}$ is always invertible and the determinant of Jacobian matrix of the linear transformation cancels out.[5] Combining this prior for the time varying parameters with an *Inverse-Wishart* prior for $\boldsymbol{\Sigma}_u$

$$p(\boldsymbol{\Sigma_u}) \sim \mathcal{IW}(\underline{\mathbf{S}}, \underline{\nu}) \tag{7}$$

leads to

$$p(\boldsymbol{\Phi}, \boldsymbol{\Sigma_u}|\lambda) \sim \mathcal{NIW} \tag{8}$$

Equation (8) states a prior for the time varying parameters and the variance covariance matrix of the innovations of the TVP-VAR which is conditional on λ, the crucial hyper-parameter determining the amount of time variation of the coefficients in the state equations (2). [6] Importantly, this prior is conjugate to the Gaussian likelihood. We exploit the conjugacy to update equation (8) with the artificial observations from the theory and obtain a *Normal-Inverse-Wishart* prior which is theory coherent, meaning that it is centered on the restrictions from the economic theory.[7]

[4] This point will be covered more in detail in Section 2.4.

[5] Notice that this is true also if we consider an AR(1) dynamics for the time varying coefficients. In that case the elements off the main diagonal of the band matrix H_{Tk} would store the AR coefficients of the state equations.

[6] To be precise this prior is also conditional on $\boldsymbol{\Phi_0}$, $\underline{\boldsymbol{S}}$ and $\underline{\nu}$. However, for the purpose of the paper I will consider $\boldsymbol{\Phi_0}$, $\underline{\boldsymbol{S}}$ and $\underline{\nu}$ as fixed.

[7] Note that equivalently we can consider a flat prior for $\boldsymbol{\Sigma_u}$, that is:

$$p(\boldsymbol{\Sigma_u}) \propto |\boldsymbol{\Sigma_u}|^{-\frac{N+1}{2}} \tag{9}$$

In the specific, updating (8) with observations from the theory we get

$$p(\boldsymbol{\Phi}, \boldsymbol{\Sigma_u}|\lambda, \boldsymbol{\theta}, \gamma) = c(\lambda, \boldsymbol{\theta}, \gamma)^{-1} p(\boldsymbol{\Phi}, \boldsymbol{\Sigma_u}|\lambda) p(\boldsymbol{Y(\theta)}|\gamma, \boldsymbol{\Phi}, \boldsymbol{\Sigma_u}) \tag{10}$$

where $p(\boldsymbol{Y(\theta)}|\boldsymbol{\Phi}, \boldsymbol{\Sigma_u}, \gamma)$ is the likelihood of γ simulated samples of $\boldsymbol{Y(\theta)}$ from the theory and

$$c(\lambda, \boldsymbol{\theta}, \gamma) = \int_{-\infty}^{\infty} p(\boldsymbol{\Phi}, \boldsymbol{\Sigma_u}|\lambda) p(\boldsymbol{Y(\theta)}|\gamma, \boldsymbol{\Phi}, \boldsymbol{\Sigma_u}) d\boldsymbol{\Phi} d\boldsymbol{\Sigma_u} \tag{11}$$

is an integrating constant which ensures that the prior density is proper and integrates to one.[8] Equation (10) makes clear that the theory-coherent prior for $(\boldsymbol{\Phi}, \boldsymbol{\Sigma_u})$ is obtained estimating a TVP-VAR on the simulated data from the theory. As a matter of fact, equation (10) is just the posterior distribution of $(\boldsymbol{\Phi}, \boldsymbol{\Sigma_u})$ obtained by updating the hierarchical prior $p(\boldsymbol{\Phi}, \boldsymbol{\Sigma_u}|\lambda)$ with the likelihood of γ simulated samples of observations generated by the model from the theory. As in Del Negro et al. (2004), in $p(\boldsymbol{Y(\theta)}|\boldsymbol{\Phi}, \boldsymbol{\Sigma_u})$ I replace the sample moments $\boldsymbol{Y(\theta)'Y(\theta)}$, $\boldsymbol{Y(\theta)'X(\theta)}$, and $\boldsymbol{X(\theta)'X(\theta)}$ by their expected values, obtaining [9]

$$p(\boldsymbol{Y(\theta)}|\gamma, \boldsymbol{\Phi}, \boldsymbol{\Sigma_u}) = (2\pi)^{-\frac{\gamma T N}{2}} |\boldsymbol{\Sigma_u}|^{-\frac{\gamma T}{2}} exp\left[-\frac{1}{2} tr\left((\boldsymbol{\Sigma_u}^{-1})(\gamma \boldsymbol{\Gamma_{yy}(\theta)} - \gamma \boldsymbol{\Phi' \Gamma_{xy}(\theta)}) - \gamma \boldsymbol{\Gamma_{yx}(\theta)\Phi} + \gamma \boldsymbol{\Phi' \Gamma_{xx}(\theta)\Phi})\right)\right] \tag{12}$$

This step is meant to avoid stochastic simulation of the artificial observations from the economic theory and it let us write the likelihood of the artificial observations directly as a function of the vector of deep parameters from the economic theory $\boldsymbol{\theta}$.[10] It is easy to show that (10) is also *Normal-Inverse-Wishart* since it is de-facto obtained as the posterior distribution that combines a *Normal-Inverse Wishart* prior for $(\boldsymbol{\Phi}, \boldsymbol{\Sigma_u})$ with a *Gaussian* likelihood for the simulated data from the theory. Therefore the theory coherent prior in equation (10) takes the form

---

[8] In the appendix A.1.3 I report the details on the integrating constant.
[9] The notation $Y(\boldsymbol{\theta})$ makes clear that we are conditioning on the vector $\boldsymbol{\theta}$.
[10] Details on $\boldsymbol{\Gamma_{xx}}$, $\boldsymbol{\Gamma_{yy}}$, $\boldsymbol{\Gamma_{xy}}$ are reported in the appendix A.1.2.

$$
\begin{aligned}
p(\boldsymbol{\Sigma}_u) &\sim \mathcal{IW}\left(\underline{\underline{\boldsymbol{S}}}, \underline{\underline{\nu}}\right) \\
p(vec(\boldsymbol{\Phi})|\boldsymbol{\Sigma}_u) &\sim \mathcal{N}(vec(\underline{\underline{\boldsymbol{\Phi}}}), \boldsymbol{\Sigma}_u \otimes \underline{\underline{\boldsymbol{\Psi}}})
\end{aligned} \tag{13}
$$

$$
\underline{\underline{\boldsymbol{S}}} = \underline{\underline{\boldsymbol{S}}} + \gamma \boldsymbol{\Gamma_{yy}}(\boldsymbol{\theta}) + \underline{\boldsymbol{\Phi_0}}'(\lambda^2 \boldsymbol{H'_{Tk}} \boldsymbol{H_{Tk}}) \underline{\boldsymbol{\Phi_0}} - \underline{\underline{\boldsymbol{\Phi}}}' \underline{\underline{\boldsymbol{\Psi}}}^{-1} \underline{\underline{\boldsymbol{\Phi}}} \tag{14}
$$

$$
\underline{\underline{\nu}} = \underline{\nu} + T\gamma \tag{15}
$$

$$
vec(\underline{\underline{\boldsymbol{\Phi}}}) = vec(\left(\gamma \boldsymbol{\Gamma_{xx}}(\boldsymbol{\theta}) + \lambda^2 \boldsymbol{H'_{Tk}} \boldsymbol{H_{Tk}}\right)^{-1} (\gamma \boldsymbol{\Gamma_{xy}}(\boldsymbol{\theta}) + \lambda^2 \boldsymbol{H'_{Tk}} \boldsymbol{H_{Tk}} \underline{\boldsymbol{\Phi_0}})) \tag{16}
$$

$$
\underline{\underline{\boldsymbol{\Psi}}} = \left(\gamma \boldsymbol{\Gamma_{xx}}(\boldsymbol{\theta}) + \lambda^2 \boldsymbol{H'_{Tk}} \boldsymbol{H_{Tk}}\right)^{-1} \tag{17}
$$

This prior for $\boldsymbol{\Phi}$ encompasses two different pieces of information about the time varying coefficients. The first piece of information is their intrinsic persistence determined by the shrinkage hyperparameter λ. The second instead, is the degree of theory coherence determined by the shrinkage hyperparameter γ which defines the tightness of the *Normal* prior around the restriction function defined by the economic theory

$$
\boldsymbol{\Phi}(\boldsymbol{\theta})^* = \boldsymbol{\Gamma_{xx}}(\boldsymbol{\theta})^{-1} \boldsymbol{\Gamma_{xy}}(\boldsymbol{\theta}) \tag{18}
$$

which implies $\boldsymbol{\Phi}_t = \boldsymbol{\Gamma}^{-1}_{\boldsymbol{xx,t}} \boldsymbol{\Gamma_{xy,t}}$ for $t = 1, \ldots, T$ where $\boldsymbol{\Gamma_{xx,t}} \equiv \mathbb{E}[\boldsymbol{x_t x'_t}|\boldsymbol{\theta}]$ and $\boldsymbol{\Gamma_{xy,t}} \equiv \mathbb{E}[\boldsymbol{x_t y'_t}|\boldsymbol{\theta}]$. A couple of considerations are worthwhile. First, the assumption of a unique hyperparameter λ determining the amount of time variation of the time varying coefficients can be easily replaced by assuming λ_j's hyper-parameters with $j = 1, \ldots, k$ such that $\boldsymbol{\Omega}$ can be factorized as

$$
\boldsymbol{\Omega} = \boldsymbol{\Sigma} \otimes (\boldsymbol{\Lambda'_k} \boldsymbol{\Lambda_k})^{-1} \tag{19}
$$

where $\boldsymbol{\Lambda_k} = diag(\lambda_1, \ldots, \lambda_k)$ and the Kroneker structure of the *Normal-Inverse-Wishart* prior (6) is preserved. Keeping the assumption of a Kronecker structure allows to have a closed form expression for the likelihood of γ, λ and $\boldsymbol{\theta}$ marginally of the time varying parameters $\boldsymbol{\Phi_1}, , \ldots, \boldsymbol{\Phi_T}$ and of $\boldsymbol{\Sigma_u}$. In practice this allows to

make inference on $\boldsymbol{\Phi}, \boldsymbol{\Sigma}_u$, the shrinking hyper-parameters γ and λ and also on the deep parameters from the theory $\boldsymbol{\theta}$ jointly as in Del Negro et al. (2004). Second, the population moments encoded in the TVP-VAR $\boldsymbol{\Gamma}_{\boldsymbol{xx}}$, $\boldsymbol{\Gamma}_{\boldsymbol{xy}}$, $\boldsymbol{\Gamma}_{\boldsymbol{yy}}$, which are assumed to be available in closed form as function of the deep parameters from the theory $\boldsymbol{\theta}$ might not be available in some applications. If it is the case, these moments can be obtained by stochastic simulation as in Loria, Matthes, and Wang (2022). Furthermore, and more importantly, the moments stored in the matrices $\boldsymbol{\Gamma}_{\boldsymbol{xx}}$, $\boldsymbol{\Gamma}_{\boldsymbol{xy}}$, $\boldsymbol{\Gamma}_{\boldsymbol{yy}}$ can be either constant or time varying conditionally on the deep parameters from the theory. That is, the prior in (13) can accommodate theories which imply both constant and time-varying restrictions. This is the main innovation to the DSGE-VAR toolkit developed by Del Negro et al. (2004).

2.2 Simulation of the prior

To showcase the properties of the proposed theory coherent prior, figure 1 shows random draws for the time series of a generic $\phi_{jt}^{(i)}$ coefficient of the TVP-VAR for $t = 1, \ldots, 250$ from the prior. The dashed black line represents the restriction function (18) defined by the economic theory for that coefficient, that I label $\phi_{jt}^{(i)}(\boldsymbol{\theta})^*$. Just for illustrative purposes for now I consider the case of a constant restriction function, however having time varying moments stored in the matrices $\boldsymbol{\Gamma}_{\boldsymbol{xx}}$ and $\boldsymbol{\Gamma}_{\boldsymbol{xy}}$ conditionally on the deep parameters of the theory, would directly imply a time varying restriction function trough (18). The first row shows five draws from the theory coherent prior for different values of λ conditioning on a positive γ, meaning that for a given degree of theory coherence I change the value of hyper-parameter which determines the persistence of the coefficient. When $\lambda = 0$, the draws for the coefficient $\phi_{jt}^{(i)}$ are independent for $t = 1, \ldots, T$ and centered on the restriction function coming from the theory. Hence, nothing prevents a coefficient at time t to be arbitrary distant from a coefficient at time $t - 1$ except the fact both coefficients are centered, with a precision determined by γ, on the restriction function defined by the theory. As λ

increases the time series of the coefficient becomes more and more persistent up to the point that when λ is very big, the coefficient becomes almost constant over time. Instead, in the second row, I let the degree of theory coherence change conditionally on a given intrinsic persistence of the time varying coefficients. In other words, I show random draws for $\phi_{jt}^{(i)}$ for different values of γ conditionally on $\lambda > 0$. When $\gamma = 0$, the coefficients are just random walks with variance equal to $\frac{\sigma_{ii}}{\lambda^2}$, and they are completely unrelated to the restriction function implied by the economic theory $\phi_{jt}^{(i)}(\boldsymbol{\theta})^*$.[11] As γ increases the draws for the coefficient are centered on the restriction function defined by the theory with an increasing precision and on the limit, with $\gamma \to \infty$, they go to $\phi_{jt}^{(i)}(\boldsymbol{\theta})^*$.[12] Importantly, the hyperparameter γ is contemporaneously shrinking all the time varying coefficients $\phi_{t,j}^{(i)}$ in the equations of the VAR towards the restrictions implied by the economic theory. This reflects the idea that it is economic theory that a priori postulates a plausible correlation structure among the coefficients of VAR model through the restriction function in (18).

[11] Note that the draws fore the time varying coefficients were initiated near the restriction functions just for visualization purposes.

[12] The coefficients would have a degenerate distribution, with point mass on the restriction function.

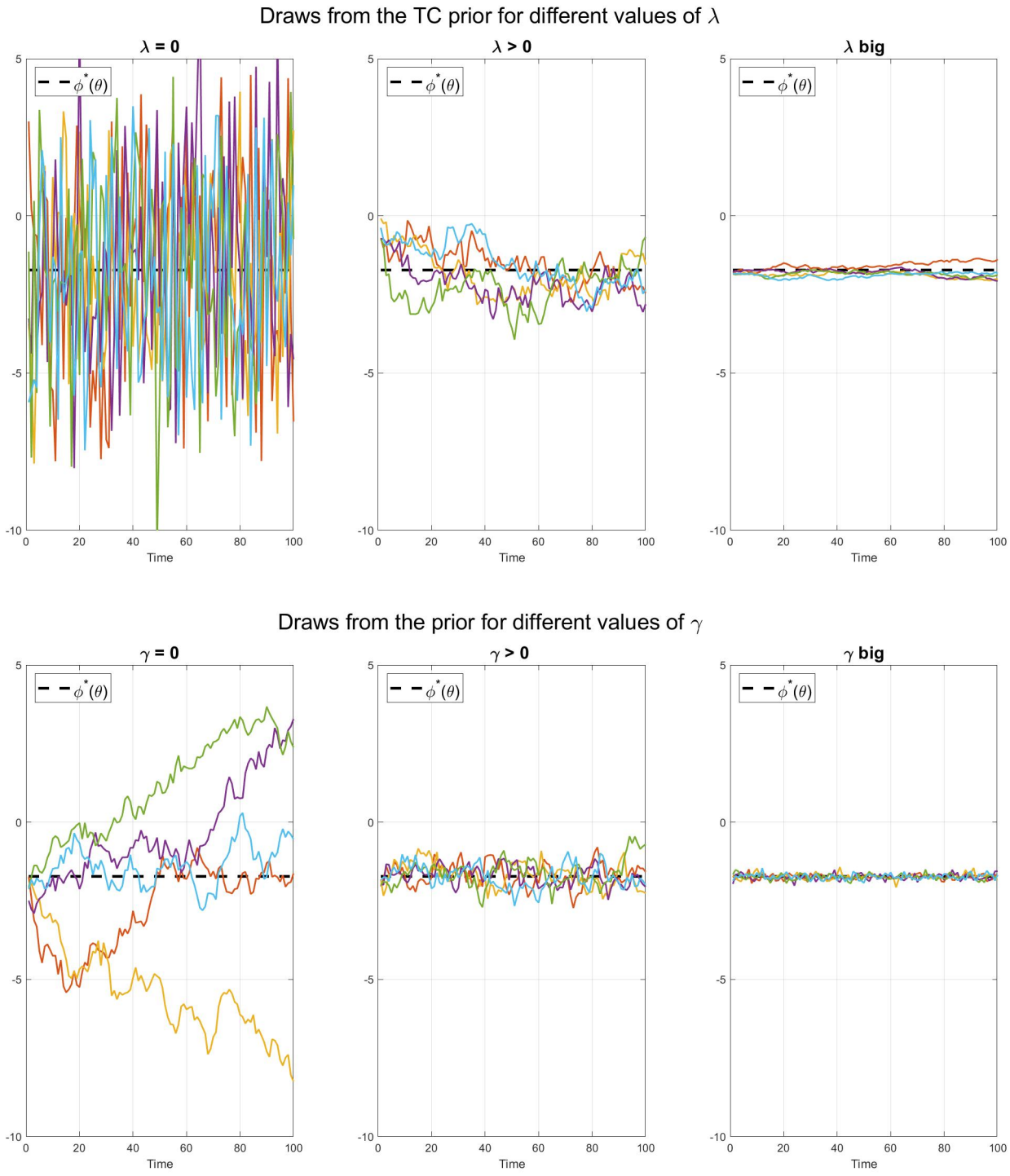


Figure 1: Each subfigure presents five draws from the theory coherent prior for a generic coefficient of the TVP-VAR for different values of the hyperparameters λ (shrinkage towards constant) and γ (theory coherent shrinkage). The first row, presents draws for $\lambda = 0$, $\lambda > 0$ and a big λ, conditionally on $\gamma > 0$. The second row presents draws for $\gamma = 0$, $\gamma > 0$ and a big γ, conditionally on $\lambda > 0$. The dashed black line represents the restriction functions coming from the theory.

It is important to remark that this framework can accommodate both constant and time-varying restrictions for the coefficients through (18). This allows to extend the DSGE-VAR toolkit developed in Del Negro et al. (2004) to a broader set of chosen economic theories, which might imply time varying restriction functions for the parameters of the VAR. To show this property, I exploit the medium scale New Keynesian model in Del Negro, Giannoni, and Schorfheide (2015) to parametrize a prior for a medium scale TVP-VAR for output growth, consumption growth, investment growth, real wage growth, hours worked, inflation and the Fed Fund rate. The NK model accounts for the ZLB and forward guidance and it is solved using the method proposed by Cagliarini et al. (2013) for linear rational equation expectation systems in the face of anticipated structural changes. The solution of the model implies a state space representation that exhibits time varying coefficients. Once the population moments implied by the state space representation are used to parameterize (18) we obtain time varying restrictions functions for the time varying coefficients.[13] Figure 2 makes this point by showing the implied time varying restrictions on the coefficient of the first lag of the inflation rate in the equation of the Fed Fund Rate. Together with the time varying restriction function, the figure shows draws from the prior for different values of the hyper-parameter γ conditioning on a given value of λ and a set of structural parameters of the NK model. As expected, as γ increases the draws for the time varying coefficient are centered with increasing precision on the restriction function defined by the economic theory. More in the specific, while outside the zero lower bound the Fed Fund rate is expected to increase when the lagged value of inflation increases, inside the zero lower bound the interest rate is not expected to respond to the lagged value of the inflation rate and therefore the prior mean for this coefficients becomes centered around zero.

[13]The details on the derivation of the moments are available in the section 4.

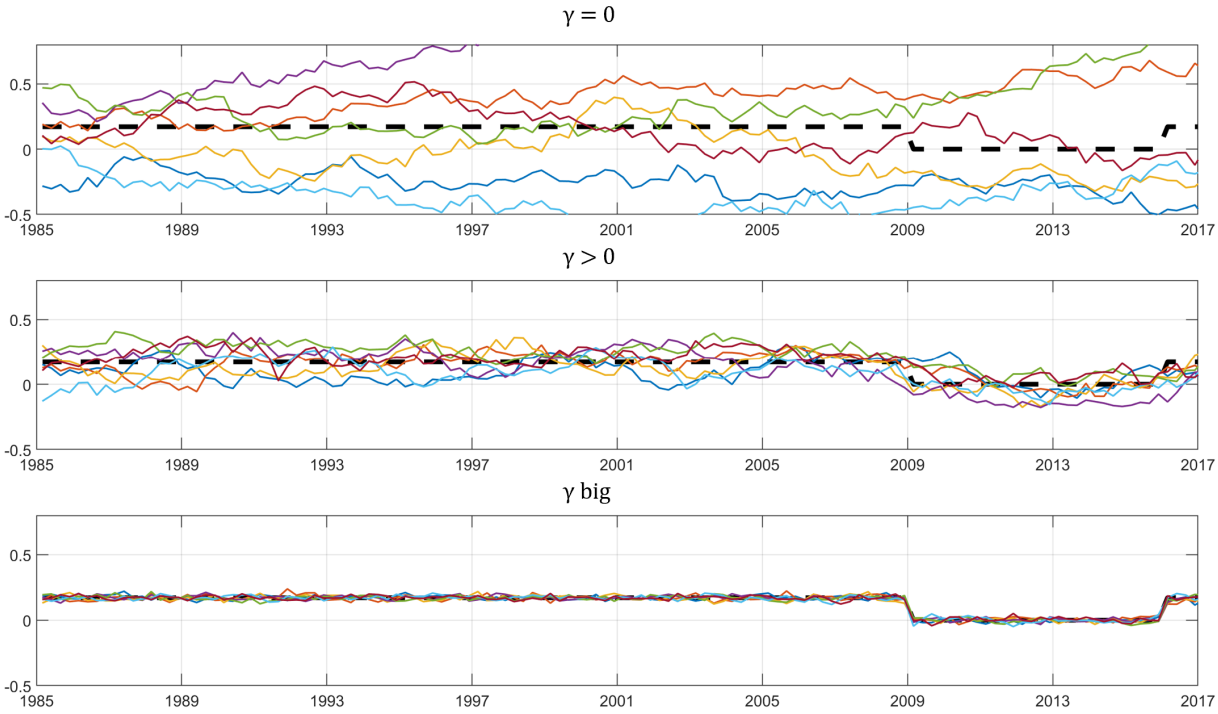


Figure 2: The figure shows draws from the prior for the coefficient on the first lag of the inflation rate in the equation of the Fed Fund rate for different values of the hyperparameter γ. The model is a seven variable TVP-VAR for the US economy. The dashed black line represents the restriction function implied by the medium scale NK model.

2.3 Conditional posterior of Φ and Σ_u

It is straightforward to derive the posterior distribution of $(\boldsymbol{\Phi}, \boldsymbol{\Sigma}_u)$ conditional on the hyper-parameters λ, γ and on the deep parameters from the theory $\boldsymbol{\theta}$ since this is just a *Normal-Inverse-Wishart* distribution obtained by updating a Gaussian likelihood with the *Normal-Inverse-Wishart* prior (13), namely

$$\begin{aligned} p(vec(\boldsymbol{\Phi})|\boldsymbol{\Sigma}_u, \gamma, \lambda, \boldsymbol{\theta}, \boldsymbol{Y}) &\sim \mathcal{N}(vec(\tilde{\boldsymbol{\Phi}}), \boldsymbol{\Sigma}_u \otimes \tilde{\boldsymbol{\Psi}}) \\ p(\boldsymbol{\Sigma}_u|\gamma, \lambda, \boldsymbol{\theta}, \boldsymbol{Y}) &\sim \mathcal{IW}\left(\tilde{\boldsymbol{S}}, \tilde{\nu}\right) \end{aligned} \tag{20}$$

$$vec(\tilde{\boldsymbol{\Phi}}) = vec(\left(\boldsymbol{X'X} + \gamma\boldsymbol{\Gamma_{xx}}(\boldsymbol{\theta}) + \lambda^2\boldsymbol{H'_{Tk}}\boldsymbol{H_{Tk}}\right)^{-1}\left(\boldsymbol{X'Y} + \gamma\boldsymbol{\Gamma_{xy}}(\boldsymbol{\theta}) + \lambda^2\boldsymbol{H}'_{Tk}\boldsymbol{H}_{Tk}\underline{\boldsymbol{\Phi}}_{\boldsymbol{0}}\right)) \tag{21}$$

$$\tilde{\boldsymbol{\Psi}} = \left(\boldsymbol{X'X} + \gamma\boldsymbol{\Gamma_{xx}}(\boldsymbol{\theta}) + \lambda^2\boldsymbol{H'_{Tk}}\boldsymbol{H_{Tk}}\right)^{-1} \tag{22}$$

$$\tilde{\boldsymbol{S}} = \boldsymbol{Y'Y} + \underline{\underline{\boldsymbol{S}}} + \underline{\underline{\boldsymbol{\Phi}}}'\underline{\underline{\boldsymbol{\Psi}}}^{-1}\underline{\underline{\boldsymbol{\Phi}}} - \tilde{\boldsymbol{\Phi}}'\tilde{\boldsymbol{\Psi}}^{-1}\tilde{\boldsymbol{\Phi}} \tag{23}$$

$$\tilde{\nu} = \underline{\underline{\nu}} + T \tag{24}$$

Equation (21) makes it clear that thanks to the conjugacy of the prior the formula for the mean of the conditional posterior of $\boldsymbol{\Phi}$ boils to a standard OLS regression formula based on a sample augmented with a set of dummy observations that determines the degree of persistence of the time varying coefficients and another set of dummy observations that centers the coefficients on the restriction function implied by the theory. While the first set of dummy observations shapes the correlation of the time varying coefficients over time by imposing a set of linear restrictions as a function of the hyper-parameter λ, the second set of dummy observations induces at each time period a correlation structure implied by the non-linear cross equation restrictions coming from the theory, as a function of the deep parameters from the theory $\boldsymbol{\theta}$ and of the shrinking parameter γ.[14] It is worth to remark that from a computational perspective equation (21) makes the estimation of the time varying coefficients very

[14] In the appendix A.1.1 I show that thanks to the Kronecker structure of the prior (6) the time variation of the coefficients can be modelled by dummy observations.

efficient since it allows to draw all the history of the time varying coefficients in all the equations of the VAR in a single step, avoiding forward filtering and backward smoothing algorithms à la Carter and Kohn (1994). [15]

2.4 λ, γ and fit complexity trade off

Tuning of the optimal degree of theory coherence and the intrinsic amount of time variation of the coefficients is a delicate matter, since it clearly involves a fit-complexity trade off. As a matter of fact, very low values of both γ and λ imply a priori that the coefficients in two consecutive time periods can potentially be very distant from each others and from the restriction functions defined by the theory.[16] Intuitively, this model will fit the data very well in sample but will perform badly for forecasting out-of-sample. Indeed, decreasing γ and λ will in general increase in-sample fit of the model at the expense of out-of-sample accuracy. Based on this argument, I recommend to base the optimal choice of both the hyper-parameters on the maximization of the marginal likelihood of the model or equivalently on the maximization of the posterior of the hyper-parameters λ and γ under a flat prior for these hyper-parameters. This translates into maximizing the one-step-ahead out-of-sample forecasting ability of the model. Indeed, the log-marginal likelihood (or Bayesian evidence) can be interpreted as the sum of the one step-ahead predictive scores, since it is equivalent to the scoring rule of the form

$$S(\boldsymbol{Y}) = \sum_{t=1}^{T} s(\boldsymbol{y}_t|\boldsymbol{y}_{t-1}) = \sum_{t=1}^{T} log(p(\boldsymbol{y}_t|\boldsymbol{y}_{t-1})) \tag{25}$$

[15] Indeed, as in the precision sampler by Chan and Jeliazkov (2009), we can draw all the latent states from $t = 1, \dots, T$ in a single step and thanks to the Kronecker structure of the posterior, we can do it for all the N equations of the TVP-VAR jointly. More in general, the Kronecker structure (19) coupled with the precision sampler by Chan et al. (2009) can be exploited to estimate medium to large scale TVP-VARs.

[16] When both $\gamma = 0$ and $\lambda = 0$ the model is left totally unrestricted, with the prior variance covariance of the coefficients being equal to infinity. Clearly, in this case there are more parameters than you can feasibly estimate with a flat prior meaning that the conditional posterior of $\boldsymbol{\Phi}$ cannot be computed due to the non-invertibility of $\boldsymbol{X'X}$ (this can be seen from equation (20)).

An attractive feature of the TC-TVP-VAR is that the marginal likelihood $p(\boldsymbol{Y}|\gamma, \lambda, \boldsymbol{\theta})$ obtained by integrating out $\boldsymbol{\Phi}, \boldsymbol{\Sigma}_u$ from the conditional posterior $p(\boldsymbol{\Phi}, \boldsymbol{\Sigma_u}|\lambda, \boldsymbol{\theta}, \gamma, \boldsymbol{Y})$ is available in closed form and it is equal to

$$p(\boldsymbol{Y}|\lambda, \boldsymbol{\theta}, \gamma) = (\pi)^{-\frac{TN}{2}} \frac{\Gamma_N\left(\frac{\tilde{\nu}}{2}\right)|\tilde{S}|^{-\frac{\tilde{\nu}}{2}}|\tilde{\boldsymbol{\Psi}}|^{\frac{N}{2}}}{\Gamma_N\left(\frac{\underline{\underline{\nu}}}{2}\right)|\underline{\underline{S}}|^{-\frac{\underline{\underline{\nu}}}{2}}|\underline{\underline{\boldsymbol{\Psi}}}|^{\frac{N}{2}}} \tag{26}$$

As a consequence, calibrating γ and λ to maximize (26) corresponds to finding γ and λ maximizing the one-step-ahead out-of-sample forecasting ability of the model. This strategy of estimating hyper-parameters by maximizing the marginal likelihood is an empirical Bayes method which has a clear frequentist interpretation. In what follows, and in particular in the estimation algorithm detailed in the next section 2.5 I will regard γ and λ as random variables and perform full posterior inference on the hyper-parameters, but analogously maximizing the posterior of the hyper-parameters will correspond to maximizing the one-step-ahead out of sample forecastability of the model. Following the same steps as in Giannone, Lenza, and Primiceri (2015), we can rewrite equation (26) as

$$p(\boldsymbol{Y}|\lambda, \boldsymbol{\theta}, \gamma) \propto |(\boldsymbol{V_\varepsilon^{post}})^{-1}\boldsymbol{V_\varepsilon^{prior}}|^{\frac{T+\underline{\underline{\nu}}}{2}} \prod_{t=1}^{T} |\boldsymbol{V_{t|t-1}}|^{-\frac{1}{2}} \tag{27}$$

where $\boldsymbol{V_\varepsilon^{post}}$ and $\boldsymbol{V_\varepsilon^{prior}}$ are the posterior and prior means (or modes) of the residual variance, while $\boldsymbol{V_{t|t-1}}$ is equal to the variance (conditional on $\boldsymbol{\Sigma_u}$) of the one-step-ahead forecast of $\boldsymbol{y}$, averaged across all possible a-priori realizations of $\boldsymbol{\Sigma_u}$. Complete formulas of these objects are reported in the appendix A.1.4 and are the analog of the formulas reported in Giannone et al. (2015). Equation (27) makes clear that the marginal likelihood involves two terms: a reward for model fit, $|(\boldsymbol{V_\varepsilon^{post}})^{-1}\boldsymbol{V_\varepsilon^{prior}}|^{\frac{T+\underline{\underline{\nu}}}{2}}$ and a penalty term for model complexity $\prod_{t=1}^{T} |\boldsymbol{V_{t|t-1}}|^{-\frac{1}{2}}$. Figure 3 plots the model fit term and the penalty term of the marginal likelihood as a function of the two hyper-parameters γ and λ conditionally on a set of deep parameters from the theory

$\boldsymbol{\theta}$. As λ decreases the tightness of the prior centering each coefficient on its realization in the previous period eases and in sample model fit improves, meaning that $\boldsymbol{V}^{post}$ decreases. However, at the same time, as λ decreases $\boldsymbol{V_{t|t-1}}$ increases, since the variance (conditional on $\boldsymbol{\Sigma}$) of the one-step-ahead forecast of $\boldsymbol{y}$ increases. Analogously, as γ decreases the restriction functions from the theory become less and less binding, meaning that the model will fit better in sample, namely $\boldsymbol{V}^{post}$ decreases, but the variance of the one step ahead forecast error $\boldsymbol{V_{t|t-1}}$ will increase.

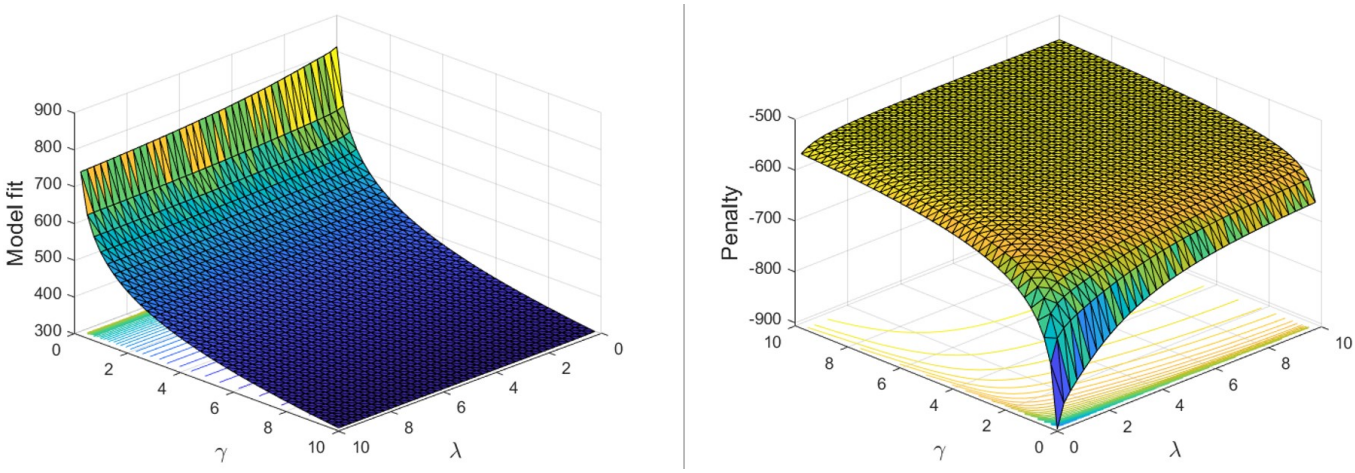


Figure 3: The figure shows the model fit term (on the left) and the penalty term (on the right) of the marginal likelihood (26) as a function of the hyper-parameters λ (persistency) and γ (theory coherence).

2.5 Estimation

In this section I describe a posterior simulation sampler to make inference on the TVP-VAR parameters $\boldsymbol{\Phi}$ and $\boldsymbol{\Sigma}_u$ and on the shrinkage hyper-parameters λ and γ together with the deep parameters from the theory $\boldsymbol{\theta}$.[17] We can simulate draws from the posterior $p(\boldsymbol{\Phi}, \boldsymbol{\Sigma_u}, \lambda, \boldsymbol{\theta}, \gamma | \boldsymbol{Y})$ exploiting the factorization:

$$p(\boldsymbol{\Phi}, \boldsymbol{\Sigma_u}, \lambda, \boldsymbol{\theta}, \gamma | \boldsymbol{Y}) = p(\boldsymbol{\Phi}, \boldsymbol{\Sigma_u} | \lambda, \boldsymbol{\theta}, \gamma, \boldsymbol{Y}) p(\lambda, \boldsymbol{\theta}, \gamma | \boldsymbol{Y}) \tag{28}$$

[17]For the purpose of the paper we fix $\boldsymbol{\Phi}_0 = \hat{\boldsymbol{\Phi}}_{\boldsymbol{0}}$ and $\underline{\boldsymbol{S}} = \hat{\boldsymbol{S}} = diag(\hat{s}_1^2, \ldots, \hat{s}_N^2)$ and $\nu = N + 2$. Alternatively inference on these hyper-parameters can be made by extending the Random Walk Metropolis step in the MCMC sampler described below.

This factorization makes clear that we can sample from the posterior of $\boldsymbol{\Phi}, \boldsymbol{\Sigma_u}, \boldsymbol{\theta}, \gamma, \lambda$ building a MCMC algorithm that iterates the following two steps:

1. Draw from $p(\lambda, \boldsymbol{\theta}, \gamma | \boldsymbol{Y}) \propto p(\boldsymbol{Y} | \lambda, \boldsymbol{\theta}, \gamma) p(\lambda, \boldsymbol{\theta}, \gamma)$.

2. Draw from $p(\boldsymbol{\Phi}, \boldsymbol{\Sigma_u} | \lambda, \boldsymbol{\theta}, \gamma, \boldsymbol{Y})$

In the first step, estimation of both the shrinkage parameters γ and λ and of deep parameters from the theory $\boldsymbol{\theta}$ happens by posterior evaluation of

$$p(\lambda, \boldsymbol{\theta}, \gamma | Y) \propto p(\boldsymbol{Y} | \lambda, \boldsymbol{\theta}, \gamma) p(\boldsymbol{\theta}) p(\gamma) p(\lambda) \tag{29}$$

where $p(\gamma)$ and $p(\lambda)$ are the priors for the shrinkage hyper-parameters [18] while $p(\boldsymbol{\theta})$ is the prior of the deep parameters from the theory. Hence, learning about the structural parameters happens implicitly by projecting the VAR estimates onto the restrictions implied by the model from the theory. More precisely, the estimates of the deep parameters minimizes the weighted discrepancy between the TVP-VAR unrestricted estimates and the restriction function (18). This approach can be thought as a Bayesian version of Smith Jr. (1993) and was pioneered by Del Negro et al. (2004). In particular, the TVP-VAR is used to summarize the statistical properties of both the observed data and the theory-simulated data and an estimate of the deep parameters from the theory is obtained by matching as close as possible TVP-VAR parameters from observed data and from the simulated data.[19] To sample from $p(\lambda, \boldsymbol{\theta}, \gamma | \boldsymbol{Y})$ I consider a two blocks random walk metropolis algorithm. This is basically a Gibbs Sampler where in the first block, I draw the hyper-parameters γ and λ conditionally on $\boldsymbol{\theta}$ while in the second block I draw $\boldsymbol{\theta}$ conditionally on the shrinking hyper-parameters γ and λ. Step 2, instead, consists just on Monte Carlo draws from the posterior of $(\boldsymbol{\Sigma_u}, \boldsymbol{\Phi})$ conditional on λ, γ and $\boldsymbol{\theta}$ which is the *Normal-Inverse-Wishart* distribution in (20) that is:

[18] For both γ and λ I consider the Uniform prior $\gamma \sim \mathcal{U}(0, b_\gamma)$ and $\lambda \sim \mathcal{U}(0, b_\lambda)$ where $b_\gamma = 10^{10}$ and $b_\lambda = 10^{10}$

[19] See Proposition 1 and Proposition 2 in Del Negro et al. (2004) for further details on this.

2.1) $p(\boldsymbol{\Sigma_u}|\boldsymbol{\Phi}, \lambda, \boldsymbol{\theta}, \gamma, \boldsymbol{Y}) \sim \mathcal{IW}\left(\tilde{\boldsymbol{S}}, \tilde{\nu}\right)$

2.2) $p(vec(\boldsymbol{\Phi})|\boldsymbol{\Sigma_u}, \lambda, \boldsymbol{\theta}, \gamma, \boldsymbol{Y}) \sim \mathcal{N}(\tilde{\boldsymbol{\Phi}}, \boldsymbol{\Sigma}_u \otimes \tilde{\boldsymbol{\Psi}})$

In this approach, the time varying parameters and the variance covariance matrix are drawn in a single step from their *Normal-Inverse-Wishart* conditional posterior distribution. It is worth to mention that in the case we allow for regressor specific shrinkage hyper-parameters λ_j for $j = 1, \ldots, k$ nothing changes conceptually, since the algorithm should just be adapted to draw all the λ_j's hyper-parameters together with γ in the second block of the random walk metropolis in Step 1. [20] As a last note, to let the estimated path of time varying coefficients $\boldsymbol{\Phi_1}, \ldots, \boldsymbol{\Phi_T}$ and the estimate of the hyper-parameter λ less affected by the choice of the initial condition $\boldsymbol{\Phi_0}$ I assume that in equation (4) $\boldsymbol{\eta_1} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma_u} \otimes 5(\sum_{t=1}^{T_{pre}} \boldsymbol{x}_t \boldsymbol{x}_t')^{-1})$ such that $\boldsymbol{\Phi_1} \sim \mathcal{N}(\hat{\boldsymbol{\Phi}}_{\boldsymbol{0}}, \boldsymbol{\Sigma_u} \otimes 5(\sum_{t=1}^{T_{pre}} \boldsymbol{x}_t \boldsymbol{x}_t')^{-1})$ where T_{pre} is the size of a pre-sample of observations while $\hat{\boldsymbol{\Phi}}_{\boldsymbol{0}} = 0$.

3 Forecasting with the TC-TVP-VAR

In this section I consider the problem of forecasting the rate of growth of GDP, the inflation rate and the Fed Fund rate using a trivariate TVP-VAR model. In the specific, I estimate a trivariate TVP-VAR for the US economy using data from 1970 up to 2019 and I compare the forecast accuracy of a standard TVP-VAR model to the forecasts from a TC-TVP-VAR.

3.1 Small scale New Keynesian model for the US economy

In the TC-TVP-VAR I exploit the New Keynesian model in Del Negro et al. (2004) to parametrize the theory coherent prior. The conceptual framework commonly denoted

[20] Formulas for the marginal likelihood and the conditional posterior distribution of $\boldsymbol{\Phi}$ and $\boldsymbol{\Sigma}_u$ can be found in the appendix A.1.5.

as the 3-equation New Keynesian model constitutes the nucleus of Michael Woodford's book "Interest and Prices" (Woodford, 2003) and underpins most of modern monetary macroeconomics models.[21] More specifically, the structural model is composed by an IS curve (30), a New Keynesian Phillips curve (31), a monetary policy rule (32) and two equations that describe the dynamics for the log-deviation from the steady state of technological process (33) and government spending (34), namely

$$\hat{y}_t = E_t[\hat{y}_{t+1}] - \frac{1}{\tau}\left(\hat{R}_t - E_t[\hat{\pi}_{t+1}]\right) + (1-\rho_g)\hat{g}_t + \frac{\rho_z}{\tau}\hat{z}_t \tag{30}$$

$$\hat{\pi}_t = \frac{\tilde{\gamma}}{r^*}E_t[\hat{\pi}_{t+1}] + \kappa(\hat{y}_t - \hat{g}_t) \tag{31}$$

$$\hat{R}_t = \rho_R\hat{R}_{t-1} + (1-\rho_R)(\psi_1\hat{\pi}_t + \psi_2\hat{y}_t) + \varepsilon_{R,t} \qquad \varepsilon_{R,t} \sim \mathcal{N}(0,\sigma_R) \tag{32}$$

$$\hat{z}_t = \rho_z\hat{z}_{t-1} + \varepsilon_{z,t} \qquad \varepsilon_{z,t} \sim \mathcal{N}(0,\sigma_z) \tag{33}$$

$$\hat{g}_t = \rho_g\hat{g}_{t-1} + \varepsilon_g \qquad \varepsilon_{g,t} \sim \mathcal{N}(0,\sigma_g) \tag{34}$$

The population moments needed to parametrize the prior are derived from the state-space representation of the New Keynesian model obtained by solving the system of non-linear rational expectation equations. In particular, the non-linear rational expectation equations are solved using the method based on matrix eigenvalue decomposition by Sims (2002) leading to a solution which has the form

$$\boldsymbol{s_t} = \mathcal{T}(\boldsymbol{\theta})\boldsymbol{s_{t-1}} + \mathcal{R}(\boldsymbol{\theta})\boldsymbol{\epsilon}_t \tag{35}$$

that is complemented with the set of observation equations

$$\boldsymbol{y}_t = \mathcal{D} + \mathcal{B}\boldsymbol{s_t} \tag{36}$$

which look like:

$$YGR_t = ln(\tilde{\gamma}) + \Delta\hat{y}_t + \hat{z}_t \tag{37}$$

[21] See Del Negro et al. (2004) for the more in depth details on the New Keynesian Model.

$$INFL_t = ln(\pi^*) + \hat{\pi}_t \tag{38}$$

$$ln(INT_t) = 4ln(r^*) + 4ln(\pi^*) + 4\hat{R}_t \tag{39}$$

and relate the unobservable latent states in (35) to the observed time series of output growth, inflation rate and the Fed Fund rate. Equations (35) and (36) are used to derive the population moments needed to parametrize the prior. The population moments $\boldsymbol{\Gamma_{xx}}(\boldsymbol{\theta})$, $\boldsymbol{\Gamma_{xy}}(\boldsymbol{\theta})$, $\boldsymbol{\Gamma_{yy}}(\boldsymbol{\theta})$ are derived conditioning on $\boldsymbol{\theta}$, the vector of the deep parameters of the NK model

$$\boldsymbol{\theta} = [ln(\tilde{\gamma}), ln(\pi^*), ln(r^*), \kappa, \tau, \psi_1, \psi_2, \rho_R, \rho_g, \rho_z, \sigma_R^2, \sigma_g^2, \sigma_z^2] \tag{40}$$

3.2 Forecast comparison

Table 1 shows the comparison of the forecasts from a standard TVP-VAR model and the TC-TVP-VAR model. The forecasting exercise is designed such that I compute the recursive one quarter, two quarters, and one year ahead forecasts starting from 1985-Q1 up to 2019-Q4.[22] To compare relative point forecast accuracy, in the table I report the Root Mean Squared Error (RMSE) while for evaluating density forecast accuracy I report the average Cumulative Ranked Probability Scores (CRPS). In the table I also include the results concerning the forecasts from a constant parameters VAR with flat prior and a constant parameters Bayesian VAR (BVAR) with Minnesota type of prior.[23] Overall, the TC-TVP-VAR provides the most accurate point and density forecasts for both output growth and inflation rate, outperforming the standard TVP-VAR model at all the horizons considered. For forecasting output growth, the standard TVP-VAR performs poorly relative to the TC-TVP-VAR, but also to the constant parameters BVAR with Minnesota prior, suggesting that the model tends to fit some noise in the time series of output growth. In line with the previous forecasting literature, allowing for time variation of the parameters of

[22] Details on the data can be found in the appendix A.2.1

[23] Appendix A.2.2 the reports details on the competing forecasting models.

Table 1: Point and density forecast accuracy 1985Q1 - 2019Q4

	GDP growth		Inflation rate		Interest Rate	
	RMSE	**CRPS**	**RMSE**	**CRPS**	**RMSE**	**CRPS**
a) *One quarter ahead*						
VAR flat prior	0.4417	0.3703	0.3451	0.2877	**0.0402**	0.0954
B-VAR min	0.3422	0.3584	0.2854	0.2637	0.0403	**0.0936**
Standard TVP-VAR	0.4283	0.3853	0.3362	0.2682	0.0451	0.1096
TC-TVP-VAR	**0.3027**	**0.3485**	**0.2715**	**0.2461**	0.0608	0.1312
b) *Two quarters ahead*						
VAR flat prior	0.5045	0.3885	0.4771	0.3423	0.1384	0.1784
B-VAR min	0.3990	0.3909	0.3819	0.3110	**0.1303**	**0.1709**
Standard TVP-VAR	0.5131	0.4151	0.3362	0.2911	0.1548	0.2037
TC-TVP-VAR	**0.3340**	**0.3668**	**0.3183**	**0.2722**	0.1689	0.2096
c) *One year ahead*						
VAR flat prior	0.4278	**0.3728**	0.4965	0.3675	0.4475	0.3385
B-VAR min	0.4037	0.4065	0.4099	0.3393	**0.4142**	**0.3229**
Standard TVP-VAR	0.6472	0.4775	0.2855	0.3039	0.5665	0.3854
TC-TVP-VAR	**0.3617**	0.3887	**0.2758**	**0.2802**	0.5156	0.3593

Notes: The Table reports the Root Mean Squared Error (RMSE) and the average Cumulative Ranked Probability Scores (CRPS). In bold the best model according to each forecast metric.

the VAR is important for obtaining accurate forecasts of the inflation rate, as the standard TVP-VAR outperforms both the VAR with flat prior and the BVAR with the Minnesota prior. However, economic shrinkage is helpful to obtain more reliable point and density forecasts when modelling the time variation of the coefficients. Indeed the TC-TVP-VAR outperforms the standard TVP-VAR at all the horizons. As a caveat, the standard Minnesota prior centering the autoregressive coefficients on a random walk process outperforms the other competitors, including the TC-TVP-VAR, for forecasting the Fed Fund Rate. This result is consistent with the results of the forecasting exercise in Del Negro et al. (2004) which use the same small scale NK model to parametrize a prior for a constant parameters VAR.

4 Response analysis at the ZLB with the TC-TVP-VAR

TVP-VAR are extensively used in applied research not only to make forecasts, but also to infer changes in the response of the economy to macroeconomic shocks. In this section, I show that the proposed shrinkage prior can be useful also to enhance inference on the impulse response functions estimated from a TVP-VAR. Recent studies in empirical macroeconomics have used TVP-VARs to assess whether the US economy's performance was affected by a binding ZLB constraint (Debortoli, Galí, and Gambetti, 2019; Benati et al., 2023). As a matter of fact, according to a standard New-Keynesian model, the economy is expected to exhibit different responses when the ZLB constraint is in effect. For example, the model predicts a distinct response of output and inflation following both demand and supply shocks when the conventional stabilizing monetary policy response to aggregate shocks is constrained as a consequence of the Federal Funds Rate hitting the ZLB. Figure 4 makes this point, by showing the responses to a pure demand shock - the risk premium shock [24] - in the Smets et al. (2007) model version considered in Del Negro et al. (2015).[25] The figure plots the cumulative response of output growth and the response of the inflation rate to an exogenous risk premium shock that reduces households' required return of assets, decreasing firms' cost of capital. As expected, when monetary policy is constrained by the policy rate hitting the zero lower bound, the response of both output and inflation to the risk premium shock is magnified on impact with the effects of the shock taking much more quarters to be reabsorbed by the economy.

[24] The risk premium shock in Smets and Wouters (2007) is an exogenous term which affects the intertemporal margins entering both the consumption and investment Euler equation. In contrast to a discount factor shock, the risk premium shock helps explaining the co-movement of consumption and investment. More details on the risk premium shocks are presented in the Appendix A.3.2.

[25] In the model, the ZLB period is treated as in Del Negro et al. (2015), and more details on the model will be explained in the next subsection.

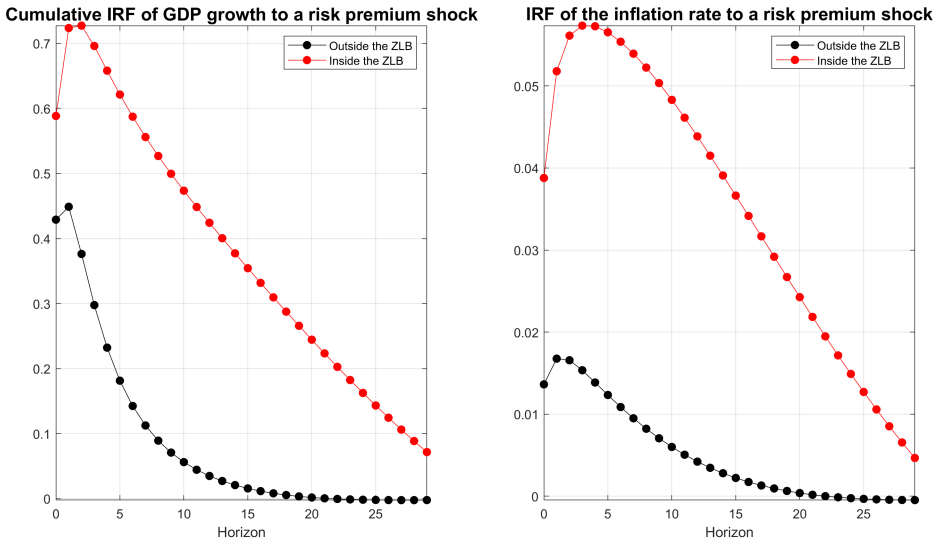


Figure 4: The two plots show the cumulative response of output growth (on the left) and the response of the inflation rate (on the right) to a risk premium shock inside and outside the ZLB period in the New Keynesian model.

Despite the sharp difference in the propagation of the shocks in the economy inside and outside the zero lower bound period predicted by a standard NK model, empirical evidences investigating this issue are mixed and many studies do not find substantial evidence supporting a different response of the US economy during the zero lower bound period. For example, Debortoli et al. (2019), support the *irrelevance hypothesis* i.e. the hypothesis that the economy's performance has not been affected by a binding ZLB constraint, embracing the view that unconventional monetary policy have been effective at getting around the zero lower bound (ZLB) constraint. Benati et al. (2023) recently showed that, given the short length of the zero-lower-bound period, inference based on a standard TVP-VAR where the time varying parameters are regarded as slow moving stochastic processes, doesn't allow to capture the changing relationship among the macroeconomic variables during the ZLB period predicted by the New Keynesian model. In what follows, based on a simulation study, I also find that a standard TVP-VAR struggles to detect the change in the responses of the economy in the ZLB period generated by a NK model. I show that the TC-TVP-VAR can in

principle be used to solve this inferential issue and therefore to recover the distinct response of the economy during the zero lower bound period.

4.1 Medium scale NK model accounting for ZLB and forward guidance

The New Keynesian model is the version of the Smets et al. (2007) model considered in Del Negro et al. (2015). The model features price and wage stickiness, investment adjustment costs, habit formation in consumption and 7 shocks being monetary policy shocks, technology shocks, price mark-up shocks, wage mark-up shocks, risk premium shocks, fiscal policy shock and shocks to the marginal efficiency of capital. In the model the monetary policy rule accounts for ZLB period and forward guidance.[26] More specifically, when accounting for the zero lower bound and forward guidance the solution of the model implies a state state space representation which exhibits time varying coefficients. The solution method follows the approach developed by Cagliarini et al. (2013) for linear stochastic rational expectations models in the face of a finite sequence of anticipated structural changes. In particular, it is assumed that at a given period t, agents expect the nominal interest rate to be at the ZLB for $\bar{H}$ periods. That is, the monetary policy rule

$$R_t = \rho_R R_{t-1} + (1-\rho_R)\left(\psi_1(\pi_t - \pi_t^*) + \psi_2(y_t - y_t^f)\right) + \psi_3\left((y_t - y_t^f) - (y_{t-1} - y_{t-1}^f)\right) + r_t^m \tag{41}$$

becomes

$$R_\tau = -R_* \tag{42}$$

[26] The log-linearized equilibrium conditions of the model can be found in the appendix A.3 The model is labelled "SW" in Del Negro et al. (2015) and assumes a constant inflation target.

for $\tau = t, \ldots, t + \bar{H}$ while it is determined by (41) for $\tau > t + \bar{H}$. This implies that the equilibrium conditions

$$\boldsymbol{\Gamma}_{2,\tau} E_\tau[\boldsymbol{s}_{\tau+1}] + \boldsymbol{\Gamma}_{0,\tau} \boldsymbol{s}_\tau = \boldsymbol{\Gamma}_{c,\tau} + \boldsymbol{\Gamma}_{1,\tau} \boldsymbol{s}_{\tau-1} + \boldsymbol{\Psi}_{\boldsymbol{\tau}} \boldsymbol{\varepsilon}_\tau \tag{43}$$

differ over time depending on whether $\tau \leq t + \bar{H}$, with the row corresponding to the policy rule differing across τ. Following Cagliarini et al. (2013) the solution for the linear rational expectations model under anticipated structural variations takes the form

$$\boldsymbol{s}_{\boldsymbol{\tau}} = \mathcal{C}_\tau^{(\tau,\bar{H})} + \mathcal{T}_\tau^{(\tau,\bar{H})} \boldsymbol{s}_{\boldsymbol{\tau-1}} + \mathcal{R}_\tau^{(\tau,\bar{H})} \boldsymbol{\epsilon}_{\boldsymbol{\tau}} \tag{44}$$

where $\mathcal{C}_\tau^{(\tau,\bar{H})}$ $\mathcal{T}_\tau^{(\tau,\bar{H})}$ and $\mathcal{R}_\tau^{(\tau,\bar{H})}$ are computed by recursion

$$\mathcal{C}_\tau^{(\tau,\bar{H})} = (\boldsymbol{\Gamma}_{2,\tau} \mathcal{T}_{\tau+1}^{(\tau,\bar{H})} + \boldsymbol{\Gamma}_{0,\tau})^{-1} (\boldsymbol{\Gamma}_{c,\tau} - \boldsymbol{\Gamma}_{2,\tau} \mathcal{C}_{\tau+1}^{(\tau,\bar{H})}) \tag{45}$$

$$\mathcal{T}_\tau^{(\tau,\bar{H})} = (\boldsymbol{\Gamma}_{2,\tau} \mathcal{T}_{\tau+1}^{(\tau,\bar{H})} + \boldsymbol{\Gamma}_{0,\tau})^{-1} \boldsymbol{\Gamma}_{1,\tau} \tag{46}$$

$$\mathcal{R}_\tau^{(\tau,\bar{H})} = (\boldsymbol{\Gamma}_{2,\tau} \mathcal{T}_{\tau+1}^{(\tau,\bar{H})} + \boldsymbol{\Gamma}_{0,\tau})^{-1} \boldsymbol{\Psi}_{\boldsymbol{\tau}} \tag{47}$$

starting from $\mathcal{T}_{t+\bar{H}+1}^{(\tau,\bar{H})} = \mathcal{T}$, $\mathcal{C}_{t+1+\bar{H}}^{(\tau,\bar{H})} = 0$ and where the superscript $(t, \bar{H})$ is used to indicate that the solution is obtained under the assumption that the announcement of zero interest rates for a duration of $\bar{H}$ periods was made in period t. Following Del Negro et al. (2015) to measure the number of quarters $\bar{H}$ that the Federal Funds Rate is expected to remain at the ZLB I exploit information based on the overnight index swap (OIS) rates. In particular I identify the ZLB period as the quarters in which the OIS rate is lower then 0.35. This classification leads to the same ZLB period considered in Benati et al. (2023) and Debortoli et al. (2019) namely 2009Q1-2015Q3 (28 quarters). Following Chen, Cúrdia, and Ferrero (2012) I assume that the number of quarters such that the policy rate is expected to stay fixed, is at most equal to four. According to the model's solution, the matrices $\mathcal{C}_t^{(\tau,\bar{H})}, \mathcal{T}_t^{(\tau,\bar{H})}, \mathcal{R}_t^{(\tau,\bar{H})}$ characterize the

transition equation of a time varying coefficients state space model

$$\boldsymbol{y}_t = \mathcal{D} + \mathcal{B}\boldsymbol{s_t} + \boldsymbol{v}_t \tag{48}$$

$$\boldsymbol{s}_t = \mathcal{C}_t + \mathcal{T}_t\boldsymbol{s_{t-1}} + \mathcal{R}_t\boldsymbol{\epsilon}_t \tag{49}$$

where the constant matrices $\mathcal{D}$, $\mathcal{B}$ and time varying matrices $\mathcal{C}_t$, $\mathcal{T}_t$, $\mathcal{R}_t$ depend on the structural parameters $\boldsymbol{\theta}$ while $\boldsymbol{v}_t$ is a measurement error. [27] Therefore, while the space representation of the model's solution features time varying coefficients, the deep structural parameters of the NK model $\boldsymbol{\theta}$ are constant. Defining $\underline{T}^{zlb}$ the first period inside the ZLB and $\bar{T}^{zlb}$ the last period inside the ZLB, for $t < \underline{T}^{zlb}$ and $t > \bar{T}^{zlb}$ we have $\mathcal{R}_t = \bar{\mathcal{R}}$ and $\mathcal{T}_t = \bar{\mathcal{T}}$ and $\mathcal{C}_t = \mathbf{0}$. Therefore the state space model becomes

$$\boldsymbol{y}_t = \mathcal{D} + \mathcal{B}\boldsymbol{s}_t \tag{50}$$

$$\boldsymbol{s}_t = \bar{\mathcal{T}}\boldsymbol{s}_{t-1} + \bar{\mathcal{R}}\boldsymbol{\varepsilon}_t \tag{51}$$

Hence, for $t = 1, \ldots, \underline{T}^{zlb} - 1$ and $t = \bar{T}^{zlb} + 1, \ldots, T$ the moments $\boldsymbol{\Gamma_{xx,t}} = \mathbb{E}[\boldsymbol{x_t x_t'}|\boldsymbol{\theta}]$, $\boldsymbol{\Gamma_{xy,t}} = \mathbb{E}[\boldsymbol{x_t y_t'}|\boldsymbol{\theta}]$ and $\boldsymbol{\Gamma_{yy,t}} \equiv \mathbb{E}[\boldsymbol{y_t y_t'}|\boldsymbol{\theta}]$ needed to parameterize the prior are computed assuming $\mathbb{E}[\boldsymbol{s}_t\boldsymbol{s}_t'] = \mathbb{E}[\boldsymbol{s}_{t-1}\boldsymbol{s}_{t-1}']$ and solving the Lyapunov equation

$$\mathbb{E}[\boldsymbol{s_t s_t'}] = \bar{\mathcal{T}}\,\mathbb{E}[\boldsymbol{s_{t-1} s_{t-1}'}]\bar{\mathcal{T}}' + \bar{\mathcal{R}}\boldsymbol{\Omega}\bar{\mathcal{R}}' \tag{52}$$

As for the periods inside the ZLB, namely for $\underline{T}^{zlb} \leq t \leq \bar{T}^{zlb}$, $\mathbb{E}[\boldsymbol{s}_t\boldsymbol{s}_t']$ and the implied population moments are computed recursevely according to the low of motion implied by (48) and (49).

[27] Details for the observation equations are in the appendix A.3 together with the list of the deep structural parameters $\boldsymbol{\theta}$ (appendix A.3.1).

4.2 Simulation study

As anticipated above and shown in figure 4, the NK model predicts a distinct response of the economy to a risk premium shock inside and outside the ZLB period. Conditioning on the vector of deep parameters of the NK model, I simulate data from the state space representation (48) and (49).[28] In the simulation I consider artificial samples with $T = 139$ mimicking quarterly observations for the period 1985Q1-2019Q3. The length of the ZLB period is 28 quarters, covering the period 2009Q1-2015Q3. I estimate a standard TVP-VAR model (Chan et al., 2009) on the simulated data to understand whether the model is able to recover the change in the response of the economy during the ZLB period generated by the NK model.[29] Figure 5 shows the estimated responses of output and inflation to a risk premium shock obtained from a standard TVP-VAR. The figure plots the responses of output growth (cumulative) and the inflation rate to a one standard deviation risk premium shock in two reference dates, one outside and the other one inside the ZLB period. In order to identify the shocks in the structural TVP-VAR I exploit the true impact matrix of the NK model given by

$$\frac{\partial \boldsymbol{y_t}}{\partial \boldsymbol{\epsilon}_t'} = \underbrace{\mathcal{B}(\boldsymbol{\theta})\boldsymbol{\mathcal{R}_t}(\boldsymbol{\theta})\boldsymbol{\Omega}(\boldsymbol{\theta})}_{\boldsymbol{A_{0,t}}(\boldsymbol{\theta})^{NK}} \tag{53}$$

where $\boldsymbol{\Omega}(\boldsymbol{\theta}) = diag(\sigma_g^2, \sigma_b^2, \sigma_\mu^2, \sigma_z^2, \sigma_{\lambda_f}^2, \sigma_{\lambda_w}^2, \sigma_r^2)$ is the diagonal matrix containing the volatility of the structural shocks. As figure 5 shows, despite conditioning on the correct identification scheme to identify the structural shocks, inference based on a standard TVP-VAR struggles to provide convincing evidences supporting a distinct response of both output growth and the inflation rate to a risk premium shock inside and outside the ZLB period. As a matter of fact, the estimates of the impulse responses are so imprecise that the model doesn't allow to detect any change in

[28] Parameters are calibrated according to the posterior mode in Del Negro et al. (2015)

[29] The standard TVP-VAR model is the model in Chan et al. (2009). To estimate the model I use the MATLAB codes kindly made available by Joshua Chan on his personal website.

the response of the economy inside and outside the ZLB. [30]

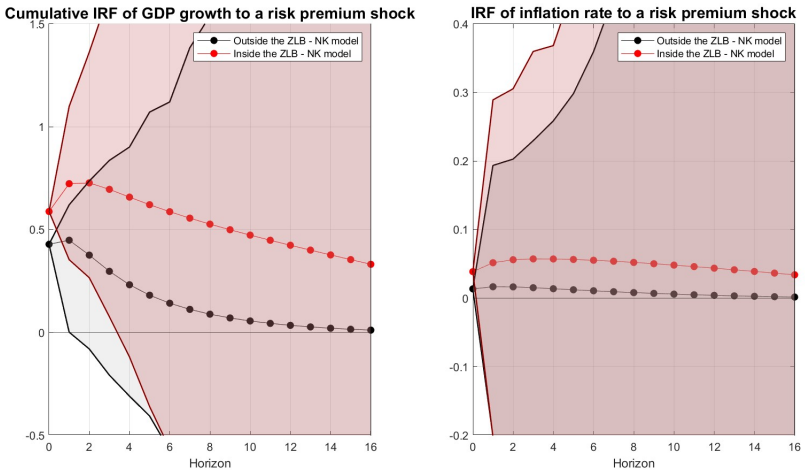


Figure 5: Estimated $90^{th} - 10^{th}$ credible sets of the cumulative response of output growth (on the left) and of the response of the inflation rate (on the right) to a risk premium shock inside (shaded red) and outside (shaded grey) the ZLB period by a standard TVP-VAR on data simulated from the NK model.

Figure 6, instead, shows the responses obtained by estimating a TC-TVP-VAR that exploits the NK model as a prior for the time varying coefficients. The moment matrices $\boldsymbol{\Gamma_{xx}}$, $\boldsymbol{\Gamma_{xy}}$, $\boldsymbol{\Gamma_{yy}}$ used to parametrize the *Normal-Inverse-Wishart* prior are derived from the equations of the state representation (48) and (49). Clearly, due to the time variation of the coefficients in the state space representation, the moments of the structural model are time varying with $\mathbb{E}[\boldsymbol{x_t x_t'}|\boldsymbol{\theta}]$, $\mathbb{E}[\boldsymbol{x_t y_t'}|\boldsymbol{\theta}]$ and $\mathbb{E}[\boldsymbol{y_t y_t'}|\boldsymbol{\theta}]$ changing over time. This, in turns, implies time varying restriction functions for the coefficients of the TVP-VAR encoded in $\boldsymbol{\Phi(\theta)^* = \Gamma_{xx}(\theta)^{-1}\Gamma_{xy}(\theta)}$. Since the time varying restrictions incorporate the change in the response of the economy foreseen by the NK model, the responses of output growth and the inflation rate to the risk premium shock are more precisely estimated as the time varying coefficients are shrinked towards these restrictions. This is shown in figure 6, which plots the responses of output growth and of the inflation rate to a risk premium shock for different fixed values

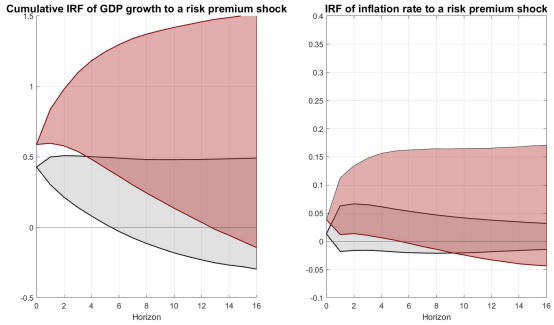
[30] As shown in the appendix in figure 8, this result is not driven by this specific simulated sample of artificial observations.

of γ, that is for a different amount of shrinkage towards the restrictions implied by the NK model.[31] As the figure shows, a sufficiently high γ is needed to detect a precise and distinct response of the economy inside and outside the ZLB period as predicted by the NK model. As γ increases and gets big enough, we get more and more precise estimates of the time varying coefficients. This allows to detect the different propagation of the shocks inside and outside the ZLB period, as predicted by the New Keynesian model. Letting $\gamma \to \infty$ the estimated model is a restricted TVP-VAR in which the time varying coefficients exactly satisfy $\boldsymbol{\Phi(\theta)^* = \Gamma_{xx}(\theta)^{-1}\Gamma_{xy}(\theta)}$. As expected, the estimated responses from this model almost exactly resemble the responses from the NK model (panel (d)).

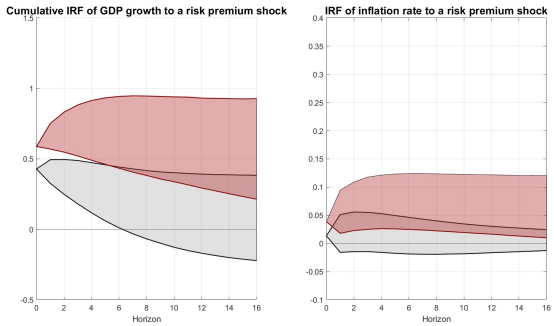
4.3 Estimating the TC-TVP-VAR on US data

When analyzing real data, it is reasonable to think of the NK model just as an approximate (most likely misspecified) tightly parameterized representation of the true data generating process. In other words, we do not expect the restrictions implied by the NK model to hold exactly. However, when encoded into a prior, these restrictions might prove to be useful to get more precise estimates of the time varying coefficients and of nonlinear functions of these coefficients, such as the impulse response functions. In particular, in our context, shrinkage might turn out to be particularly useful to detect the change in the response of the economy during the ZLB period predicted by the NK model. In order to analyze to what extent the data support the changing behavior during the ZLB, my approach consists in exploiting the New Keynesian model as a prior for the TC-TVP-VAR. The extent to which the predictions from the New Keynesian model will be supported by the data will depend on the estimated posterior distribution of γ, the hyper-parameter governing the degree of shrinkage of the parameters towards the restriction implied by the New Keynesian model. If the

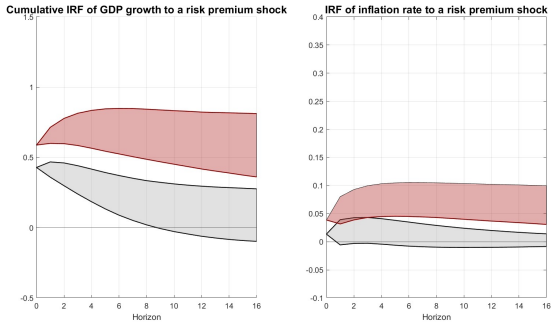
[31] Also in this case, in the TC-TVP-VAR we condition on the correct identification scheme for the structural shocks.



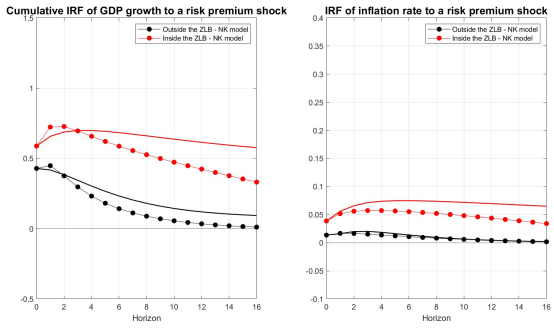
(a) $\gamma = 10$



(b) $\gamma = 100$



(c) $\gamma = 300$



(d) $\gamma = \infty$

Figure 6: Estimated $10^{th} - 90^{th}$ credible sets of the cumulative response of output growth (left) and the response of the inflation rate (right) to a risk premium shock inside (shaded red) and outside (shaded grey) the ZLB period for different values of γ. Data are simulated from the NK model. Horizons are in quarters. a) $\gamma = 10$ b) $\gamma = 100$ c) $\gamma = 300$ d) $\gamma = \infty$. For $\gamma = \infty$ the figure also reports the true responses from the NK model.

posterior distribution of γ is concentrated relatively far from zero, the coefficients of the TVP-VAR will be shrinked towards the restrictions implied by the NK model and the estimated responses from the TC-TVP-VAR will resemble the predictions of the NK model. This would happen in practice if the restrictions from the NK model find enough support on the data. Conversely, if the restrictions from the NK are deemed implausible by the data, the posterior distribution of γ will be concentrated around zero and the estimates of the time varying coefficients will not reflect the restrictions coming from the prior. The estimated model is a 7-variable TC-TVP-VAR for the US economy including output growth, consumption growth, investment growth, real wage growth, hours worked, inflation and the Fed Fund rate and it is estimated over the sample 1985Q1-2019Q3.[32] In order to identify the shocks in the structural TVP-VAR, I exploit the impact matrix of the NK model. The deep parameters of the NK model are treated as unknown and estimated along with the other parameters of the model, which implies that on impact there is uncertainty on the effect of the shocks on the variables of the system. Figure 7 shows the estimated responses from the TC-TVP-VAR where the hyper-parameter γ is estimated along with the other parameters of the model. The posterior distribution of γ is estimated to be concentrated far from zero, meaning that the time varying restriction functions from the NK model find sufficient support on the data. This, in turns, is reflected on the estimates of the $10^{th} - 90^{th}$ credible sets which provide some evidences supporting a distinct responses of the economy inside and outside the ZLB period. As predicted by the NK model, the estimates suggest that when monetary policy is constrained by the nominal rate hitting the ZLB, the risk premium shock is reabsorbed at a much slower pace by the economy. Following the positive demand shock both output growth and the inflation rate increase, with the effect on output growth being more precisely estimated at shorter horizons. The effect on output peaks after almost two quarters outside the ZLB period, while about after five quarters inside the ZLB period. As for inflation,

[32]Details on the variables and their transformation are available in the appendix A.3.

the effect peaks after one quarter outside the ZLB period and after two quarters inside the ZLB period. After the peak, the effect of the risk premium shock on both output and inflation dies at a much slower pace inside the ZLB period, as foreseen by the NK model. Clearly, given the symmetric nature of the impulse response functions in our econometric model, figure 7 implies that in the face of an increase in the risk premium, the economy would experience more persistent decreases of both inflation and output inside the ZLB period. Hence, risk premium shocks become more important when the ZLB binds as found in Gourio and Ngo (2020). More in general, this finding is compatible with the findings in Aruoba, Mlikota, Schorfheide, and Villalvazo (2022), which building on Mavroeidis (2021) develop a structural VAR in which an occasionally-binding constraint generates censoring of one of the dependent variables. They find that the presence of the ZLB is empirically relevant for the propagation of macroeconomic shocks. Differently from both Mavroeidis (2021) and Aruoba et al. (2022), but similarly to Debortoli et al. (2019) and Benati et al. (2023) our approach assumes the ZLB period to be completely observable. As well, while both Mavroeidis (2021) and Aruoba et al. (2022) leverage censoring to find identifying information on the propagation of macroeconomic shocks we directly resort to economic theory encoded in the NK model to identify the macroeconomic shocks. One key difference is that while in Aruoba et al. (2022) the structural coefficients switch across unobserved regimes, in our TVP-SVAR the autoregressive time varying parameters are allowed to slowly drift inside, outside and in the transition to and from the ZLB period. [33]

[33] As for the impact matrix $A_{0,t}(\boldsymbol{\theta})$, in our framework deterministically changes over time (conditioning on the NK parameters $\boldsymbol{\theta}$) as a function of $\bar{H}_t$, the number of periods agents think the ZLB remains binding.

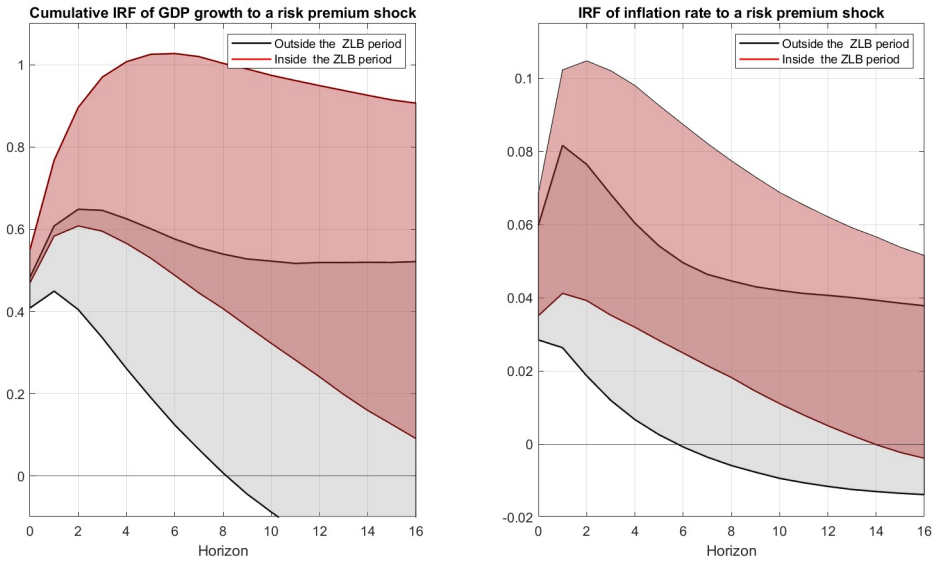


Figure 7: Estimated $90^{th} - 10^{th}$ credible sets of the cumulative response of output growth (on the left) and of the response of the inflation rate (on the right) to a risk premium shock inside (shaded red) and outside (shaded grey) the ZLB period from the TC-TVP-VAR on US data.

5 Conclusion

TVP-VAR are flexible statistical models used both for prediction and policy analysis in macroeconomics. Despite their flexibility allows to capture changes in the dynamic relationship among the macroeconomic variables, these models can easily become too flexible with the risk of over-fitting the data. This translate into poor forecasting performances and imprecise inference on the time-varying parameters and of typical objects of interests such as the impulse responses. On the other side, models from the economic theory typically provide a more tightly parameterized representation of the macroeconomy and therefore have the opposite tendency of fitting the data rather poorly. This paper exploits the restrictions implied by economic theory to formulate a prior for the parameters of TVP-VARs so as to enhance inference within this class of models. More specifically, the paper introduces a shrinkage prior that centers the

time varying coefficients at each time period on the cross equation restrictions implied by an underlying economic theory about the variables in the system. The paper shows that "economic shrinkage" can be successfully used to obtain more accurate forecasts and more precise estimates of the impulse response functions. Furthermore, exploiting the restrictions of the classical 3-equation New Keynesian block to form a prior for the TVP-VAR through the proposed framework improves both point point and density forecast accuracy of both output growth and inflation at all the horizons considered. Furthermore, the paper exploits a medium scale New Keynesian model that accounts for forward guidance and the Zero Lower bound period, to formulate a prior for a medium scale TVP-VAR. Using this theory coherent shrinkage prior allows to estimate more precisely the response of the economy to macroeconomic shocks inside and outside the ZLB period, helping to solve the inferential problems faced by the standard TVP-VAR. On US data, the paper finds indeed a distinct effect of a pure demand shock - the risk premium shock - inside and outside the ZLB period, with the effects of the shock reabsorbing at a much slower pace inside the ZLB period. This finding has important implications for the conduct of both fiscal and macroprudential policy at the ZLB.

Future research For future research the prior proposed in this paper could be modified to accommodate for multiple competing theories à la Loria et al. (2022). As well, the model could be extended to embed stochastic volatility, but this extension is non-trivial since in TVP-VARs a la Primiceri (2005) stochastic volatility breaks the Kronecker structure of the likelihood. More in general, restrictions from economic theory could successfully be used to reduce overfitting and sharpen inference in non-parametric VARs and Gaussian Procesess VARs (Hauzenberger, Huber, Marcellino, and Petz, 2022).

A Appendix

A.1 Theory Coherent TVP-VAR

A.1.1 Time Varying Parameters by dummy observations

Starting from:

$$\underbrace{\boldsymbol{y}_t'}_{1\times N} = \underbrace{\boldsymbol{x}_t'}_{1\times k}\underbrace{\boldsymbol{\Phi}_t}_{k\times N} + \underbrace{\boldsymbol{u_t}}_{1\times N} \qquad\qquad \boldsymbol{u_t} \sim \mathcal{N}(\mathbf{0}_{1\times N}, \boldsymbol{\Sigma_u})$$

we can write the TVP-VAR in static compact form as:

$$\underbrace{\boldsymbol{Y}}_{T\times N} = \underbrace{\boldsymbol{X}}_{T\times Tk}\underbrace{\boldsymbol{\Phi}}_{Tk\times N} + \underbrace{\boldsymbol{U}}_{T\times N} \qquad\qquad \boldsymbol{U} \sim MVN(\mathbf{0}, \boldsymbol{I_T}, \boldsymbol{\Sigma_u})$$

Suppose we want to specify independent RW stochastic processes for all the coefficients in $\boldsymbol{\Phi}$ as:

$$\phi_{jt}^{(i)} = \phi_{jt-1}^{(i)} + \eta_{jt}^{(i)} \qquad\qquad \eta_j^{(i)} \sim \mathcal{N}\left(0, \frac{\Sigma_{ii}}{\lambda_j^2}\right) \tag{54}$$

for $t = 1, \ldots, T$ $i = 1, \ldots, N$ and $j = 1, \ldots, k$. Defining $\boldsymbol{\Lambda_k} = diag(\lambda_1, \ldots, \lambda_k)$ we can use a set of dummy observations to model the time variation of the coefficients. The dummy observations imply the linear *fuzzy* restrictions in (54) on the coefficients in $\boldsymbol{\Phi}$, in the specific we can write:

$$\underbrace{\begin{bmatrix} \boldsymbol{\Lambda}_k\boldsymbol{\Phi_0} \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}}_{\boldsymbol{Y^*}} = \underbrace{\begin{bmatrix} \boldsymbol{\Lambda_k} & \mathbf{0} & \ldots & \mathbf{0} \\ 0 & \boldsymbol{\Lambda_k} & \ldots & \mathbf{0} \\ \mathbf{0} & 0 & \ldots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & 0 & \boldsymbol{\Lambda_k} \end{bmatrix}\begin{bmatrix} \boldsymbol{I}_k & \mathbf{0} & \ldots & \mathbf{0} \\ -\boldsymbol{I}_k & \boldsymbol{I}_k & \ldots & \mathbf{0} \\ \mathbf{0} & -\boldsymbol{I}_k & \ldots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & -\boldsymbol{I}_k & \boldsymbol{I}_k \end{bmatrix}}_{\boldsymbol{X^*}}\begin{bmatrix} \boldsymbol{\Phi_1} \\ \boldsymbol{\Phi_2} \\ \boldsymbol{\Phi_3} \\ \vdots \\ \boldsymbol{\Phi_T} \end{bmatrix} + \underbrace{\begin{bmatrix} \boldsymbol{u}_1^* \\ \boldsymbol{u}_2^* \\ \boldsymbol{u}_3^* \\ \vdots \\ \boldsymbol{u}_T^* \end{bmatrix}}_{\boldsymbol{U^*}} \tag{55}$$

$$\boldsymbol{Y}^* = \boldsymbol{X}^*\boldsymbol{\Phi} + \boldsymbol{U}^* \tag{56}$$

This is just another way of writing:

$$\underbrace{\begin{bmatrix} \boldsymbol{I}_k & \boldsymbol{0} & \dots & \boldsymbol{0} \\ -\boldsymbol{I}_k & \boldsymbol{I}_k & \dots & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{I}_k & \dots & \boldsymbol{0} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{I}_k & \boldsymbol{I}_k \end{bmatrix}}_{\boldsymbol{H_{Tk}}} \underbrace{\begin{bmatrix} \boldsymbol{\Phi_1} \\ \boldsymbol{\Phi_2} \\ \boldsymbol{\Phi_3} \\ \vdots \\ \boldsymbol{\Phi_T} \end{bmatrix}}_{\boldsymbol{\Phi}} = \underbrace{\begin{bmatrix} \boldsymbol{\Phi_0} \\ \boldsymbol{0} \\ \boldsymbol{0} \\ \vdots \\ \boldsymbol{0} \end{bmatrix}}_{\boldsymbol{\Phi_{00}}} - \underbrace{\begin{bmatrix} \boldsymbol{\Lambda_k^{-1}} & \boldsymbol{0} & \dots & \boldsymbol{0} \\ 0 & \boldsymbol{\Lambda_k^{-1}} & \dots & \boldsymbol{0} \\ \boldsymbol{0} & 0 & \dots & \boldsymbol{0} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & 0 & \boldsymbol{\Lambda_k^{-1}} \end{bmatrix}}_{\boldsymbol{\Lambda^{-1}}} \begin{bmatrix} \boldsymbol{u}_1^* \\ \boldsymbol{u}_2^* \\ \boldsymbol{u}_3^* \\ \vdots \\ \boldsymbol{u}_T^* \end{bmatrix} \tag{57}$$

Since the generic t^{th} block of dimension $k \times N$ reads as follows :

$$\boldsymbol{\Phi_t} = \boldsymbol{\Phi}_{t-1} + \boldsymbol{\Lambda_k^{-1}}\boldsymbol{u_t^*} \qquad\qquad vec(\boldsymbol{u}_t) \sim \mathcal{N}(0, \boldsymbol{\Sigma_u} \otimes \boldsymbol{I}_k) \tag{58}$$

defining $\boldsymbol{\eta_t} = \boldsymbol{\Lambda_k^{-1}}\boldsymbol{u_t}$ we obtain:

$$vec(\boldsymbol{\eta_t}) \sim \mathcal{N}(0, \boldsymbol{\Sigma_u} \otimes (\boldsymbol{\Lambda_k'\Lambda_k})^{-1}) \tag{59}$$

that is equivalent to the dynamic linear model :

$$\underbrace{\boldsymbol{y}_t'}_{1 \times N} = \underbrace{\boldsymbol{x}_t'}_{1 \times k}\underbrace{\boldsymbol{\Phi}_t}_{k \times N} + \underbrace{\boldsymbol{u}_t'}_{1 \times N} \qquad\qquad \boldsymbol{u_t} \sim \mathcal{N}(\boldsymbol{0}_{N\times 1}, \boldsymbol{\Sigma_u}) \tag{60}$$

$$vec(\boldsymbol{\Phi_t}) = vec(\boldsymbol{\Phi_{t-1}}) + \boldsymbol{\eta_t} \qquad\qquad \boldsymbol{\eta_t} \sim \mathcal{N}(0, \boldsymbol{\Sigma} \otimes (\boldsymbol{\Lambda_k'\Lambda_k})^{-1}) \tag{61}$$

where $\boldsymbol{\Omega} = (\boldsymbol{\Lambda_k'\Lambda_k})^{-1}$

A.1.2 Details on the moment matrices

The likelihood of the observations simulated from the model of the economic theory $\boldsymbol{Y}(\boldsymbol{\theta})$ is:

$$p(\boldsymbol{Y}(\boldsymbol{\theta})|\boldsymbol{\Phi},\boldsymbol{\Sigma_u}) = (2\pi)^{-\frac{TN}{2}}|\boldsymbol{\Sigma_u}|^{-\frac{T}{2}} exp\left[-\frac{1}{2}tr\left((\boldsymbol{\Sigma_u}^{-1})(\boldsymbol{Y}(\boldsymbol{\theta})'\boldsymbol{Y}(\boldsymbol{\theta}) - \boldsymbol{\Phi}'\boldsymbol{X}(\boldsymbol{\theta})'\boldsymbol{Y}(\boldsymbol{\theta})) - \boldsymbol{Y}(\boldsymbol{\theta})'\boldsymbol{X}(\boldsymbol{\theta})\boldsymbol{\Phi} + \boldsymbol{\Phi}'\boldsymbol{X}(\boldsymbol{\theta})'\boldsymbol{X}(\boldsymbol{\theta})\boldsymbol{\Phi})\right)\right] \tag{62}$$

Considering γ replications of $Y(\boldsymbol{\theta})$ we get

$$p(\boldsymbol{Y}(\boldsymbol{\theta})|\boldsymbol{\Phi},\boldsymbol{\Sigma_u}) = (2\pi)^{-\frac{\gamma TN}{2}}|\boldsymbol{\Sigma_u}|^{-\frac{\gamma T}{2}} exp\left[-\frac{1}{2}tr\left((\boldsymbol{\Sigma_u}^{-1})(\boldsymbol{\gamma Y}(\boldsymbol{\theta})'\boldsymbol{Y}(\boldsymbol{\theta}) - \gamma\boldsymbol{\Phi}'\boldsymbol{X}(\boldsymbol{\theta})'\boldsymbol{Y}(\boldsymbol{\theta})) - \gamma\boldsymbol{Y}(\boldsymbol{\theta})'\boldsymbol{X}(\boldsymbol{\theta})\boldsymbol{\Phi} + \gamma\boldsymbol{\Phi}'\boldsymbol{X}(\boldsymbol{\theta})'\boldsymbol{X}(\boldsymbol{\theta})\boldsymbol{\Phi})\right)\right] \tag{63}$$

Next, following the approach of Del Negro et al. (2004) we replace the sample moments $\boldsymbol{X}(\boldsymbol{\theta})'\boldsymbol{X}(\boldsymbol{\theta})$, $\boldsymbol{X}(\boldsymbol{\theta})'\boldsymbol{Y}(\boldsymbol{\theta})$, $\boldsymbol{Y}(\boldsymbol{\theta})'\boldsymbol{Y}(\boldsymbol{\theta})$ by their expected values. Taking expectations conditionally on $\boldsymbol{\theta}$ we define:

$$\boldsymbol{\Gamma_{xx}}(\boldsymbol{\theta}) \equiv \mathbb{E}[\boldsymbol{X'X}|\boldsymbol{\theta}] = \begin{bmatrix} \boldsymbol{\Gamma_{xx,1}}(\boldsymbol{\theta}) & \boldsymbol{0}_k & \dots & \boldsymbol{0}_k \\ \boldsymbol{0}_k & \boldsymbol{\Gamma_{xx,2}}(\boldsymbol{\theta}) & \dots & \boldsymbol{0}_k \\ \vdots & \boldsymbol{0}_k & \ddots & \boldsymbol{0}_k \\ \boldsymbol{0}_k & \dots & \boldsymbol{0}_k & \boldsymbol{\Gamma_{xx,T}}(\boldsymbol{\theta}) \end{bmatrix} \tag{64}$$

$$\boldsymbol{\Gamma_{xy}}(\boldsymbol{\theta}) \equiv \mathbb{E}[\boldsymbol{X'X}|\boldsymbol{\theta}] = \begin{bmatrix} \boldsymbol{\Gamma_{xy,1}}(\boldsymbol{\theta}) \\ \boldsymbol{\Gamma_{xy,2}}(\boldsymbol{\theta}) \\ \dots \\ \boldsymbol{\Gamma_{xy,T}}(\boldsymbol{\theta}) \end{bmatrix} \tag{65}$$

$$\boldsymbol{\Gamma_{yy}}(\boldsymbol{\theta}) \equiv \mathbb{E}[\boldsymbol{Y'Y}|\boldsymbol{\theta}] = \sum_{t=1}^{T} \boldsymbol{\Gamma_{yy,t}}(\boldsymbol{\theta}) \tag{66}$$

where $\boldsymbol{\Gamma_{xx,t}} = \mathbb{E}[\boldsymbol{x_t x_t'}|\boldsymbol{\theta}]$, $\boldsymbol{\Gamma_{xy,t}} = \mathbb{E}[\boldsymbol{x_t y_t'}|\boldsymbol{\theta}]$ and $\boldsymbol{\Gamma_{yy,t}} \equiv \mathbb{E}[\boldsymbol{y_t y_t'}|\boldsymbol{\theta}]$ for $t = 1, \dots, T$. Substituting in the likelihood we get

$$p(\boldsymbol{Y}(\boldsymbol{\theta})^*|\gamma,\boldsymbol{\Phi},\boldsymbol{\Sigma_u}) = (2\pi)^{-\frac{\gamma TN}{2}}|\boldsymbol{\Sigma_u}|^{-\frac{\gamma T}{2}} exp\left[-\frac{1}{2}tr\left((\boldsymbol{\Sigma_u}^{-1})(\gamma\boldsymbol{\Gamma_{yy}}(\boldsymbol{\theta}) - \gamma\boldsymbol{\Phi}'\boldsymbol{\Gamma_{xy}}(\boldsymbol{\theta})) - \gamma\boldsymbol{\Gamma_{yx}}(\boldsymbol{\theta})\boldsymbol{\Phi} + \gamma\boldsymbol{\Phi}'\boldsymbol{\Gamma_{xx}}(\boldsymbol{\theta})\boldsymbol{\Phi})\right)\right] \tag{67}$$

A.1.3 Integrating constant of the theory coherent prior

The integrating constant of the *Normal-Inverse-Wishart* prior

$$c(\lambda, \boldsymbol{\theta}, \gamma) = \int_{-\infty}^{\infty} p(\boldsymbol{\Phi}, \boldsymbol{\Sigma_u}|\lambda) p(\boldsymbol{Y(\theta)}|\gamma, \boldsymbol{\Phi}, \boldsymbol{\Sigma_u}) d\boldsymbol{\Phi} d\boldsymbol{\Sigma_u} \tag{68}$$

is given by:

$$c(\boldsymbol{\lambda}, \boldsymbol{\theta}, \gamma) = (\pi)^{-\frac{\gamma N}{2}} \frac{\Gamma_N\left(\frac{\underline{\underline{\nu}}}{2}\right) |\underline{\underline{S}}|^{\frac{-\underline{\underline{\nu}}}{2}} |\underline{\underline{\boldsymbol{\Psi}}}|^{\frac{N}{2}}}{\Gamma_N\left(\frac{\underline{\nu}}{2}\right) |\underline{S}|^{-\frac{\underline{\nu}}{2}} |\underline{\boldsymbol{\Psi}}|^{\frac{N}{2}}} \tag{69}$$

where the definitions for $\underline{\underline{\boldsymbol{S}}}$, $\underline{\underline{\boldsymbol{\Psi}}}$, $\underline{\underline{\boldsymbol{\Phi}}}$,$\underline{\boldsymbol{\Psi}}$, $\underline{\boldsymbol{S}}$, $\underline{\boldsymbol{\Psi}}$, are given above in the text while $\boldsymbol{\Gamma}_N(.)$ is the Gamma function.

A.1.4 Marginal likelihood and fit-complexity trade off

The marginal likelihood is given by:

$$p(\boldsymbol{Y}|\lambda, \boldsymbol{\theta}, \gamma) = (\pi)^{-\frac{TN}{2}} \frac{\Gamma_N\left(\frac{\tilde{\nu}}{2}\right) |\tilde{\boldsymbol{S}}|^{-\frac{\tilde{\nu}}{2}} |\tilde{\boldsymbol{\Psi}}|^{\frac{N}{2}}}{\Gamma_N\left(\frac{\underline{\underline{\nu}}}{2}\right) |\underline{\underline{\boldsymbol{S}}}|^{-\frac{\underline{\underline{\nu}}}{2}} |\underline{\underline{\boldsymbol{\Psi}}}|^{\frac{N}{2}}} \tag{70}$$

Following the same steps as in (Giannone et al., 2015) it can be re-written as :

$$p(\boldsymbol{Y}|\lambda, \boldsymbol{\theta}, \gamma) = const |(\boldsymbol{V_\varepsilon^{post}})^{-1} \boldsymbol{V_\varepsilon^{prior}}|^{\frac{T+\underline{\underline{\nu}}}{2}} \prod_{t=1}^{T} |\boldsymbol{V_{t|t-1}}|^{-\frac{1}{2}} \tag{71}$$

where

$$vec(\hat{\boldsymbol{\varepsilon}}_{\boldsymbol{t}}) = vec(\boldsymbol{Y} - \boldsymbol{X}\tilde{\boldsymbol{\Phi}}) \tag{72}$$

$$\boldsymbol{V}_\varepsilon^{prior} = E[\boldsymbol{\Sigma}] = \frac{\underline{\underline{\boldsymbol{S}}}}{\underline{\underline{\nu}} - N - 1} \tag{73}$$

$$\boldsymbol{V}_\varepsilon^{post} = E[\boldsymbol{\Sigma}|\boldsymbol{Y}] = \frac{\underline{\underline{\boldsymbol{S}}} + \hat{\boldsymbol{\varepsilon}}'\hat{\boldsymbol{\varepsilon}} + (\tilde{\boldsymbol{\Phi}} - \underline{\underline{\boldsymbol{\Phi}}})' \underline{\underline{\boldsymbol{\Psi}}}^{-1} (\tilde{\boldsymbol{\Phi}} - \underline{\underline{\boldsymbol{\Phi}}})}{T + \underline{\underline{\nu}} - N - 1} \tag{74}$$

$$\boldsymbol{V}_{t|t-1} = \frac{\underline{\underline{\boldsymbol{S}}}}{\underline{\underline{\nu}} - N - 1} \otimes \left(1 + \boldsymbol{X}_{\boldsymbol{t}}'(\boldsymbol{X}_{\boldsymbol{t-1}}'\boldsymbol{X}_{\boldsymbol{t-1}} + \underline{\underline{\boldsymbol{\Psi}}}^{-\mathbf{1}})^{-\mathbf{1}}\boldsymbol{X}_{\boldsymbol{t}}\right) \tag{75}$$

$$const = \left(\frac{1}{\pi}\right)^{\frac{NT}{2}} \frac{\Gamma_N\left(\frac{T+\underline{\underline{\nu}}}{2}\right)}{\Gamma_N\left(\frac{\underline{\underline{\nu}}}{2}\right)} \frac{(T + \underline{\underline{\nu}} - N - 1)^{\underline{\underline{\nu}}/2}}{(\underline{\underline{\nu}} - N - 1)^{T+\underline{\underline{\nu}}/2}} \tag{76}$$

A.1.5 Formulas with distinct λ_j for $j = 1, \ldots, K$

In the case we consider regressor specific hyper-parameters λ_j with $j = 1, \ldots, K$, defining $\boldsymbol{\Lambda_k} = diag(\lambda_1, \ldots, \lambda_k)$ the theory coherent prior becomes:

$$p(vec(\boldsymbol{\Phi})|\boldsymbol{\Sigma}_u, \boldsymbol{\Lambda}, \boldsymbol{\theta}, \gamma) \sim \mathcal{N}(vec(\underline{\underline{\boldsymbol{\Phi}}}), \boldsymbol{\Sigma}_u \otimes \underline{\underline{\boldsymbol{\Psi}}}) \tag{77}$$

$$vec(\underline{\underline{\boldsymbol{\Phi}}}) = vec((\gamma\boldsymbol{\Gamma_{xx}}(\boldsymbol{\theta}) + \boldsymbol{H_{Tk}'}\boldsymbol{H_{Tk}}(\boldsymbol{I_T} \otimes \boldsymbol{\Lambda_k'}\boldsymbol{\Lambda_k}))^{-1} (\gamma\boldsymbol{\Gamma_{xy}}(\boldsymbol{\theta}) + \boldsymbol{H_{Tk}'}\boldsymbol{H_{Tk}}(\boldsymbol{I_T} \otimes \boldsymbol{\Lambda_k'}\boldsymbol{\Lambda_k})\underline{\boldsymbol{\Phi_0}})) \tag{78}$$

$$\underline{\underline{\boldsymbol{\Psi}}} = ((\gamma\boldsymbol{\Gamma_{xx}}(\boldsymbol{\theta}) + \boldsymbol{H_{Tk}'}\boldsymbol{H_{Tk}}(\boldsymbol{I_T} \otimes \boldsymbol{\Lambda_k'}\boldsymbol{\Lambda_k}))^{-1} \tag{79}$$

$$p(\boldsymbol{\Sigma}_u|\boldsymbol{\Lambda}, \boldsymbol{\theta}, \gamma) \sim \mathcal{IW}\left(\underline{\underline{\boldsymbol{S}}}, \underline{\underline{\nu}}\right) \tag{80}$$

$$\underline{\underline{\boldsymbol{S}}} = \underline{S} + \gamma\boldsymbol{\Gamma_{yy}}(\boldsymbol{\theta}) + \underline{\boldsymbol{\Phi_0}}'(\boldsymbol{H_{Tk}'}\boldsymbol{H_{Tk}})(\boldsymbol{I_T} \otimes \boldsymbol{\Lambda'}\boldsymbol{\Lambda})\underline{\boldsymbol{\Phi_0}} - \underline{\underline{\boldsymbol{\Phi}}}'\underline{\underline{\boldsymbol{\Psi}}}^{-1}\underline{\underline{\boldsymbol{\Phi}}} \tag{81}$$

$$\underline{\underline{\nu}} = \underline{\nu} + \gamma \tag{82}$$

The formula of the conditional posterior of $\boldsymbol{\Phi}$ and $\boldsymbol{\Sigma_u}$ becomes

$$\begin{aligned} p(vec(\boldsymbol{\Phi})|\boldsymbol{\Sigma}_u, \boldsymbol{\Lambda}, \boldsymbol{\theta}, \gamma, \boldsymbol{Y}) &\sim \mathcal{N}(vec(\tilde{\boldsymbol{\Phi}}), \boldsymbol{\Sigma}_u \otimes \tilde{\boldsymbol{\Psi}}) \\ p(\boldsymbol{\Sigma}_u|\boldsymbol{\Lambda}, \boldsymbol{\theta}, \gamma, \boldsymbol{Y}) &\sim \mathcal{IW}\left(\tilde{\boldsymbol{S}}, \tilde{\nu}\right) \end{aligned} \tag{83}$$

$$vec(\tilde{\boldsymbol{\Phi}}) = vec(\left(\boldsymbol{X'X} + \gamma\boldsymbol{\Gamma_{xx}}(\boldsymbol{\theta}) + \boldsymbol{H_{Tk}'}\boldsymbol{H_{Tk}}(\boldsymbol{I_T} \otimes \boldsymbol{\Lambda_k'}\boldsymbol{\Lambda_k})\right)^{-1} (\boldsymbol{X'Y} + \gamma\boldsymbol{\Gamma_{xy}}(\boldsymbol{\theta}) + \boldsymbol{H_{Tk}'}\boldsymbol{H_{Tk}}(\boldsymbol{I_T} \otimes \boldsymbol{\Lambda_k'}\boldsymbol{\Lambda_k})\underline{\boldsymbol{\Phi_0}})) \tag{84}$$

$$\tilde{\boldsymbol{\Psi}} = (\boldsymbol{X'X} + \gamma\boldsymbol{\Gamma_{xx}}(\boldsymbol{\theta}) + \boldsymbol{H_{Tk}'}\boldsymbol{H_{Tk}}(\boldsymbol{I_T} \otimes \boldsymbol{\Lambda_k'}\boldsymbol{\Lambda_k}))^{-1} \tag{85}$$

$$\tilde{\boldsymbol{S}} = \boldsymbol{Y'Y} + \underline{\underline{\boldsymbol{S}}} + \underline{\underline{\boldsymbol{\Phi}}}'\underline{\underline{\boldsymbol{\Psi}}}^{-1}\underline{\underline{\boldsymbol{\Phi}}} - \tilde{\boldsymbol{\Phi}}'\tilde{\boldsymbol{\Psi}}^{-1}\tilde{\boldsymbol{\Phi}} \tag{86}$$

$$\tilde{\nu} = \underline{\underline{\nu}} + T \tag{87}$$

The marginal likelihood is just (26), updated with the new definitions of $\underline{\underline{\nu}}$, $\underline{\underline{\boldsymbol{S}}}$, $\underline{\underline{\boldsymbol{\Psi}}}$, $\underline{\underline{\boldsymbol{\Phi}}}$,$\tilde{\boldsymbol{\Psi}}$, $\tilde{\boldsymbol{S}}$, $\tilde{\boldsymbol{\Psi}}$.

A.2 Forecasting exercise and small scale NK model

A.2.1 Data

The data for the out of sample forecasting exercise in Section 3 are taken from the FRED-QD Dataset from the Federal bank of St. Louis. The series IDs of the time series are `GDPC1`, `CPIAUCSL`, `FEDFUNDS`. Quarterly GDP growth and quarterly inflation rate are obtained transforming the series according to $\Delta\%y_t = 100\left(\frac{y_t - y_{t-1}}{y_{t-1}}\right)$. As for the Fed Funds interest rate it is transformed by taking the logarithm.

A.2.2 Competing models in the forecasting exercise:

The competing models in the out of sample forecasting exercise in Section 3 are

- A constant parameters VAR with flat prior.
- A constant parameters VAR with *Normal Inverse-Wishart* prior.
- A TVP-VAR model.

The VAR with *Normal Inverse-Wishart* prior is given by:

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\Pi} + \boldsymbol{U} \qquad\qquad \boldsymbol{U} \sim MVN(\boldsymbol{0}, \boldsymbol{\Sigma}_u, \boldsymbol{I_T}) \tag{88}$$

where $\boldsymbol{Y}$ is $T \times N$, $\boldsymbol{X}$ is $T \times k$ with $k = Np + 1$, $\boldsymbol{\Pi}$ is $k \times n$, $\boldsymbol{U}$ is $T \times N$ and MVN stands for the matricvariate normal. The prior for the autoregressive coefficients and the variance covariance matrix is:

$$vec(\boldsymbol{\Pi}) \sim \mathcal{N}(vec(\underline{\boldsymbol{\mu}_{\boldsymbol{\Pi}}}), \boldsymbol{\Sigma}_u \otimes \underline{\boldsymbol{\Omega}_{\boldsymbol{\Pi}}}) \tag{89}$$

$$\boldsymbol{\Sigma} \sim \mathcal{IW}(\boldsymbol{S_0}, v_0) \tag{90}$$

where $vec(\boldsymbol{\Pi})$ centers the non-stationary series on a random walk process and the stationary ones on a white noise process and $\underline{\boldsymbol{\Omega_\Pi}}$ is a diagonal matrix with first element element equal to

$$\omega_1 = 100 \tag{91}$$

and

$$\omega_s = \frac{\theta_1}{\sigma_i^2 l^2} \tag{92}$$

for the other elements $s = 2, \ldots, Np$ with $\theta_1 = 0.1$ and σ_i^2 is set equal to $\hat{\sigma}_i^2$ being the estimated variance of the i^{th} variable in a VAR model using a pre-sample of observations. l is the lag order of the variable associated to that variable. $\boldsymbol{S_0}$ is set equal to $\hat{\boldsymbol{S}}_\mathbf{0}$ the variance covariance matrix estimated variance from a VAR model on a pre-sample of observations and $v_0 = N + 2$. The standard TVP-VAR model and the corresponding choice of the prior distributions follows Chan et al. (2009). The model is given by:

$$\underbrace{\boldsymbol{y_t'}}_{1 \times N} = \underbrace{\boldsymbol{x_t'}}_{1 \times k} \underbrace{\boldsymbol{\Phi}_t}_{k \times N} + \underbrace{\boldsymbol{u_t'}}_{1 \times N} \qquad \boldsymbol{u_t} \sim \mathcal{N}(\mathbf{0}_{N \times 1}, \boldsymbol{\Sigma_u}) \tag{93}$$

$$vec(\boldsymbol{\Phi_t}) = vec(\boldsymbol{\Phi_{t-1}}) + \boldsymbol{\eta}_t \qquad \boldsymbol{\eta}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}) \tag{94}$$

where $\boldsymbol{\Omega} = diag(\omega_1, \ldots, \omega_{Nk})$ [34]

$$\omega_i \sim \mathcal{IG}\left(\nu, s^2\right) \tag{95}$$

where $\nu = 3$ $s^2 = 0.005$ and the prior for the variance covariance matrix is set as in the constant parameter BVAR with *Normal-Inverse-Wishart* prior.

[34] Note that in this model the variances in the state equation of the time varying coefficients are not constrained to be proportional across equations (the structure of $\boldsymbol{\Omega}$ is not constrained to be a Kronecker product $\boldsymbol{\Omega} = \boldsymbol{\Sigma}_u \otimes \boldsymbol{\Lambda_k}$).

A.2.3 Prior for the DSGE parameters

Table 2: Prior distribution for the parameters of the New Keynesian model

Parameter	Prior distribution	Mean	Standard Deviation
$ln(\tilde{\gamma}s)$	Normal	0.500	0.250
$ln(\pi^*)$	Normal	1.000	0.500
$ln(r^*)$	Gamma	0.500	0.250
κ	Gamma	0.300	0.150
τ	Gamma	2.000	0.500
ψ_1	Gamma	1.500	0.250
ψ_2	Gamma	0.500	0.200
ρ_R	Beta	0.500	0.250
ρ_g	Beta	0.800	0.100
ρ_z	Beta	0.300	0.100
σ_R	Inverse Gamma	0.251	0.139
σ_g	Inverse Gamma	0.630	0.323
σ_z	Inverse Gamma	0.875	0.430

The table reports the details on the prior distribution of the parameters of the New Keynesian model.

A.3 Medium scale New Keynesian model

The model is taken from Del Negro et al. (2015) and it is a version of the popular medium scale New Keynesian model in Smets et al. (2007). The set of log-linearized equilibrium conditions of the model is

$$c_t = -\frac{(1-he^{-\gamma})}{\sigma_c(1+he^{-\gamma})}(R_t - E_t[\pi_{t+1}] + b_t) + \frac{(he^{-\gamma})}{(1+he^{-\gamma})}(c_{t-1} - z_t) + \frac{1}{(1+he^{-\gamma})}E_t[c_{t+1} + z_{t+1}] + \frac{(\sigma_c - 1)}{\sigma_c(1+he^{-\gamma})}\frac{w_* l_*}{c_*}(i_t - E_t[l_{t+1}]) \tag{96}$$

$$q_t^k = S''e^{2\gamma}(1+\bar{\beta})\left(i_t - \frac{1}{1+\bar{\beta}}(i_{t-1} - z_t) - \frac{\bar{\beta}}{1+\bar{\beta}}E_t[i_{t+1} + z_{t+1}] - \mu_t\right) \tag{97}$$

$$\bar{k}_t = \left(1 - \frac{i_*}{k_*}\right)(\bar{k}_{t-1} - z_t) + \frac{i_*}{k_*}i_t + \frac{i_*}{k_*}S''e^{2\gamma}(1+\bar{\beta})\mu_t \tag{98}$$

$$\frac{r_*^k}{r_*^k + (1-\delta)}E_t[r_{t+1}^k] + \frac{1-\delta}{r_*^k + (1-\delta)}E_t[q_{t+1}^k] - q_t^k = R_t + b_t - E_t[\pi_{t+1}] \tag{99}$$

$$k_t = u_t - z_t + \bar{k}_{t-1} \tag{100}$$

$$\frac{1-\psi}{\psi}r_t^k = u_t \tag{101}$$

$$mc_t = w_t + \alpha l_t - \alpha k_t \tag{102}$$

$$k_t = w_t - r_t^k + l_t \tag{103}$$

$$y_t = \Phi_p(\alpha k_t + (1-\alpha)l_t) + \mathcal{I}\{\rho_z < 1\}(\Phi_p - 1)\frac{1}{1-\alpha}\tilde{z}_t \tag{104}$$

$$y_t = g_t + \frac{c_*}{y_*}c_t + \frac{i_*}{y_*}i_t + \frac{r_*^k k_*}{y_*}u_t - \mathcal{I}\{\rho_z < 1\}\frac{1}{1-\alpha}\tilde{z}_t \tag{105}$$

$$g_t = \rho_g g_{t-1} + \sigma_g \varepsilon_{g,t} + \eta_{gz}\sigma_z \varepsilon_{z,t} \tag{106}$$

$$\pi_t = \kappa mc_t \frac{\iota_p}{1+\iota_p\bar{\beta}}\pi_{t-1} + \frac{\bar{\beta}}{1+\iota_p\bar{\beta}}E_t[\pi_{t+1}] + \lambda_{f,t} \tag{107}$$

$$w_t = \frac{(1-\zeta_w\bar{\beta})(1-\zeta_w)}{(1+\bar{\beta})\zeta_w((\lambda_w - 1)\epsilon_w + 1)}(w^h - w_t) - \frac{1+\iota_w\bar{\beta}}{(1+\bar{\beta})}\pi_t + \frac{1}{1+\bar{\beta}}(w_{t-1} - z_t - \iota_w\pi_{t-1}) + \frac{\bar{\beta}}{1+\bar{\beta}}E_t[w_{t+1} + z_{t+1} + \pi_{t+1}] + \lambda_{w,t} \tag{108}$$

$$w_t^h = \frac{1}{1-he^{-\gamma}}(c_t - he^{-\gamma}c_{t-1} + he^{-\gamma}z_t) + \nu_l l_t \tag{109}$$

$$\lambda_{f,t} = \rho_{\lambda_f}\lambda_{f,t-1} + \sigma_{\lambda_f}\varepsilon_{\lambda_f,t} - \eta_{\lambda_f}\sigma_{\lambda_f}\varepsilon_{\lambda_f,t-1} \tag{110}$$

$$\lambda_{w,t} = \rho_{\lambda_w}\lambda_{w,t-1} + \sigma_{\lambda_w}\varepsilon_{\lambda_w,t} - \eta_{\lambda_w}\sigma_{\lambda_w}\varepsilon_{\lambda_w,t-1} \tag{111}$$

$$R_t = \rho_R R_{t-1} + (1-\rho_R)\left(\psi_1(\pi_t - \pi_t^*) + \psi_2(y_t - y_t^f)\right) + \psi_3\left((y_t - y_t^f) - (y_{t-1} - y_{t-1}^f)\right) + r_t^m \tag{112}$$

$$E_t[\tilde{R}_{t+1}^k - R_t] = b_t \tag{113}$$

$$\tilde{R}_{t+1}^k - \pi_t = \frac{r_*^k}{r_*^k + (1-\delta)} r_t^k + \frac{(1-\delta)}{r_*^k + (1-\delta)} q_t^k - q_{t-1}^k \tag{114}$$

$$b_t = \rho_b b_{t-1} + \sigma_b \varepsilon_{b,t} \tag{115}$$

$$\mu_t = \rho_\mu \mu_{t-1} + \sigma_\mu \varepsilon_{\mu,t} \tag{116}$$

$$\tilde{z}_t = \rho_z \tilde{z}_{t-1} + \sigma_z \varepsilon_{z,t} \tag{117}$$

$$r_t^m = \rho_r \tilde{z}_{t-1} + \sigma_r \varepsilon_{r,t} \tag{118}$$

The observation equations are:

$$\text{Output growth} = \gamma + 100(y_t - y_{t-1} + z_t) \tag{119}$$

$$\text{Consumption growth} = \gamma + 100(c_t - c_{t-1} + z_t) \tag{120}$$

$$\text{Investment growth} = \gamma + 100(i_t - i_{t-1} + z_t) \tag{121}$$

$$\text{Real wage growth} = \gamma + 100(w_t - w_{t-1} + z_t) \tag{122}$$

$$\text{Hours worked} = \bar{l} + 100 l_t \tag{123}$$

$$\text{Inflation} = \pi_* + 100\pi_t \tag{124}$$

$$\text{FFR} = R_* + 100(R_t) \tag{125}$$

Data sources and transformations are as in Del Negro et al. (2015) to which I refer for further details. During the ZLB period, the observation equation for the interest

rate includes a measurement error, namely:

$$\text{FFR} = R_* - R_* + v_{r,t} \tag{126}$$

with $\mathbb{E}[v_{r,t}] = 0$ and $var(v_{r,t}) = 0.001$. This is to account for the fact that the Fed Fund Rate was not exactly equal to zero, remaining slightly above zero in the ZLB.

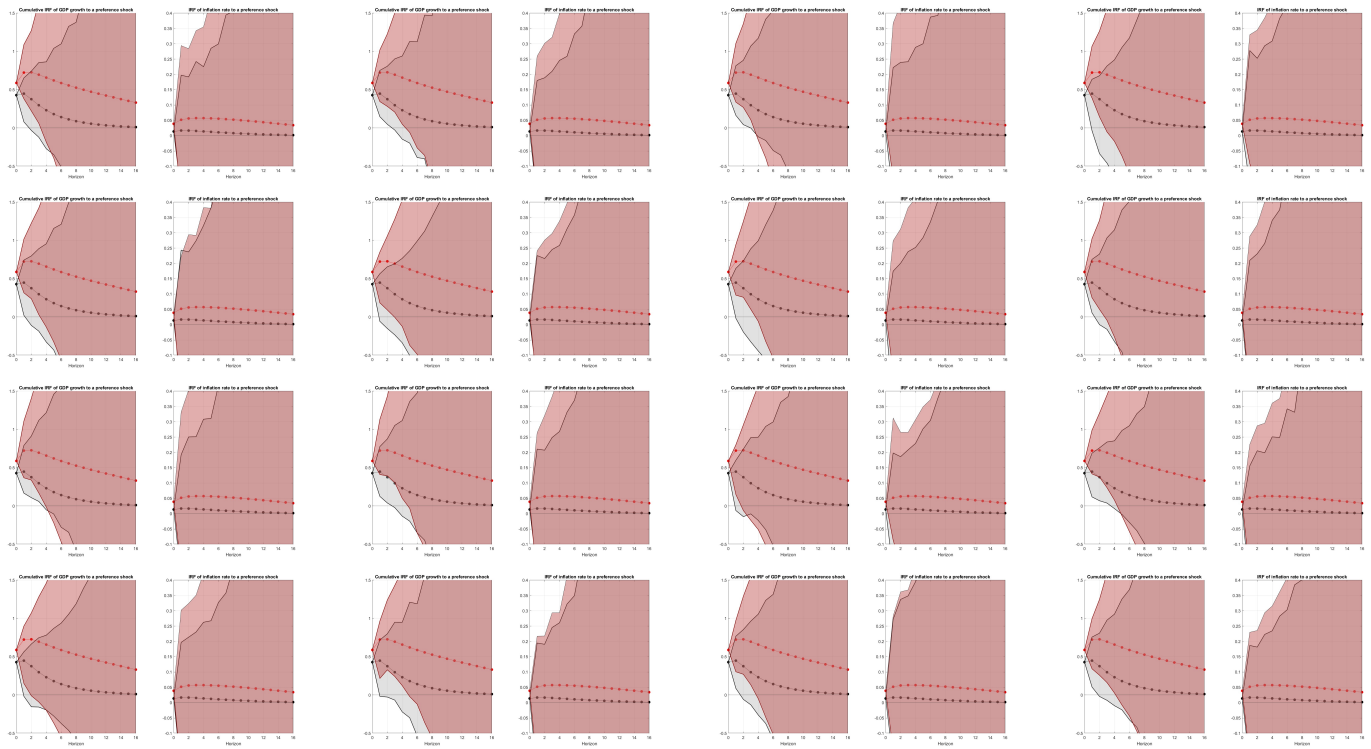


Figure 8: Estimated 10^{th} and 90^{th} credible intervals of responses of output growth (cumulative) and inflation rate inside (in shaded red) and outside (in shaded grey) the ZLB period from a standard TVP-VAR on distinct simulated samples from the NK model. The dotted line plots are the responses inside (in red) and outside (in black) the ZLB period in the NK model.

A.3.1 Prior for the structural parameters

Table 3: Prior distribution for the parameters of the New Keynesian model

Parameter	Prior distribution	Mean	Standard Deviation	Parameter	Prior distribution	Mean	Standard Deviation
ψ_1	Gamma	1.500	0.250	ρ_R	Beta	0.750	0.100
ψ_2	Gamma	0.120	0.050	ψ_2	Beta	0.500	0.200
ψ_3	Gamma	1.12	0.050	ψ_3	Inverse-Gamma	0.10	2.000
ζ_p	Beta	0.500	0.100	ζ_w	Beta	0.500	0.100
α	Normal	0.300	0.050	π^*	Gamma	0.750	0.400
Φ	Normal	1.250	0.120	γ	Normal	0.400	0.100
h	Beta	0.700	0.100	S''	Normal	4.000	1.500
ν_l	Normal	2.000	0.750	σ_c	Normal	1.500	0.370
ι_p	Beta	0.500	0.500	ι_w	Beta	0.500	0.150
r_*	Gamma	1.500	0.250	ψ	Beta	0.500	0.150
ρ_z	Beta	0.500	0.200	σ_z	Inverse-Gamma	0.100	2.000
ρ_b	Beta	0.500	0.250	σ_b	Inverse-Gamma	0.100	2.000
ρ_{λ_f}	Beta	0.500	0.200	σ_{λ_f}	Inverse-Gamma	0.100	2.000
ρ_{λ_w}	Beta	0.500	0.200	σ_{λ_w}	Inverse-Gamma	0.100	2.000
ρ_μ	Beta	0.500	0.200	σ_μ	Inverse-Gamma	0.100	2.000
ρ_g	Beta	0.500	0.200	σ_g	Inverse-Gamma	0.100	2.000
η_{λ_f}	Beta	0.500	0.200	η_{λ_f}	Beta	0.500	0.200
η_{gz}	Beta	0.500	0.200				

The table reports the details on the prior distribution of the parameters of the medium scale New Keynesian model which accounts for the ZLB period and forward guidance. The choice of the prior follows Del Negro, Giannoni, and Schorfheide (2015).

A.3.2 Risk premium shocks

Risk premium shocks in the Smets et al. (2007) are financial shocks that drive a wedge between the risk-free interest rate and the actual interest rates faced by borrowers. Figure 9 shows the impulse response functions to a negative risk premium shock in the Smets et al. (2007) model. The shock propagates through the economy like a pure demand shock, by raising both real activity and inflation.

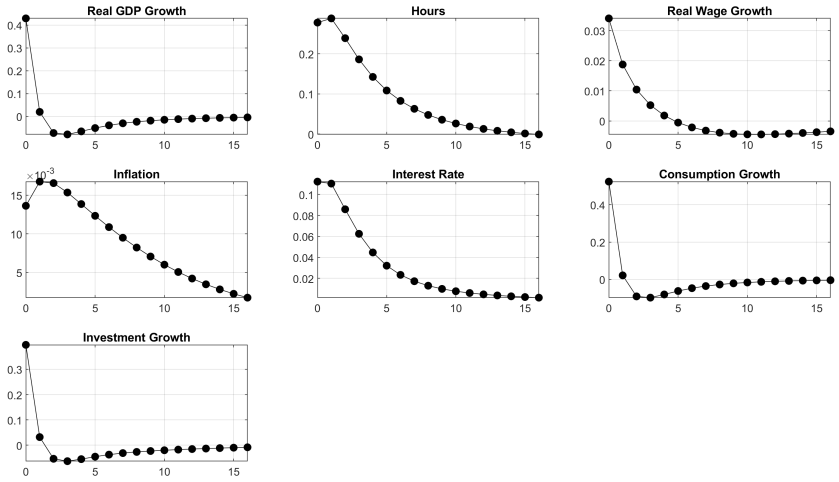


Figure 9: The figure shows the impulse response functions to one standard deviation risk premium shock in the Smets and Wouters (2007) model.

Considering a standard VAR model with the seven variables of the Smets et al. (2007) model augmented with the Excess Bond Premium (EBP) series by Gilchrist and Zakrajšek (2012), we think of the orthogonal component of the EPB as akin to the risk premium shocks in the Smets and Wouters model. Figure 10, shows the impulse response function to a one standard deviation shock to the EBP Gilchrist et al. (2012). The shocks are identified through recursive identification by ordering the EBP series last in the VAR, that is allowing the excess bond premium series to contemporaneously respond to the other shocks in the VAR.

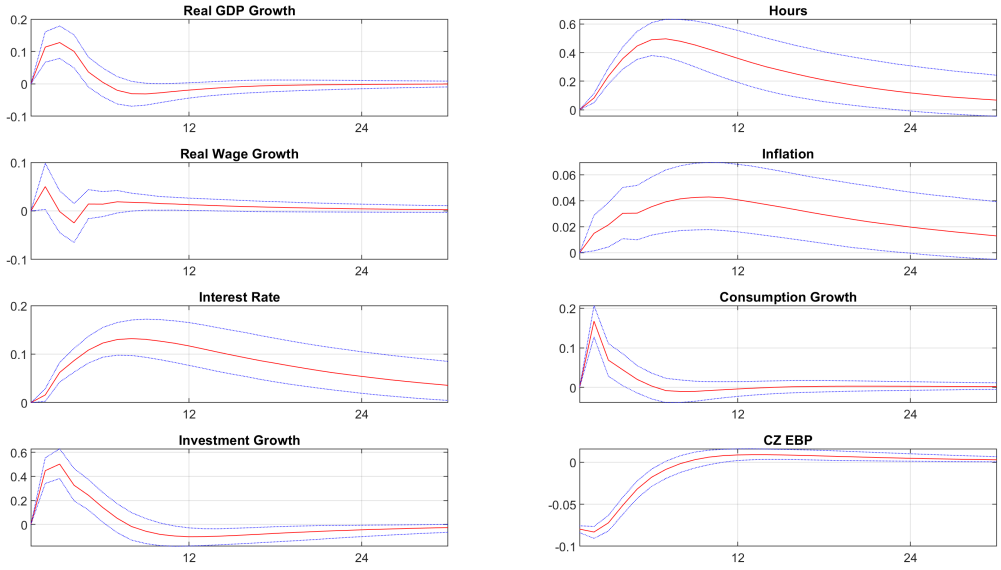


Figure 10: The figure reports the impulse response functions to a negative one standard deviation shock to the excess bond premium of Gilchrist and Zakrajšek (2012). In red the posterior median, while in blue dotted line the $84^{th} - 16^{th}$ credible intervals.

.

References

Aruoba, S. Borağan, Marko Mlikota, Frank Schorfheide, and Sergio Villalvazo (2022). "SVARs with occasionally-binding constraints". In: *Journal of Econometrics* 231.2. Special Issue: The Econometrics of Macroeconomic and Financial Data, pp. 477–499.

Ascari, Guido, Paolo Bonomolo, and Qazi Haque (2023). "The long-run phillips curve is... a curve". In.

Ascari, Guido and Argia M. Sbordone (2014). "The Macroeconomics of Trend Inflation". In: *Journal of Economic Literature* 52.3, pp. 679–739.

Belmonte, Miguel A.G., Gary Koop, and Dimitris Korobilis (2014). "Hierarchical Shrinkage in Time-Varying Parameter Models". In: *Journal of Forecasting* 33.1, pp. 80–94.

Benati, Luca and Thomas A. Lubik (June 2023). *Impulse Response Analysis at the Zero Lower Bound*. Diskussionsschriften dp2306. Universitaet Bern, Departement Volkswirtschaft.

Bitto, Angela and Sylvia Frühwirth-Schnatter (2019). "Achieving shrinkage in a time-varying parameter model framework". In: *Journal of Econometrics* 210.1. Annals Issue in Honor of John Geweke "Complexity and Big Data in Economics and Finance: Recent Developments from a Bayesian Perspective", pp. 75–97.

Cagliarini, Adam and Mariano Kulish (Mar. 2013). "Solving Linear Rational Expectations Models with Predictable Structural Changes". In: *The Review of Economics and Statistics* 95.1, pp. 328–336.

Carriero, Andrea, Todd E. Clark, and Massimiliano Marcellino (2021). "No-arbitrage priors, drifting volatilities, and the term structure of interest rates". In: *Journal of Applied Econometrics* 36.5, pp. 495–516.

Carter, C. K. and R. Kohn (1994). "On Gibbs Sampling for State Space Models". In: *Biometrika* 81.3, pp. 541–553.

Chan, Joshua C.C., Eric Eisenstat, and Rodney W. Strachan (2020). "Reducing the state space dimension in a large TVP-VAR". In: *Journal of Econometrics* 218.1, pp. 105–118.

Chan, Joshua CC and Ivan Jeliazkov (2009). "Efficient simulation and integrated likelihood estimation in state space models". In: *International Journal of Mathematical Modelling and Numerical Optimisation* 1.1-2, pp. 101–120.

Chen, Han, Vasco Cúrdia, and Andrea Ferrero (2012). "The Macroeconomic Effects of Large-scale Asset Purchase Programmes*". In: *The Economic Journal* 122.564, F289–F315.

Cogley, Timothy and Thomas Sargent (2002). "Evolving Post-World War II US Inflation Dynamics". In: *NBER Macroeconomics Annual 2001, Volume 16*. National Bureau of Economic Research, Inc, pp. 331–388.

Cogley, Timothy and Thomas J. Sargent (2005). "Drifts and volatilities: monetary policies and outcomes in the post WWII US". In: *Review of Economic Dynamics* 8.2. Monetary Policy and Learning, pp. 262–302.

Cogley, Timothy and Argia M. Sbordone (Dec. 2008). "Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve". In: *American Economic Review* 98.5, pp. 2101–26.

Coulombe, Philippe Goulet (2021). *Time-Varying Parameters as Ridge Regressions*.

D'Agostino, Antonello, Luca Gambetti, and Domenico Giannone (2013). "Macroeconomic forecasting and structural change". In: *Journal of Applied Econometrics* 28.1, pp. 82–101.

Debortoli, Davide, Jordi Galí, and Luca Gambetti (Oct. 2019). "On the Empirical (Ir)relevance of the Zero Lower Bound Constraint". In: *NBER Macroeconomics Annual 2019, volume 34*. NBER Chapters. National Bureau of Economic Research, Inc, pp. 141–170.

Del Negro, Marco, Marc P. Giannoni, and Frank Schorfheide (Jan. 2015). "Inflation in the Great Recession and New Keynesian Models". In: *American Economic Journal: Macroeconomics* 7.1, pp. 168–96.

Del Negro, Marco and Frank Schorfheide (2004). "Priors from General Equilibrium Models for VARS*". In: *International Economic Review* 45.2, pp. 643–673.

Farmer, Roger E.A., Daniel F. Waggoner, and Tao Zha (2009). "Understanding Markov-switching rational expectations models". In: *Journal of Economic Theory* 144.5, pp. 1849–1867.

Fessler, Pirmin and Maximilian Kasy (Oct. 2019). "How to Use Economic Theory to Improve Estimators: Shrinking Toward Theoretical Restrictions". In: *The Review of Economics and Statistics* 101.4, pp. 681–698.

Frühwirth-Schnatter, Sylvia and Helga Wagner (2010). "Stochastic model specification search for Gaussian and partial non-Gaussian state space models". In: *Journal of Econometrics* 154.1, pp. 85–100.

Gallant, A. Ronald and Robert E. McCulloch (2009). "On the Determination of General Scientific Models With Application to Asset Pricing". In: *Journal of the American Statistical Association* 104.485, pp. 117–131.

Giacomini, Raffaella and Giuseppe Ragusa (2014). "Theory-coherent forecasting". In: *Journal of Econometrics* 182.1. Causality, Prediction, and Specification Analysis: Recent Advances and Future Directions, pp. 145–155.

Giannone, Domenico, Michele Lenza, and Giorgio E. Primiceri (May 2015). "Prior Selection for Vector Autoregressions". In: *The Review of Economics and Statistics* 97.2, pp. 436–451.

Gilchrist, Simon and Egon Zakrajšek (2012). "Credit Spreads and Business Cycle Fluctuations". In: *American Economic Review* 102.4, pp. 1692–1720.

Gourio, François and Phuong Ngo (2020). "Risk premia at the ZLB: a macroeconomic interpretation". In.

Hauzenberger, Niko, Florian Huber, Massimiliano Marcellino, and Nico Petz (2022). *Gaussian Process Vector Autoregressions and Macroeconomic Uncertainty*.

Huber, Florian, Gary Koop, and Luca Onorante (July 2021). "Inducing Sparsity and Shrinkage in Time-Varying Parameter Models". In: *Journal of Business & Economic Statistics* 39.3, pp. 669–683.

Ingram, Beth F. and Charles H. Whiteman (Dec. 1994). "Supplanting the 'Minnesota' prior: Forecasting macroeconomic time series using real business cycle model priors". In: *Journal of Monetary Economics* 34.3, pp. 497–510.

Kalli, Maria and Jim E. Griffin (2014). "Time-varying sparsity in dynamic regression models". In: *Journal of Econometrics* 178.2, pp. 779–793.

Loria, Francesca, Christian Matthes, and Mu-Chun Wang (2022). "Economic theories and macroeconomic reality". In: *Journal of Monetary Economics* 126, pp. 105–117.

Mavroeidis, Sophocles (2021). "Identification at the zero lower bound". In: *Econometrica* 89.6, pp. 2855–2885.

Milani, Fabio (2007). "Expectations, learning and macroeconomic persistence". In: *Journal of Monetary Economics* 54.7, pp. 2065–2082.

Primiceri, Giorgio E (2005). "Time varying structural vector autoregressions and monetary policy". In: *The Review of Economic Studies* 72.3, pp. 821–852.

Sims, Christopher (2002). "Solving Linear Rational Expectations Models". In: *Computational Economics* 20.1-2, pp. 1–20.

Sims, Christopher A. (1980). "Macroeconomics and Reality". In: *Econometrica* 48.1, pp. 1–48.

Smets, Frank and Rafael Wouters (June 2007). "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach". In: *American Economic Review* 97.3, pp. 586–606.

Smith Jr., A. A. (1993). "Estimating nonlinear time-series models using simulated vector autoregressions". In: *Journal of Applied Econometrics* 8.S1, S63–S84.

Stevanovic, Dalibor (Apr. 2016). "Common time variation of parameters in reduced-form macroeconomic models". In: *Studies in Nonlinear Dynamics & Econometrics* 20.2, pp. 159–183.

Stock, James H. and Mark W. Watson (Dec. 2001). "Vector Autoregressions". In: *Journal of Economic Perspectives* 15.4, pp. 101–115.

Wind, Joris de and Luca Gambetti (Mar. 2014). *Reduced-rank time-varying vector autoregressions*. CPB Discussion Paper 270. CPB Netherlands Bureau for Economic Policy Analysis.

Woodford, Michael (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

Modelling and Forecasting Macroeconomic Risk with Time Varying Skewness Stochastic Volatility Models *

Andrea Renzetti†

Abstract

Monitoring downside risk and upside risk to the key macroeconomic indicators is critical for effective policymaking aimed at maintaining economic stability. In this paper I propose a parametric framework for modelling and forecasting macroeconomic risk based on stochastic volatility models with *Skew-Normal* and *Skew-t* shocks featuring time varying skewness. Exploiting a mixture stochastic representation of the *Skew-Normal* and *Skew-t* random variables, in the paper I develop efficient posterior simulation samplers for Bayesian estimation of both univariate and VAR models of this type. In an application, I use the models to predict downside risk to GDP growth in the US and I show that these models represent a competitive alternative to semi-parametric approaches such as quantile regression. Finally, estimating a medium scale VAR on US data I show that time varying skewness is a relevant feature of macroeconomic and financial shocks.

J.E.L Classification Code: C22, C32, C53

Keywords: Stochastic Volatility, Stochastic Skewness, Bayesian VARs, Macroeconomic tail risk

*This paper is part of my PhD dissertation at the University of Bologna. I am grateful to Andrea Carriero for invaluable support and guidance. I thank Luca Fanelli and Umberto Cherubini for comments and suggestions and participants at the March 2021 Unibo PhD Forum, the July 2022 Unibo PhD Forum and the Oslo IAAE 2023 conference.

†Department of Economics, Alma Mater Studiorium Università di Bologna, Piazza Scaravilli 2, 40126 Bologna, Italy

1 Introduction

Central banks and policy institutions play a critical role in maintaining financial stability and fostering economic growth. A key challenge they face is effectively monitoring the likelihood of severe events that could have adverse effects on the economy. Failing to adequately assess these risks can lead to underestimation of potential losses and insufficient policy responses. To address this challenge, it is essential to develop econometric tools that can accurately predict and assess tail risk in macroeconomic outcomes. In this paper, I propose an econometric framework specifically designed for modeling and forecasting macroeconomic tail risk. The framework relies on fully parametric univariate and multivariate stochastic volatility models with *Skew-Normal* and *Skew-t* shocks featuring stochastic skewness. These models aim to capture and predict persistent time-varying asymmetries in the future distribution of the variables of interest. Capturing these asymmetries is especially relevant given the risk management nature of the problem of policymaking faced by central banks and policy institutions (Kilian et al. 2003).

The paper begins by extending the well-known univariate stochastic volatility model introduced by Jacquier et al. (1994) to explicitly account for time-varying conditional skewness in the predictive distribution of a single target variable. Then, building upon the univariate approach, the paper introduces a Bayesian Vector Autoregressive (VAR) model with stochastic volatility and time-varying skewness. By allowing to track changes in the shape of the predictive distribution of multiple time series, this model is suitable for quantification and forecasting of tail risk to multiple target variables. Importantly, the model retains all the advantages and familiar toolkit for policy analysis and scenario analysis associated to the VAR framework. The model is estimated through an efficient *Gibbs sampler* that exploits a convenient mixture stochastic representation of the *Skew-Normal* and *Skew-t* shocks. To test the effectiveness of the proposed framework, I use the time-varying skewness stochastic volatility models to monitor downside risk to GDP growth in the US economy. The findings of this analysis align with the main conclusions of Adrian et al. (2019), revealing a nonlinear and asymmetric impact of financial conditions on the future distribution of GDP growth. Additionally, the models provide slightly more accurate out-of-sample forecasts of downside risk compared to quantile regression, which is often considered as the benchmark model in this literature. Furthermore, estimating a medium-scale VAR model of monetary policy, I show that shocks to financial and macroeconomic time series exhibit both time-varying volatility and time-varying skewness, suggesting that taking into considera-

tion both of these features might be of particularly relevance for accurately assessing upside and downside risk to macroeconomic indicators.

Related literature A fast-growing body of studies recently used univariate quantile regression methods for modelling and predicting asymmetries in the future distribution of the macroeconomic variables of interest. For example, Giglio et al. (2016) used predictive quantile regression to investigate whether systemic risk indicator and financial distress indicators predict changes in the lower quantiles of future macroeconomics shocks. As well, Kiley (2018) used quantile regression to examine fluctuations in the risk of a large increase in unemployment. More recently, Adrian et al. (2019) used a two step-procedure based on predictive quantile regression and quantile interpolation to model changes in downside risk to future GDP growth as a function of current financial and economic conditions.[1] Despite its popularity, the quantile regression method of Adrian et al. (2019) typically fails in the presence of a large information set where fully parametric models often produce more accurate forecasts of downside risks (Carriero et al. 2020). As a matter of fact, when using quantile regression, including multiple lags of the dependent and independent variables so as to capture the rich autocorrelation structure of macroeconomic and financial time series becomes very impractical and often leads to imprecise estimates of the coefficients and problems such as quantile crossing. Moreover, the entire predictive distribution of the target variables can only be obtained in two steps by interpolating the estimated quantiles with a flexible distribution. In the light of these limitations a new wave of studies have recently brought some evidences in favour of the use of fully parametric models to assess and predict tail risk to macroeconomic outcomes. Brownlees et al. (2021) for example, show that standard GARCH models have superior forecasting performance with respect to quantile regression methods for forecasting downside risk to GDP growth. As well, Carriero et al. (2020) show that a Bayesian VAR with stochastic volatility performs comparably to quantile regression for estimating and forecasting tail risks. Here I follow and extend this line of research by considering fully parametric models featuring both time varying volatility and time varying skewness, as recently done by Delle Monache et al. (2021), Iseringhausen (2021), Wolf (2021)

1. The two step approach based on quantile regression of Adrian et al. (2019) gained substantial popularity in the literature and has been employed in many other frameworks to assess and predict tail risk to economic outcomes. Among the others, López-Salido et al. (2020) used the two step approach of Adrian et al. (2019) for assessing and predicting downside and upside risk to inflation while Gelos et al. (2022) used the same approach for predicting the probability of large capital out-flows and in-flows to emerging markets.

and Montes-Galdón et al. (2022). While the first three contributions are all univariate [2] in this paper I model time varying volatility together with time varying skewness both in a univariate and in a multivariate framework. The main advantages of the multivariate framework is that it allows to jointly model the dynamic relationship between the target variables and the risk factors and to explicitly model tail risk to multiple macroeconomic outcomes of interest. The multivariate model that I propose in this paper is a VAR model in which Bayesian shrinkage can be conveniently used to avoid over-fitting when exploiting a potential large information set due both to the inclusion of larger number of macroeconomic variables and of a meaningful number of lags needed to properly account for the rich autocorrelation structure of the macroeconomic and financial time series. The model features two distinct stochastic processes respectively governing the time varying volatility and the time varying skewness of the shocks. By considering distinct stochastic processes for the skewness and the volatility of the shocks, this model is different from the Bayesian VAR with *Skew-Normal* shocks introduced by Montes-Galdón et al. (2022) where the latent stochastic process governing the shape of the shocks influences not only the conditional skewness, but also the conditional mean and the conditional variance of the variables in the system. As well, the model differs from Karlsson et al. (2023) who recently proposed a general class of generalized hyperbolic skew Student's distribution with stochastic volatility for the shocks of the VAR in which the time variation in the volatility of the shocks drives also time variation in their skewness. To my knowledge, this is the first paper that estimates a VAR with two distinct stochastic processes for the volatility and the skewness of the shocks.

Outline The rest of the article is organized as follows. In Section 2.1 I present the univariate stochastic volatility models with *Skew-Normal* and *Skew-t* shocks featuring time varying skewness. Then in Section 2.2 I exploit the same conceptual framework to model time varying skewness together with time varying volatility in the shocks of a VAR model. In both sections I present posterior simulation samplers used for Bayesian estimation of these models. In Section 3 I and use the models to predict downside risk to GDP growth and compare the forecasting performances to the popular two step approach based on quantile regression by Adrian et al. (2019).

2. Delle Monache et al. (2021) propose a score driven model with *Skew-t* innovations. Iseringhausen (2020) is the first paper to introduce time varying conditional skewness in a univariate stochastic volatility model by exploiting a *Noncentral-t* distribution for the innovations. Wolf (2021) exploits the *Skew-Normal* distributions but considers a different parametrization for the shocks with respect to the univariate model that I consider in Section 2.1 relying as well on a different estimation strategy.

Finally, in Section 4 I estimate a medium scale VAR model and show that many macroeconomic and financial variables exhibit time varying conditional skewness.

2 Models

2.1 Univariate time varying skewness stochastic volatility model

Stochastic volatility models currently represent the state of the art for modelling and forecasting macroeconomic and financial time series. The basic stochastic volatility model of Jacquier et al. (1994) specifies a *log-normal* auto-regressive process for the conditional variance with independent innovations in the conditional mean and conditional variance equation. In a second contribution, Jacquier et al. (2004) introduce a stochastic volatility model that features correlation between the volatility and mean innovations (*leverage effects*) allowing for conditional skewness, but without modelling it explicitly. Cappuccio et al. (2004) present a stochastic volatility model where the shocks feature a *Skew-GED* distribution while Abanto-Valle et al. (2015) introduce a stochastic volatility with *Skew-t* innovations. Both contributions explicitly model conditional skewness, but do not allow for time varying conditional skewness. Here I present a direct extension of the univariate stochastic volatility model of Jacquier et al. (1994) that instead explicitly allows for time varying conditional skewness.

In order to model asymmetries in the conditional distribution of the dependent variable, I assume that the innovations in an otherwise standard stochastic volatility model follow a potentially asymmetric distribution, being the *Skew-Normal* (Azzalini 1986) and the *Skew-t* (Azzalini et al. 2003) distribution. The $Skew-Normal(\zeta, \omega^2, \lambda)$ is an asymmetric distribution fully characterized by three parameters: the location parameter ζ, the scale parameter ω^2 and the shape parameter λ. The shape parameter λ governs the skewness of this distribution. As $\lambda = 0$ the *Skew-Normal* becomes symmetric and collapses to the *Normal*. Positive values of λ are associated with a right skewed distribution while negative values of λ are associated with a left skewed distribution. [3] To model time variation in the shape of the shocks, I treat the shape parameter λ as an additional stochastic process in the model:

$$y_t = \boldsymbol{x_t}\boldsymbol{\pi} + \sqrt{h_t}\varepsilon_t \qquad \varepsilon_t \sim Skew\text{-}Normal(\zeta_t, \omega_t^2, \lambda_t) \tag{1}$$

3. See Appendix A.1 for details on the *Skew-Normal and Skew-t*.

$$log(h_t) = \phi_h log(h_{t-1}) + \eta_t \qquad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2) \tag{2}$$

$$\lambda_t = \phi_\lambda \lambda_{t-1} + \xi_t \qquad \xi_t \sim \mathcal{N}(0, \sigma_\xi^2) \tag{3}$$

where y_t is the dependent variable observed over the periods $t = 1, \ldots, T$, while $\boldsymbol{x_t}$ is a row vector of that might contain lags of the dependent variable and other exogenous regressors and $\boldsymbol{\pi}$ is the column vector of coefficients. I assume that the *Skew-Normal* shocks have zero mean and unit variance, that is $\mathbb{E}[\varepsilon_t] = 0$ and $var(\varepsilon_t) = 1$, which implies the following constraints on the location and scale parameters:

$$\zeta_t = -\omega_t \delta_t \sqrt{\frac{2}{\pi}} \qquad \vee\, t \tag{4}$$

$$\omega_t^2 = \left[1 - \frac{2}{\pi}\delta_t^2\right]^{-1} \qquad \vee\, t \tag{5}$$

where $\delta_t = \frac{\lambda_t}{\sqrt{1+\lambda_t^2}}$, with $-1 < \delta_t < 1$. This parametrization ensures that $\mathbb{E}[y_t|\mathcal{I}_{t-1}] = \boldsymbol{x_t}\boldsymbol{\pi}$. In this regard, it is important to remark that imposing $\zeta_t = 0$ instead of (4) would imply $\mathbb{E}[\varepsilon_t] \neq 0$, and in general $\mathbb{E}[\varepsilon_t|\mathcal{I}_{t-1}] \neq 0$.[4] As well, this parametrization ensures that y_t features both time varying conditional volatility and time varying conditional skewness with the former exclusively governed by the stochastic process in equation (2) while the latter by the stochastic process in (3).[5]

In order to explicitly model heavy-tails, together with time-varying skewness, I also consider an alternative specification where the innovations are distributed as a $Skew-t(\zeta_t, \omega_t^2, \lambda_t, \nu)$ (Azzalini et al. 2003). The parameter of the degrees of freedom ν determines the tail thickness of the *Skew-t* distribution: as $\nu \to \infty$ the *Skew-t* converges to the *Skew-Normal* while when $\lambda = 0$ the *Skew-t* collapses to a *Student-t* with ν degrees of freedom. In this case the constraints

4. Imposing $\zeta_t = 0$ instead of (4) leads to a model with a time varying intercept, shifting the conditional mean of y_t proportionally to λ_{t-1}.

5. It is possible to have a model that features both time varying volatility and time varying skewness by assuming:

$$\begin{aligned} y_t &= \boldsymbol{x_t}\boldsymbol{\pi} + \varepsilon_t & \varepsilon_t &\sim \textit{Skew-Normal}(\zeta_t, \omega_t^2, \lambda_t) \\ \lambda_t &= \phi_\lambda \lambda_{t-1} + \xi_t & \xi_t &\sim \mathcal{N}(0, \sigma_\xi^2) \end{aligned}$$

assuming $\mathbb{E}(\varepsilon_t) = 0$ (hence (4) still holds) and imposing $\omega^2 = 1$ which implies $var(\varepsilon_t) \neq 1 = \left(1 - \frac{2\delta_t^2}{\pi}\right)$. However, in this case the parameter λ_t would drive both conditional skewness and conditional volatility. This is not desirable in general, since we might want to model these two distinct features using different dynamics.

on the location and scale parameters that ensure $\mathbb{E}[\varepsilon_t] = 0$ and $var(\varepsilon_t) = 1$ become:

$$\zeta_t = -\omega_t \delta_t k_1 \sqrt{\frac{2}{\pi}} \qquad \vee t \tag{6}$$

$$\omega_t^2 = \left(k_2 - \frac{2}{\pi} k_1^2 \delta_t^2 \right)^{-1} \qquad \vee t \tag{7}$$

where $k_1 = \sqrt{\frac{\nu}{2}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})}$, $k_2 = \frac{\nu}{\nu-2}$ and $\Gamma(.)$ is the Gamma function. This stochastic volatility model with *Skew-t* shocks includes as special cases both the stochastic volatility model with heavy tails without conditional skewness of Jacquier et al. (2004) and the model with heavy tails and constant conditional skewness of Abanto-Valle et al. (2015).[6] It is straightforward to modify this specification by assuming a different dynamics for the log-volatility and the shape parameter in the state equations (2) and (3). For example if we suspect that some of the variables in $\boldsymbol{x_t}$ affect not only the conditional mean, but also the conditional variance and the conditional skewness of y_t, we can include them in the state equations of these two distinct stochastic processes. For example, as it will be shown in the application to the Growth at Risk framework in Section 3, motivated by the findings of Adrian et al. (2019) and subsequent work by Delle Monache et al. (2021), Montes-Galdón et al. (2022) and Wolf (2021) I consider a specification in which financial condition affect not only the conditional mean but also the conditional skewness of the future GDP growth distribution.

2.1.1 Priors and estimation of the univariate TVSSV model

This section develops a posterior simulation sampler which allows for Bayesian estimation of the univariate models presented above. For what concerns the specification of the prior distribution for the parameters of the model, I assume a *Normal* prior for the regression coefficients ($\boldsymbol{\pi}$) and for the coefficients in the state equations (ϕ_λ and ϕ_h) while I specify an Inverse Gamma Prior for the variances of the innovations to the log-volatility and to the shape parameter (σ_η^2 and σ_ξ^2). The estimation strategy leverages on the fact that $\varepsilon_t \sim$ *Skew Normal*$(\zeta_t, \omega_t^2, \lambda_t)$ has the

6. The stochastic volatility model with heavy tails of Jacquier et al. (2004) is a particular version of this model where the shape parameter is constant and equal to 0, that is $\lambda_t = 0 \quad \vee t$. As well, the stochastic volatility model with skewness and heavy tails of is a particular version of this model where $\sigma_\xi^2 \to 0$ and $\phi_\lambda = 1$, namely the shape parameter λ_t is constant.

following stochastic representation :

$$\varepsilon_t = \zeta_t + \delta_t \omega_t v_t + \sqrt{(1-\delta_t^2)}\omega_t z_t \tag{8}$$

where $v_t \overset{i.i.d}{\sim}$ Truncated Normal$_{[0,\infty)}(0,1)$ and $z_t \overset{i.i.d}{\sim} \mathcal{N}(0,1)$. Equation (8) implies that conditioning on the mixing variable v_t and on δ_t, which is one to one map to λ_t, the random variable ε_t is distributed as a *Normal* . This result greatly simplifies the derivation of the full conditional distributions in the *Gibbs Sampler* and allows to exploit and adapt many of the results used for the estimation of the standard stochastic volatility model with Gaussian innovations (Jacquier et al. 1994). In particular, in the model with *Skew-Normal* shocks, once I have obtained a draw from the full conditional posterior distribution of the mixing variable v_t and from the full conditional distribution of the shape parameter λ_t, I can exploit the conditionally *Normal* distribution of ε_t in the derivation of formulas of the conditional distributions of the other parameters and the latent states of the model. Moreover ζ_t, ω_t and δ_t are neither parameters nor latent states to be estimated. ζ_t and ω_t satisfy the constraints (4) and (5) and ensure the correct parameterization of the shocks at each time period $t = 1, \dots, T$, while δ_t is a one to one map to λ_t, namely $\delta_t = \frac{\lambda_t}{\sqrt{1+\lambda_t^2}}$.

Table 1 presents the details on the *Gibbs Sampler* while Appendix A.2 reports the derivations of the full conditional posterior distributions. In Step 1) I sample the mixing variables $\{v_t\}_{t=1}^T$ from the full conditional posterior distribution $p(v_t|\boldsymbol{\Theta}, \boldsymbol{\lambda}, \boldsymbol{h}, \boldsymbol{y})$ which is a *Truncated Normal* distribution. Steps 2) 3) 4) 5) 6) are pretty standard: I draw the regression coefficients $\boldsymbol{\pi}$ in the observation equation (1) and the autoregressive coefficients and the variances in the two state equations (2) (3) from their respective full conditional posterior distributions. In Step 7) and Step 9) I draw the initial states for the volatility h_0 and the shape parameter λ_0, while in Steps 8) and 10) I draw the entire history for the volatilities and the shape parameters. Since it is not feasible to directly sample from the full conditional distributions of the volatilities $p(h_1, \dots, h_T|\boldsymbol{\Theta}, \boldsymbol{v}, \boldsymbol{\lambda}, \boldsymbol{y})$ and the shape parameters $p(\lambda_{i1}, \dots, \lambda_{iT}|\boldsymbol{\Theta}, \boldsymbol{v}, \boldsymbol{h}, \boldsymbol{y})$ I rely on the particle filter to approximate these distributions. In alternative to the particle step, to draw both the log-volatilities and the shape parameters it is possible to consider an independence Metropolis Hastings step but I experienced that the algorithm based on the particle filter has smaller

mixing times.[7] In the particle approximation, I use the transition equations (2) and (3) as importance densities and compute the weights accordingly. The details on the particle steps used to approximate the full conditional posterior distribution of the volatilities and the shape parameters can be find in Table 4 in the Appendix A.4. As well, in the Appendix A.4, I report the details on the steps of the alternative algorithm which relies on the independence Metropolis Hastings steps to draw the volatilities and the shape parameters.

Table 1: MCMC algorithm for the univariate TVSSV model

MCMC for the univariate TVSSV model
Initialize $\boldsymbol{\Theta}^{(0)}, \boldsymbol{s}^{(0)}$
For $m = 0$: Total MCMC draws
1) Draw $\{v_t\}_{t=1}^{T^{(m+1)}}$ from $p(v_1 \dots, v_T \vert \boldsymbol{\Theta}^{(m)}, \boldsymbol{\lambda}^{(m)}), \boldsymbol{h}^{(m)}, \boldsymbol{y})$
2) Draw $\boldsymbol{\pi}^{(m+1)}$ from $p(\boldsymbol{\pi} \vert \boldsymbol{\Theta}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{\lambda}^{(m)}), \boldsymbol{h}^{(m)}, \boldsymbol{y})$
3) Draw $\sigma_\eta^{2(m+1)}$ from $p(\sigma_\eta^2 \vert \boldsymbol{\Theta^{(m)}}, \boldsymbol{v^{(m)}}, \boldsymbol{\lambda}^{(m)}), \boldsymbol{h}^{(m)}, \boldsymbol{y})$
4) Draw $\sigma_\xi^{2(m+1)}$ from $p(\sigma_\xi^2 \vert \boldsymbol{\Theta^{(m)}}, \boldsymbol{v^{(m)}}, \boldsymbol{\lambda}^{(m)}), \boldsymbol{h}^{(m)}, \boldsymbol{y})$
5) Draw $\phi_h^{(m+1)}$ from $p(\phi_h \vert \boldsymbol{\Theta^{(m)}}, \boldsymbol{v^{(m)}}, \boldsymbol{\lambda}^{(m)}), \boldsymbol{h}^{(m)}, \boldsymbol{y})$
6) Draw $\phi_\lambda^{(m+1)}$ from $p(\phi_\lambda \vert \boldsymbol{\Theta^{(m)}}, \boldsymbol{v^{(m)}}, \boldsymbol{\lambda}^{(m)}), \boldsymbol{h}^{(m)}, \boldsymbol{y})$
7) Draw $h_0^{(m+1)}$ from $p(h_0 \vert \boldsymbol{\Theta}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{\lambda}^{(m)}), \boldsymbol{h}^{(m)}, \boldsymbol{y})$
8) Draw $\{h_t\}_{t=1}^{T^{(m+1)}}$ from $p(h_1, \dots, h_T \vert \boldsymbol{\Theta^{(m)}}, \boldsymbol{v^{(m)}}, \boldsymbol{\lambda}^{(m)}, \boldsymbol{y})$
Particle Step
9) Draw $\lambda_0^{(m+1)}$ from $\lambda_0^{(m+1)}$ from $p(\lambda_0 \vert \boldsymbol{\Theta}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{\lambda}^{(m)}, \boldsymbol{h}^{(m)}, \boldsymbol{y})$
10) Draw $\{\lambda_t\}_{t=1}^{T^{(m+1)}}$ from $p(\lambda_{i1}, \dots, \lambda_{iT} \vert \boldsymbol{\Theta^{(m)}}, \boldsymbol{v^{(m)}}), \boldsymbol{h}^{(m)}, \boldsymbol{y})$
Particle Step
end

To estimate the version of the model with *Skew-t* innovations, I just exploit the fact that $\varepsilon_t \sim$ *Skew-t*$(\zeta_t, \omega_t^2, \lambda_t, \nu)$ has in turn a convenient stochastic representation, namely:

$$\varepsilon_t = \zeta_t + \delta_t \omega_t o_t^{-0.5} v_t + \sqrt{(1 - \delta_t^2)} \omega_t o_t^{-0.5} z_t \tag{9}$$

$v_t \overset{i.i.d}{\sim}$ Truncated Normal$_{[0,\infty)}(0,1)$, $z_t \overset{i.i.d}{\sim} \mathcal{N}(0,1)$ and $o_t \overset{i.i.d}{\sim} \mathcal{G}(\frac{\nu}{2}, \frac{\nu}{2})$

This is the same same representation of the *Skew-Normal* except for the additional mixing variable o_t. Therefore, conditioning on both the two mixing variables $m_t = \{v_t, o_t\}$ and on δ_t, which is a one to one map with λ_t, the shock ε_t is distributed as a *Normal*. Therefore, also in this case, I can exploit and adapt the derivations of the standard model with Gaussian shocks when

7. In the particle steps, in order to alleviate path degeneracy, I exploit the *Ancestor Sampling* procedure developed in Lindsten et al. (2014) which enables fast mixing even when using seemingly few particles. Lindsten et al. (2014) study the properties of the sampler and provide the formal proof for the convergence of the algorithm.

deriving the full conditional posterior distribution in the *Gibbs Sampler*. In order to estimate the model it is just needed to consider a further initial step to draw from $p(o_1 \dots, o_T|\boldsymbol{\Theta}, \boldsymbol{v}, \boldsymbol{\lambda}, \boldsymbol{h}, \boldsymbol{y})$, namely: [8]

$$\text{Draw } \{o_t\}_{t=1}^{T^{(m+1)}} \text{ from } p(o_1 \dots, o_T|\boldsymbol{\Theta}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{\lambda}^{(m)}), \boldsymbol{h}^{(m)}, \boldsymbol{y})$$

and then adapt Steps 2) to 10) in Table (1) with the new formulas of the full conditional distributions derived by conditioning on the further mixing variables $\{o_t\}_{t=1}^{T}$. In this case, since it is not possible to directly sample from the full conditional distribution of the mixing variable o_t, I use Metropolis Hastings to simulate draws from this distribution. Appendix A.3 reports the details of this step.

2.2 Time varying skewness stochastic volatility VAR model

Given the risk management nature of the problem of policymaking, it is often the case that the objective of interest is to quantify and predict tail risk to multiple macroeconomic outcomes (Kilian et al. 2003). In particular, from a modelling perspective, we might be interested in a multivariate model that can characterize asymmetries in the future distribution of multiple macroeconomic time-series. VAR models (Sims 1980) emerged as the natural tool to capture the rich dynamic interrelationship between multiple macroeconomic time series. They currently represent the workhouse in empirical macroeconomics and are routinely used for forecasting and policy analyses (Stock et al. 2001). In this section I exploit the conceptual framework presented in the previous section to jointly model the dynamic behaviour of multiple time series in a Bayesian VAR model and capture time varying skewness in the conditional distribution of the variables in the system. The model is given by:

$$\boldsymbol{y_t} = \boldsymbol{\Pi_0} + \boldsymbol{\Pi_1 y_{t-1}} + \dots + \boldsymbol{\Pi_p y_{t-p}} + \boldsymbol{A}^{-1}\boldsymbol{H}_t^{0.5}\boldsymbol{\varepsilon_t} \tag{10}$$

where $\boldsymbol{y}_t$ is an $N \times 1$ vector of variables observed over the periods $t = 1, \dots, T$. $\boldsymbol{H_t}$ is a diagonal matrix that contains the volatilities on its main diagonal, namely $\boldsymbol{H_t} = diag(h_{1,t} \dots, h_{N,t})$ and $\boldsymbol{A}^{-1}$ is a lower triangular matrix with ones on its main diagonal. The log-volatilities evolve

8. In the estimation of the model with heavy tails (*Skew-t* shocks), I fix the tail thickness parameters ν to 5. Given the relative short time series length of macroeconomic data, it is particularly difficult to make inference on this parameter. In general, you can draw this parameter adding another Metropolis Hastings step to draw from $p(\nu|\boldsymbol{\Theta}, \boldsymbol{v}, \boldsymbol{o}, \boldsymbol{\lambda}, \boldsymbol{y})$

over time according to:

$$log(h_{i,t}) = \phi_{h,i} log(h_{i,t-1}) + \eta_{i,t} \qquad \eta_{i,t} \sim N(0, \sigma^2_{\eta,i}) \tag{11}$$

for $i = 1, \ldots, N$. In the Gaussian stochastic volatility model of Cogley et al. (2005) and Primiceri (2005) it is assumed $\boldsymbol{\varepsilon_t} \sim N(0, \boldsymbol{I})$. In our specification, $\boldsymbol{\varepsilon_t}$ is a vector of *Skew-Normal* shocks, namely:

$$\boldsymbol{\varepsilon_t} = [\varepsilon_{1t}, \ldots, \varepsilon_{Nt}]' \qquad \varepsilon_{it} \sim \textit{Skew-Normal}(\zeta_{it}, \omega^2_{it}, \lambda_{it}) \tag{12}$$

where the shape parameters λ_{it} evolve according to:

$$\lambda_{i,t} = \phi_{\lambda,i} \lambda_{i,t-1} + \xi_{i,t} \qquad \xi_{i,t} \sim N(0, \sigma^2_{\xi,i}) \tag{13}$$

In order to have $\mathbb{E}[\boldsymbol{\varepsilon_t}] = \mathbf{0}$ and $var(\boldsymbol{\varepsilon_t}) = \boldsymbol{I}$ the shocks are parameterized imposing the constraints on the location parameters ζ_{it} and on the scale parameters ω_{it} discussed in the previous section. As in the univariate framework, I can explicitly model heavy-tails, together with time-varying skewness, by considering an alternative specification where:

$$\boldsymbol{\varepsilon_t} = [\varepsilon_{1t}, \ldots, \varepsilon_{Nt}]' \qquad \varepsilon_{it} \sim \textit{Skew-t}(\zeta_{it}, \omega^2_{it}, \lambda_{it}, \nu) \tag{14}$$

The model nests the constant coefficients version of the popular VAR model with stochastic volatility introduced by Cogley et al. (2005) and Primiceri (2005) and considered in Carriero et al. (2019).[9] In these models, as long as the short run restrictions implied by the Cholesky ordering are satisfied, the shocks can be interpreted as structural.[10] This means that, other than for forecasting purposes, the model can be practically used for policy analysis and structural scenario analyses. Also in this multivariate framework it is straightforward to modify the specification of the state equations of the log-volatilities and the shape parameters by assuming a different dynamics in (11) and (13). For example, as it will be shown in the empirical applica-

9. As well, the stochastic volatility VAR with fat tails in Clark et al. (2015) is also a special case of this model with $\lambda_{i,t} = 0 \vee i, t$. Karlsson et al. (2023) stochastic volatility model VAR with *Skew-t* orthogonal residual is as well a particular version of this model with $\phi_{\lambda_i} = 1$ and $\sigma^2_{\xi,i} \to 0 \quad \vee i$.

10. It is worth to mention that due to the "Cholesky type" specification of the stochastic volatility VAR model considered here, the order in which the variables enter in the VAR matters not only for the identification of the shocks but also for the estimation of the model. This is fact was stressed first by Primiceri (2005) and more recently by Arias et al. (2021) and Chan et al. (2021). On the lines of the work of Chan et al. (2021) I am currently working on a order invariant version of the model considered in this paper.

tion in Section 3, I can capture the nonlinear relationship between two variables in the VAR by including the lags of one variable in the state equations of the log-volatility and/or the shape parameter of the shocks to the other variable.

2.2.1 Priors and estimation of the TVS-SV VAR

For what concerns the choice of the prior distributions for the parameters of the model, I assume a Normal prior for the autoregressive coefficients $vec(\mathbf{\Pi})$. As well, following Cogley et al. (2005), I specify a Normal prior for the free elements in the matrix $\boldsymbol{A}$. Finally, as in the univariate framework, I specify independent Inverse Gamma priors for the variance of the innovations to the log-volatilities and to the shape parameters ($\sigma^2_{\eta,i}$ and $\sigma^2_{\xi,i}$) and Normal priors for the coefficients in the state equations ($\phi_{h,i}$ and $\phi_{\lambda,i}$). The estimation strategy for the VAR model is just a generalization of the one for the univariate model that again leverages on the stochastic representation of the *Skew Normal* (8) and *Skew-t* (9) shocks. Exploiting this representation, I can write the vector of *Skew-Normal* shocks $\boldsymbol{\varepsilon_t}$ as follows: [11]

$$\boldsymbol{\varepsilon_t} = \boldsymbol{\zeta_t} + \boldsymbol{\Omega_t}\boldsymbol{\Delta_t}\boldsymbol{v_t} + \boldsymbol{\Omega_t}(\boldsymbol{I}_N - \boldsymbol{\Delta}_{\boldsymbol{t}}^2)^{0.5}\boldsymbol{z_t} \tag{15}$$

where:

$\boldsymbol{\zeta_t} = [\zeta_{1,t}, \dots, \zeta_{N,t}]'$

$\boldsymbol{\Omega_t} = diag(\omega_{1t} \dots \omega_{Nt})$

$\boldsymbol{\Delta_t} = diag(\delta_{1t} \dots \delta_{Nt})$

$\boldsymbol{v_t} = [v_{1,t}, \dots, v_{N,t}]'$ $\qquad v_{i,t} \sim TruncatedNormal_{(0,\infty)}(0,1)$

$\boldsymbol{z_t} = [z_{1,t}, \dots, z_{N,t}]'$ $\qquad z_{it} \sim N(0,1).$

As in the univariate framework, I can exploit this result when deriving the full conditional posterior distributions of the parameters and the unobserved states in the *Gibbs Sampler*. As a matter of fact, also in this case, ζ_{it} and ω_{it} respectively stored in the column vector $\boldsymbol{\zeta_t}$ and in the diagonal matrix $\boldsymbol{\Omega_t}$ are neither parameters nor latent states to be estimated. ζ_{it} and ω_{it} are fixed to satisfy the constraints (4) and (5) and ensure the correct parameterization of the shocks in each equation of the VAR $i = 1, \dots, N$ and at each time period $t = 1, \dots, T$. As well,

11. Note that the powers on the matrices refer all to diagonal matrices. For example $(\boldsymbol{I}_N - \boldsymbol{\Delta}_{\boldsymbol{t}}^2) = diag(\sqrt{1-\delta_{1,t}^2}, \dots, \sqrt{1-\delta_{N,t}^2})$ or afterwords $\boldsymbol{O_t}^{-0.5} = diag\left(\frac{1}{\sqrt{o_{1,t}}}, \dots, \frac{1}{\sqrt{o_{N,t}}}\right)$

the elements in the diagonal matrix $\boldsymbol{\Delta_t}$ (that is δ_{it}) are one to one map of the latent states λ_{it}.

Table (2) presents the details of the sampler. In Step 1) I draw the mixing variables $\{v_{it}\}_{t=1}^{T}$ for $i = 1, \ldots, N$. In Step 2) I draw the coefficients of the VAR coefficients adapting to my framework the correct version of the triangular algorithm developed in Carriero et al. (2019) and corrected in Carriero et al. (2022). This approach allows to reduce the computational burden associated to the system-wide estimation of Bayesian VAR with stochastic volatility and non-conjugate priors by exploiting a triangularization of the system. In Step 3), I adapt the approach of Cogley et al. (2005) to draw the free elements in the matrix $\boldsymbol{A}$. In Step 4) 5) and 6) 7) I draw the variances and the autoregressive coefficients of the state equations while in Step 8) and 10) I draw the initial state for the volatilities h_{i0} and the shape parameters λ_{i0} . In Step 9) and 11) I draw the entire path for the volatilities and the shape parameters, using the *Particle Step with Ancestor Sampling* described in Table 4 in the Appendix A.4.

Table 2: MCMC algorithm for the TVSSV VAR model

Particle Gibbs Sampler for the TVSSV-VAR model	
Initialize $\boldsymbol{\Theta}^{(0)}, \boldsymbol{s}^{(0)}, \boldsymbol{v}^{(0)}$	
For $m = 0$: Total MCMC draws	
1) Draw $\{\boldsymbol{v_{it}}\}_{t=1}^{T^{(m+1)}}$ from $p(\boldsymbol{v_{i1}} \ldots, v_{iT} \vert \boldsymbol{\Theta}^{(m)}, \boldsymbol{s}^{(m)}, \boldsymbol{Y})$	$i = 1, \ldots, N$
2) Draw $\boldsymbol{\Pi}^{(m+1)}$ from $p(\boldsymbol{\Pi} \vert \boldsymbol{\Theta}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{s}^{(m)}, \boldsymbol{Y})$	
3) Draw $\boldsymbol{A}^{(m+1)}$ from $p(\boldsymbol{A} \vert \boldsymbol{\Theta}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{s}^{(m)}, \boldsymbol{Y})$	
4) Draw $\sigma_{\xi,i}^{2(m+1)}$ from $p(\sigma_{\xi,i}^{2} \vert \boldsymbol{\Theta^{(m)}}, \boldsymbol{s^{(m)}}, \boldsymbol{v^{(m)}}, \boldsymbol{Y})$	$i = 1, \ldots, N$
5) Draw $\sigma_{\eta,i}^{2(m+1)}$ from $p(\sigma_{\eta,i}^{2} \vert \boldsymbol{\Theta^{(m)}}, \boldsymbol{s^{(m)}}, \boldsymbol{v^{(m)}}, \boldsymbol{Y})$	$i = 1, \ldots, N$
6) Draw $\phi_{h,i}^{(m+1)}$ from $p(\phi_{h,i} \vert \boldsymbol{\Theta^{(m)}}, \boldsymbol{s^{(m)}}, \boldsymbol{v^{(m)}}, \boldsymbol{Y})$	$i = 1, \ldots, N$
7) Draw $\phi_{\lambda,i}^{(m+1)}$ from $p(\phi_{\lambda,i} \vert \boldsymbol{\Theta^{(m)}}, \boldsymbol{s^{(m)}}, \boldsymbol{v^{(m)}}, \boldsymbol{Y})$	$i = 1, \ldots, N$
8) Draw $h_{i,0}^{(m+1)}$ from $p(h_{i,0} \vert \boldsymbol{\Theta}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{s}^{(m)}, \boldsymbol{Y})$	$i = 1, \ldots, N$
9) Draw $\{h_{it}\}_{t=1}^{T^{(m+1)}}$ from $p(h_{i1}, \ldots, h_{iT} \vert \boldsymbol{\Theta^{(m)}}, \boldsymbol{v^{(m)}}, \boldsymbol{s}^{(m)}, \boldsymbol{Y})$	$i = 1, \ldots, N$
Particle step	
10) Draw $\lambda_{i,0}^{(m+1)}$ from $\lambda_{i,0}^{(m+1)}$ from $p(\lambda_{i,0} \vert \boldsymbol{\Theta}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{s}^{(m)}, \boldsymbol{Y})$	$i = 1, \ldots, N$
11) Draw $\{\lambda_{it}\}_{t=1}^{T^{(m+1)}}$ from $p(\lambda_{i1}, \ldots, \lambda_{iT} \vert \boldsymbol{\Theta^{(m)}}, \boldsymbol{v^{(m)}}, \boldsymbol{s}^{(m)}, \boldsymbol{Y})$	$i = 1, \ldots, N$
Particle step	
end	

As in the univariate framework, it is easy to adapt the sampler to a version of the VAR model with *Skew-t* shocks. In this case (9) becomes:

$$\boldsymbol{\varepsilon_t} = \boldsymbol{\zeta_t} + \boldsymbol{\Omega_t}\boldsymbol{\Delta_t}\boldsymbol{O_t}^{-0.5} + \boldsymbol{\Omega_t}(\boldsymbol{I}_n - \boldsymbol{\Delta_t}^2)^{0.5}\boldsymbol{O_t}^{-0.5}\boldsymbol{z_t} \tag{16}$$

where $\boldsymbol{O_t} = diag(o_{1t} \ldots o_{Nt})$ $\qquad$ $o_{it} \sim Gamma(\frac{\nu}{2}, \frac{\nu}{2})$.

It is enough to adapt the Gibbs Sampler by adding another initial step to draw the mixing variables $\{o_{it}\}_{t=1}^{T}$ for $i = 1, \ldots, N$

$$\text{Draw } \{o_{it}\}_{t=1}^{T^{(m+1)}} \text{ from } p(o_{i1} \ldots, o_{iT} | \boldsymbol{\Theta}^{(m)}, \boldsymbol{v}^{(m)}, \boldsymbol{s_t}^{(m)}, \boldsymbol{Y}) \quad i = 1, \ldots, N$$

and then to update the formulas of the full conditional posterior distributions in order to account for the extra terms. Again I use Metropolis Hastings to simulate draws from $p(o_{i1} \ldots, o_{iT} | \boldsymbol{\Theta}, \boldsymbol{v}, \boldsymbol{s_t})$ for $i = 1, \ldots, N$, since it is not directly possible to sample from these distributions.

3 Growth at Risk

The work of Adrian et al. (2019) (henceforth ABG) pioneered a recently growing body of research, which examines the main sources of tail risk to GDP growth in relationship to changes in economic and financial conditions. This section compares the out of sample Growth-at-Risk (GaR) estimates for the U.S from our time varying skewness stochastic volatility models to the two step approach based on quantile regression of ABG . In order to model asymmetric changes in the conditional distribution of GDP growth as a function of changes in financial conditions, I consider the following specification of the univariate TVSSV model:

$$\begin{aligned} gdpgrowth_t = \pi_0 + \pi_1 gdpgrowth_{t-1} + \pi_2 gdpgrowth_{t-2} + \pi_3 NFCI_{t-1} + \sqrt{h_t}\varepsilon_t \\ \varepsilon_t \sim Skew - Normal(\zeta_t, \omega_t, \lambda_t) \\ \text{or} \\ \varepsilon_t \sim Skew - t(\zeta_t, \omega_t, \lambda_t, \nu) \end{aligned} \tag{17}$$

$$log(h_t) = \phi_h log(h_{t-1}) + \eta_t \qquad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2) \tag{18}$$

$$\lambda_t = \phi_\lambda \lambda_{t-1} + \beta_1 NFCI_{t-1} + \xi_t \qquad \xi_t \sim \mathcal{N}(0, \sigma_\xi^2) \tag{19}$$

In this specification the NFCI directly affects the conditional skewness of the future GDP growth distribution. More specifically, the coefficient β_1 captures changes in the skewness of the conditional distribution of GDP growth as a function of financial conditions. This coefficient is meant to capture the non-linear relationship between deteriorating financial conditions and future GDP growth distribution found in ABG. Since our focus is to model the asymmetric effect of the NFCI on the future GDP growth distribution, I threat the log-volatilities as exogenous

autoregressive processes, not affected by the NFCI. As a matter of fact, augmenting the state equation for the log-volatilities with the NFCI index, as it is done in the state equations of the shape parameters, implies that financial conditions would affect symmetrically both a upper and the lower quantiles of the future GDP growth distribution. Together with the univariate model, I consider as well a bivariate TVSSV-VAR(2) model where $\boldsymbol{y_t} = [gdpgrowth, NFCI]'$ and:

$$\begin{aligned} \boldsymbol{y_t} = \boldsymbol{\Pi_0} + \boldsymbol{\Pi_1 y_{t-1}} + \boldsymbol{\Pi_2 y_{t-2}} + \boldsymbol{A}^{-1}\boldsymbol{H_t^{0.5}}\boldsymbol{\varepsilon_t} \\ \varepsilon_{it} \sim \textit{Skew-Normal}(\zeta_{it}, \omega_{it}^2, \lambda_{it}) \\ \text{or} \\ \varepsilon_{it} \sim Skew - t(\zeta_{it}, \omega_{it}, \lambda_{it}, \nu) \end{aligned} \tag{20}$$

$$log(h_{it}) = \phi_{h,i} log(h_{it-1}) + \eta_{it} \qquad \eta_{it} \sim \mathcal{N}(0, \sigma_{i,\eta}^2) \;\; i = gdpgrowth,\; NFCI \tag{21}$$

$$\lambda_{gdpgrowth,t} = \phi_{\lambda,1}\lambda_{gdpgrowth,t-1} + \beta_1 NFCI_{t-1} + \xi_{gdpgrowth,t} \qquad \xi_{i,t} \sim N(0, \sigma_{\xi,i}^2) \tag{22}$$

$$\lambda_{NFCI,t} = \phi_{\lambda,2}\lambda_{NFCI,t-1} + \xi_{NFCI,t} \qquad \xi_{i,t} \sim N(0, \sigma_{\xi,i}^2) \tag{23}$$

In this VAR, the dynamic relationship between GDP growth and financial conditions in modelled jointly. In particular, in this specification, due to the triangular structure of $\boldsymbol{A}^{-1}$ shocks to GDP growth contemporaneously affect the financial markets, while shocks to NFCI do not affect GDP growth within the quarter. To understand whether the models perform well in forecasting downside risk, in what follows I will compare the forecast from the TVSSV models to the forecasts from the quantile regression based method of ABG. Their approach is based on a two step procedure where in the first step they use predictive quantile regression to estimate the quantiles of the conditional distribution:

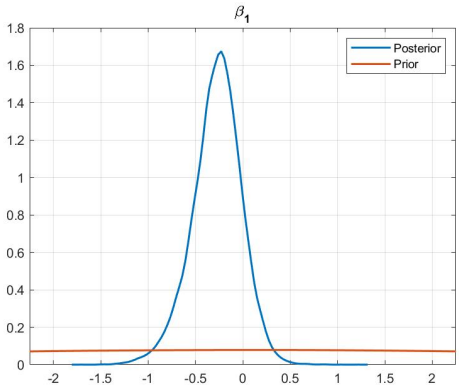
$$\hat{Q}_{gdpgrowth_{t+h}|\mathcal{I}_t}(\tau) = \hat{\boldsymbol{\beta}}^{\tau} X_t \qquad for \;\; \tau = 0.05, \ldots, 0.95 \tag{24}$$

Then, in the second step, the estimated quantiles are interpolated using a flexible *Skew-t* distribution, so as to obtain a complete predictive density for GDP growth. We specify equation (24) collecting two lags of GDP growth and one lag of NFCI in the vector X_t, so as to capture changes in the future GDP growth distribution as a function of current financial and economic conditions.

3.1 Results

This section presents the results from the estimates of both the univariate TVSSV models and the VAR TVSSV models with *Skew-Normal* and *Skew-t* shocks. The estimation sample starts in 1971Q1 and the forecasting exercise covers the period 1995Q1 - 2019Q4. Fig. 1 presents the estimated posterior distribution for the coefficient β_1 from the univariate time varying skewness stochastic volatility model. This is the coefficient that in the state equation of the skewness parameter (22) summarizes how the shape of the conditional distribution of GDP growth changes as a function of financial conditions in the previous quarter. As shown in Fig. 1, tighter financial conditions (increases in the $NFCI$) are on average associated to a decrease in the skewness of current GDP growth (the posterior mean estimate is $\hat{\beta}_1 = -0.26$). Hence, equation (19), captures the main finding of ABG, which is that deteriorating financial condition are associated to movements in the lower quantiles of future GDP growth distribution.

Figure 1: Posterior estimate of β_1

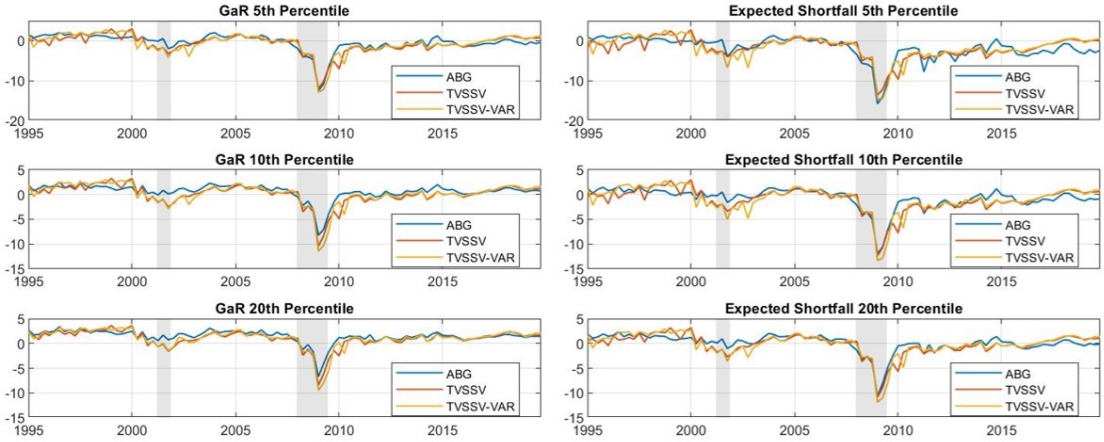


Note: The figure shows the estimated posterior distribution of the coefficient on the $NFCI_{t-1}$ in the equation of the shape parameter in the TVSSV model with *Skew-t* shocks.

Ascertained that the model is able to capture the same asymmetric effect of financial conditions on the future GDP growth distribution found in ABG, it is important to understand what is the potential of the model to assess and predict risk out of sample. Fig. 2 shows the out-of-sample forecasts of Growth at Risk and Expected Shortfall for the 5^{th}, 10^{th} and 20^{th} percentiles while Fig. 3 shows the one quarter ahead estimated recession probability. I report the results from the stochastic volatility stochastic skewness model with *Skew-t* shocks, since the results from the model with *Skew-Normal* shocks do not differ qualitatively. The figure shows

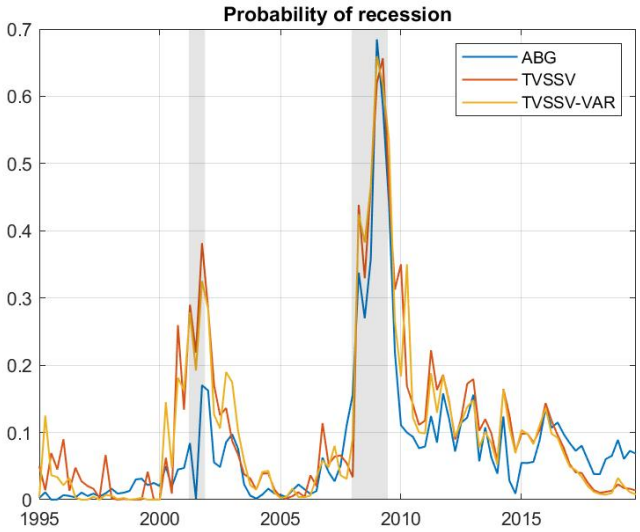
that during the Financial Crisis our parametric models predicts as much downside risk to GDP growth as the quantile regression method of ABG. As shown in Fig. 3 both the TVSSV and the TVSSV-VAR models, assign higher probability of recession to the mild contraction of the U.S. economy following the dotcom bubble in 2000s with respect to the two step method based on quantile regression.

Figure 2: One quarter ahead Growth at Risk (GaR) and Expected Shortfall (1995Q1-2019Q4)



Note: The figure shows the estimated 5^{th}, 10^{th}, 20^{th} percentiles of the one quarter ahead GDP growth predictive distribution (left panel) and the 5^{th}, 10^{th}, 20^{th} one quarter ahead expected shortfall (right panel). In blue estimates from the two step quantile regression based method by ABG, in red from the TVSSV univariate model with *Skew-t* shocks and in yellow the estimates from the TVSSV VAR model.

Figure 3: One quarter ahead recession probability (1995Q1-2019Q4)



Note: The figure shows the estimated one quarter ahead recession probabilities. In blue estimates from the two step quantile regression based method by ABG, in red from the TVSSV univariate model with *Skew-t* shocks and in yellow the estimates from the TVSSV VAR model.

In terms of forecast accuracy, Table 3 compares the forecasts from our parametric approaches to the forecasts from the method of ABG. The first two columns report the results for the average Log Scores and the average Cumulative Ranked Probability Scores (CRPS), since these two measures are the most commonly used to evaluate the relative density forecast accuracy of different models. [12] Looking at average Logscores, the first column reports the difference between the forecasts from two step procedure of ABG and the forecasts from the time varying skewness stochastic volatility models (values greater than zero are associated to more accurate density forecast w.r.t ABG). According to the average Log-scores, our parametric models provide more accurate one quarter ahead density forecasts with respect to ABG. In parenthesis I report the *p-values* from the Diebold and Mariano test (Diebold et al. 1995) of equal forecast accuracy

12. Defining y the realization of the series to predict, $f(.)$ the density forecast and $F(.)$ corresponding the cumulative distribution, Logscores and CRPS are respectively defined as:

$$Logscores(f, y) = -log(f(y)) \tag{25}$$

$$CRPS(f, y) = \int_{-\infty}^{\infty} PS(F(z), \mathbb{1}\{y \leq z\})dz = \int_{0}^{1} QS_{\alpha}(F^{-1}(\alpha), y)d\alpha \tag{26}$$

where $PS(F(z), \mathbb{1}\{y \leq z\}) = (F(z) - \mathbb{1}\{y \leq z\})^2$ is the Brier probability score and $QS_{\alpha}(F^{-1}(\alpha), y) = 2(\mathbb{1}\{y \leq F^{-1}(\alpha)\} - \alpha)(F^{-1}(\alpha) - y)$ is the Quantile Score.

and find that for the TVSSV with *Skew-t* shocks I am able to reject the null hypothesis of equal forecast accuracy. For what concerns average CRPS, on the second column, the table reports the ratio with respect to the model of ABG (values lower than 1 are associated to more accurate density forecast with respect to ABG). As you can notice, based on this metrics, the time varying skewness stochastic volatility models perform as good if not even better than the two step procedure based on quantile regression. However, in all the cases I am not able to reject the null of equal forecast accuracy.

Since I aim to assess the ability of the model to correctly characterize downside risk predictions, on the third column I report the average Quantile Weighted CRPS introduced by Gneiting et al. (2011) [13] and on the fourth, fifth and sixth column I report the average Quantile Scores for the 5^{th}, 10^{th} and 20^{th} percentiles commonly associated with the tick loss function (Giacomini et al. 2005). Also in this case I report the ratio with respect to the two step approach based on quantile regression (values lower than 1 are associated to more accurate density forecast with respect to ABG) and the *p-values* from the Diebold-Mariano test in parenthesis. As you can notice, in terms of the ability of the model to correctly characterize downside risk predictions, I find that the stochastic volatility models performs comparably if not even better than ABG. In particular for the TVSSV-VAR with *Skew-t* shocks I am able to reject the null of equal forecast accuracy with respect to ABG. The time series with the CRPS and left Tail Weighted CRPS, can be found in the Appendix B.3 (Fig. 6). As well, in the Appendix B.3 the histogram with the PITs (Fig. 7) reveals that the forecasts from the TVSSV models, are better-calibrated with respect to the forecasts from the two-step quantile regression based method. Summing up, TVSSV models are able to reproduce the main finding in ABG, namely that deteriorating financial conditions are associated to shifts of the lower quantiles of the future GDP growth distribution. At the same time TVSSV models perform comparably if not even better than quantile regression based methods for forecasting macroeconomic tail risk.

13. The Quantile Weighted CRPS are computed as:

$$twCRPS = \int_{-\infty}^{\infty} PS(F(z), \mathbb{1}\{y \leq z\})^2 w(z) dz = \int_0^1 QS_\alpha(F^{-1}(\alpha), y) v(\alpha) d\alpha \tag{27}$$

where $v(\alpha) = (1-\alpha)^2$ assigns higher weights to the lower quantiles of the distribution function.

Table 3: One quarter ahead out of sample forecasts (1995Q1-2019Q4)

	Log scores	**CRPS**	**TwL CRPS**	$QS5^{th}$	$QS10^{th}$	$QS20^{th}$
ABG	2.4840	1.1943	0.3623	0.2503	0.3926	0.5842
TVSSV ***Skew Normal***	0.1946	0.9757	0.9777	1.0334	0.9858	0.9832
	(0.1788)	(0.7380)	(0.2282)	(0.6027)	(0.3838)	(0.3278)
TVSSV ***Skew-t***	**0.3530**	0.9609	0.9823	1.0334	0.9911	0.9882
	(0.0276)	(0.8659)	(0.3045)	(0.4725)	(0.4472)	(0.3934)
TVSSV VAR ***Skew Normal***	0.1620	0.9805	0.9644	0.9678	0.9718	0.9666
	(0.2287)	(0.6781)	(0.1669)	(0.4038)	(0.3209)	(0.2589)
TVSSV VAR ***Skew-t***	0.0662	0.9700	**0.9610**	1.0119	0.9979	0.9633
	(0.1050)	(0.5942)	**(0.0366)**	(0.5493)	(0.4778)	(0.1052)

Note: For the average Logscores, the first row reports the values from the ABG method while the other rows report the difference between the two step procedure and the time varying skewness stochastic volatility models. For the other metrics I report the ratio w.r.t the ABG method. Inside the parenthesis *p-values* from the one sided Diebold-Mariano w.r.t the two step method of Adrian et al. (2019). The bold character indicates rejection of equal forecast accuracy at 5%.

4 Time varying skewness in a medium scale VAR

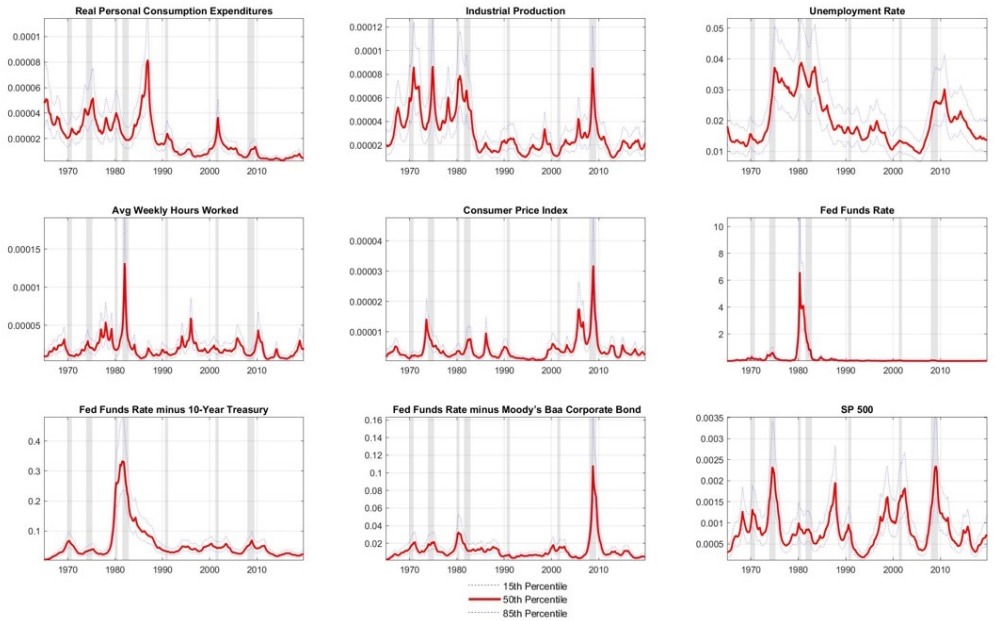
One of the main advantages of the VAR model presented in Section 2.2 is that it allows to explicitly capture time varying conditional skewness of multiple time series. In this section I estimate a medium scale VAR model which includes macroeconomic and financial monthly time series and I investigate the time varying asymmetric behaviour of the shocks to the variables in the system. I consider a VAR model with 8 variables being Real personal consumption expenditures, Industrial Production, Unemployment Rate, average Weekly Hours Worked, Consumer Price Index, Fed Funds Rate, the spread between 10-Year Treasury and the Fed Funds Rate, the spread between Moody's Baa Corporate Bond and the Fed Funds Rate and the Standard and Poors Index. The variables are in monthly frequency and are taken from the FRED-MD.[14] I present the results from the VAR with *Skew-t* shocks.[15] I include 13 lags and assume a Minnesota prior structure for the variance covariance matrix of the regression coefficients. [16] The estimation sample is January 1965 - December 2019. Fig. 4 shows the estimated volatilities while Fig. 5 shows the estimated shape parameters. The dotted line in blue are the $85^{th} - 15^{th}$ credible sets while the red line is the estimated posterior median.

It is interesting to notice that shocks to the CPI were on average positively skewed before the 2000s while became left skew for the rest of the sample that ends on 2019. This switch in the sign of the shape parameter indicates that conditionally on the past and on the contemporaneous

14. Table 7 in the Appendix reports the variable transformation.
15. For the VAR with *Skew-Normal* the estimated path for the volatilities and shape parameters are almost the same.
16. See the Appendix for the details on the hyper-parameters of the Minnetota Prior.

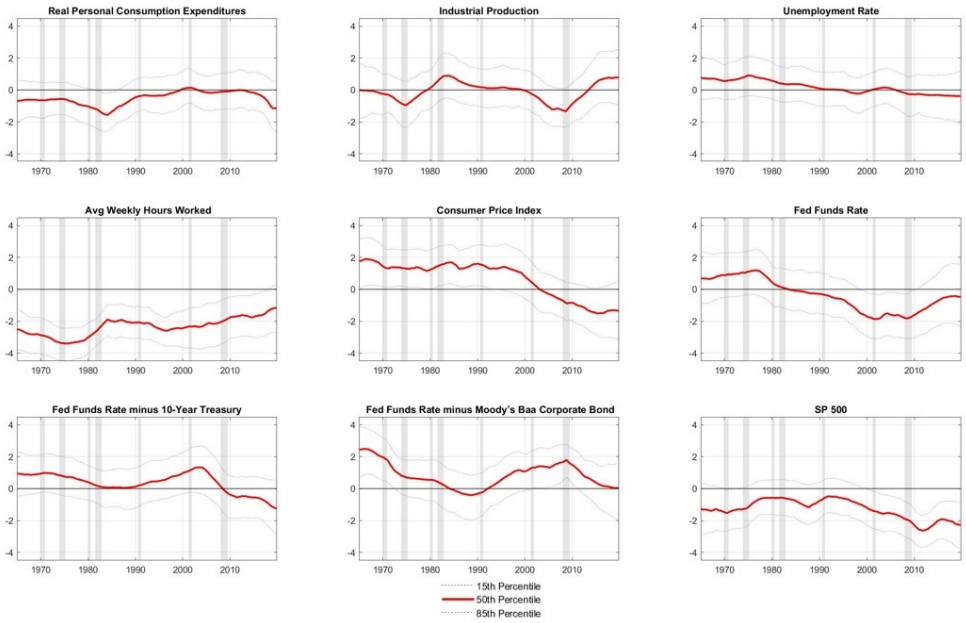
realization of Real Personal Consumption Expenditures, Industrial Production, Unemployment Rate and average Weekly Hours Worked, the distribution of CPI was right skewed in the 1980s, becoming instead left skew from the 2000s. In other words, risk switched from the upside to the downside. As for the monetary policy shocks, in the 1980s large positive hikes of the Fed Fund Rate were more frequent, while from the early 2000s large negative shocks to the Fed Fund Rate become more likely. Shocks to the average Weekly Hours Worked are skewed to the left over the entire sample, which means that negative large shocks have been systematically more frequent than positive large shocks. As well, shocks to the stock market (SP 500 index) are skewed to the left over the entire sample. This is in line with the large body of the financial econometrics literature that studies conditional skewness in asset returns (Harvey et al. 2000). As for the spread between 10-Year Treasury and the Fed Funds Rate and the spread between Moody's Baa Corporate Bond and the Fed Funds Rate, for most of the sample both the shocks are skewed to the right meaning that the probability of large positive shocks has been greater than the probability of large negative shocks. This finding vanishes starting from 2009 and might be linked to the unconventional monetary policy following the Great Financial Crisis.

Figure 4: Estimated volatilities



Note: The figure shows the estimated volatilities of the shocks in the TVSSV-VAR with *Skew-t* shocks. In red the estimated median, in blue dashed lines the $85^{th} - 15^{th}$ credible sets.

Figure 5: Estimated shape parameters λ_t



Note: The figure shows the estimated shape parameters of the shocks in the TVSSV-VAR with *Skew-t* shocks. In red the estimated median, in blue dashed lines the $85^{th} - 15^{th}$ credible sets.

5 Conclusion

In this paper I propose a fully parametric framework based on time varying skewness stochastic volatility models with *Skew-Normal* and *Skew-t* shocks for assessing and forecasting macroeconomic tail risk. First, I consider an extension of the univariate stochastic volatility model of Jacquier et al. (1994) that explicitly accounts for time varying skewness in the predictive distribution of the dependent variable. Then, I introduce a Bayesian VAR model with stochastic volatility and stochastic skewness to provide an explicit treatment of conditional skewness when modelling the dynamics of multiple time series. I compare the time varying skewness stochastic volatility models to the quantile regression method of Adrian et al. (2019) to assess and predict tail risk to GDP growth. I find that the time varying skewness stochastic volatility models considered in this paper are able to reproduce the main findings of Adrian et al. (2019), that is the nonlinear and asymmetric effect of financial conditions on the future GDP growth distribution. The models predict as much risk as quantile regression during the Financial crisis while provide slightly more accurate out of sample forecasts of downside risk over the entire sample. Finally, estimating a standard medium scale VAR model I find that time varying skewness is a relevant

feature of macroeconomic and financial shocks.

Future research For future research, the VAR model considered in this paper could be used to study the probability of joint tail events and for constructing structural scenarios of "at-risk" measures. For example, it could be used to study and assess stagflation risk, or to analyze scenarios for inflation at-risk and labour-at risk under different monetary policy paths. As a methodological extension, particularly interesting would be to consider an order invariant version of this model.

References

Abanto-Valle, CA, VH Lachos, and Dipak K Dey. 2015. "Bayesian estimation of a skew-student-t stochastic volatility model." *Methodology and Computing in Applied Probability* 17 (3): 721–738.

Adrian, Tobias, Nina Boyarchenko, and Domenico Giannone. 2019. "Vulnerable Growth." *American Economic Review* 109, no. 4 (April): 1263–89.

Andrieu, Christophe, Arnaud Doucet, and Roman Holenstein. 2010. "Particle markov chain monte carlo methods." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 72 (3): 269–342.

Arias, Jonas E., Juan F. Rubio-Ramirez, and Minchul Shin. 2021. *Macroeconomic Forecasting and Variable Ordering in Multivariate Stochastic Volatility Models.* Working Papers 21-21. Federal Reserve Bank of Philadelphia, June.

Azzalini, Adelchi. 1986. "Further results on a class of distributions which includes the normal ones." *Statistica* 46 (2): 199–208.

Azzalini, Adelchi, and Antonella Capitanio. 2003. "Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 65 (2): 367–389.

Brownlees, Christian, and André B.M. Souza. 2021. "Backtesting global Growth-at-Risk." *Journal of Monetary Economics* 118:312–330.

Cappuccio, Nunzio, Diego Lubian, and Davide Raggi. 2004. "MCMC Bayesian Estimation of a Skew-GED Stochastic Volatility Model." *Studies in Nonlinear Dynamics & Econometrics* 8 (2).

Carriero, Andrea, Joshua Chan, Todd E. Clark, and Massimiliano Marcellino. 2022. "Corrigendum to "Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors" [J. Econometrics 212 (1) (2019) 137–154]." *Journal of Econometrics* 227 (2): 506–512.

Carriero, Andrea, Todd E Clark, and Massimiliano Marcellino. 2019. "Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors." *Journal of Econometrics* 212 (1): 137–154.

———. 2015. "Realtime nowcasting with a Bayesian mixed frequency model with stochastic volatility." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 178 (4): 837–862.

———. 2020. *Capturing Macroeconomic Tail Risks with Bayesian Vector Autoregressions.* Working Papers 20-02R. Federal Reserve Bank of Cleveland, January.

Chan, Joshua CC, Gary Koop, and Xuewen Yu. 2021. "Large order-invariant Bayesian VARs with stochastic volatility." *arXiv preprint arXiv:2111.07225.*

Clark, Todd E., and Francesco Ravazzolo. 2015. "Macroeconomic Forecasting Performance under Alternative Specifications of Time-Varying Volatility." *Journal of Applied Econometrics* 30 (4): 551–575.

Cogley, Timothy, and Thomas J. Sargent. 2005. "Drifts and volatilities: monetary policies and outcomes in the post WWII US." Monetary Policy and Learning, *Review of Economic Dynamics* 8 (2): 262–302.

Delle Monache, Davide, Andrea De Polis, and Ivan Petrella. 2021. *Modeling and forecasting macroeconomic downside risk.* Temi di discussione (Economic working papers) 1324. Bank of Italy, Economic Research and International Relations Area, March.

Diebold, Francis, and Roberto Mariano. 1995. "Comparing Predictive Accuracy." *Journal of Business & Economic Statistics* 13 (3): 253–63.

Gelos, Gaston, Lucyna Gornicka, Robin Koepke, Ratna Sahay, and Silvia Sgherri. 2022. "Capital flows at risk: Taming the ebbs and flows." *Journal of International Economics* 134:103555.

Giacomini, Raffaella, and Ivana Komunjer. 2005. "Evaluation and Combination of Conditional Quantile Forecasts." *Journal of Business & Economic Statistics* 23 (4): 416–431.

Giglio, Stefano, Bryan Kelly, and Seth Pruitt. 2016. "Systemic risk and the macroeconomy: An empirical evaluation." *Journal of Financial Economics* 119 (3): 457–471.

Gneiting, Tilmann, and Roopesh Ranjan. 2011. "Comparing Density Forecasts Using Threshold- and Quantile-Weighted Scoring Rules." *Journal of Business & Economic Statistics* 29 (3): 411–422.

Harvey, Campbell R, and Akhtar Siddique. 2000. "Conditional skewness in asset pricing tests." *The Journal of finance* 55 (3): 1263–1295.

Iseringhausen, Martin. 2020. "The time-varying asymmetry of exchange rate returns: A stochastic volatility – stochastic skewness model." *Journal of Empirical Finance* 58:275–292.

———. 2021. "A time-varying skewness model for Growth-at-Risk."

Jacquier, Eric, Nicholas G Polson, and Peter Rossi. 1994. "Bayesian Analysis of Stochastic Volatility Models." *Journal of Business & Economic Statistics* 12 (4): 371–89.

Jacquier, Eric, Polson, and Peter Rossi. 2004. "Bayesian analysis of stochastic volatility models with fat-tails and correlated errors." *Journal of Econometrics* 122 (1): 185–212.

Karlsson, Sune, Stepan Mazur, and Hoang Nguyen. 2023. "Vector autoregression models with skewness and heavy tails." *Journal of Economic Dynamics and Control* 146:104580.

Kiley, Michael T. 2018. "Unemployment risk." *Journal of Money, Credit and Banking.*

Kilian, Lutz, and Simone Manganelli. 2003. *The central bank as a risk manager: quantifying and forecasting inflation risks.* Working Paper Series 226. European Central Bank, April.

Lindsten, Fredrik, Michael I Jordan, and Thomas B Schon. 2014. "Particle Gibbs with ancestor sampling." *Journal of Machine Learning Research* 15:2145–2184.

López-Salido, J. David, and Francesca Loria. 2020. *Inflation at Risk.* Finance and Economics Discussion Series 2020-013. Board of Governors of the Federal Reserve System (U.S.), February.

Montes-Galdón, Carlos, and Eva Ortega. 2022. "Skewed SVARS: Tracking the structural sources of macroeconomic tail risks." In *Essays in Honour of Fabio Canova,* 44:177–210. Emerald Publishing Limited.

Primiceri, Giorgio E. 2005. "Time varying structural vector autoregressions and monetary policy." *The Review of Economic Studies* 72 (3): 821–852.

Sims, Christopher A. 1980. "Macroeconomics and reality." *Econometrica: journal of the Econometric Society,* 1–48.

Stock, James H., and Mark W. Watson. 2001. "Vector Autoregressions." *Journal of Economic Perspectives* 15, no. 4 (December): 101–115.

Wolf, Elias. 2021. "Estimating growth at risk with skewed stochastic volatility models." *Available at SSRN 4030094.*

A Appendix

A.1 Skew Normal and Skew-t: distributions and parameterization

The *Skew Normal* (Azzalini 1986) distribution is:

$$p(\varepsilon_t|\zeta,\omega^2,\lambda) = \frac{2}{\omega}\phi\left(\frac{\varepsilon_t-\zeta}{\omega}\right)\Phi\left(\lambda\left(\frac{\varepsilon_t-\zeta}{\omega}\right)\right)$$

where $\phi(.)$ and $\Phi(.)$ are respectively the pdf and cdf of the standard Normal. In general, $\varepsilon_t \sim Skew-Normal(\zeta,\omega^2,\lambda)$ has the following stochastic representation:

$$\varepsilon_t = \zeta + \delta\omega v_t + \sqrt{(1-\delta^2)}\omega z_t \tag{28}$$

where:

$v_t \overset{i.i.d}{\sim}$ Truncated Normal$_{[0,\infty)}(0,1)$

$z_t \overset{i.i.d}{\sim} \mathcal{N}(0,1)$

$\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$, with $-1 < \delta < 1$.

The mean and the variance of ε_t are given by:

$$\mathbb{E}[\varepsilon_t] = \zeta + \omega\delta\sqrt{\frac{2}{\pi}} \tag{29}$$

$$var(\varepsilon_t) = \omega^2\left(1-\frac{2\delta^2}{\pi}\right) \tag{30}$$

Assuming $\mathbb{E}[\varepsilon_t] = 0$ and $var(\varepsilon_t) = 1$ leads to the following constraints on the location and scale parameters: $\zeta = -\omega\delta\sqrt{\frac{2}{\pi}}$ and $\omega^2 = \left(1-\frac{2\delta^2}{\pi}\right)^{-1}$. Once we impose these constraints on the location and scale parameters, with $\lambda = 0$ the distribution collapses to the Standard Normal.

The *Skew-t* distribution (Azzalini et al. 2003) is:

$$p(\varepsilon_t|\zeta,\omega^2,\lambda,\nu) = \frac{2}{\omega}t_\nu\left(\frac{\varepsilon_t-\zeta}{\omega}\right)T_{\nu+1}\left(\lambda\left(\frac{\varepsilon_t-\zeta}{\omega}\right)\sqrt{\frac{\nu+1}{\nu\left(\frac{\varepsilon_t-\zeta}{\omega}\right)^2}}\right) \tag{31}$$

where $t(.)$ and $T(.)$ are respectively the pdf and cdf of the Student-t with ν degrees of freedom. $\varepsilon_t \sim$ *Skew-t*$(\zeta,\omega^2,\lambda,\nu)$ has the following stochastic representation:

$$\varepsilon_t = \zeta + \delta\omega o_t^{-0.5} v_t + \sqrt{(1-\delta^2)}\omega o_t^{-0.5} z_t \tag{32}$$

where:
$v_t \overset{i.i.d}{\sim}$ Truncated Normal$_{[0,\infty)}(0,1)$
$z_t \overset{i.i.d}{\sim} \mathcal{N}(0,1)$
$o_t \overset{i.i.d}{\sim} \mathcal{G}(\frac{\nu}{2}, \frac{\nu}{2})$
$\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$, with $-1 < \delta < 1$.

The mean and the variance of ε_t are given by:

$$\mathbb{E}[\varepsilon_t] = \omega\delta k_1 \sqrt{\frac{2}{\pi}} \tag{33}$$

$$var(\varepsilon_t) = \omega^2 \left(k_2 - \frac{2}{\pi} k_1^2 \delta^2 \right) \tag{34}$$

with $k_1 = \sqrt{\frac{\nu}{2}} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})}$, $k_2 = \frac{\nu}{\nu-2}$.

Assuming $\mathbb{E}[\varepsilon_t] = 0$ and $var(\varepsilon_t) = 1$ leads to the following constraints on the location and scale parameters: $\zeta = -\omega\delta k_1 \sqrt{\frac{2}{\pi}}$ and $\omega^2 = \left(k_2 - \frac{2}{\pi} k_1^2 \delta^2\right)^{-1}$. Once we impose these constraints on the location and scale parameters with $\lambda = 0$ the distribution collapses to a *Student-t* distribution properly re-scaled to have unit variance (and zero mean).

A.2 Full conditional posterior distributions

A.2.1 Univariate time varying skewness stochastic volatility model: skew normal shocks

The full conditional distribution of $\{v_t\}_{t=1}^T$ is given by:

$$p(v_t|.) \propto exp\left[-\frac{1}{2} \left(\frac{1}{(1-\delta_t^2)} v_t^2 - \frac{2\delta_t h_t^{-0.5}}{\omega_t (1-\delta_t^2)} (y_t - \boldsymbol{x_t}\boldsymbol{\pi} - \zeta_t \sqrt{h_t}) v_t \right) \right] \mathbb{I}(0 \leq v_t < \infty) \tag{35}$$

this is a *Truncated Normal*$\left(\frac{\delta_t h_t^{-0.5}[y_t - \boldsymbol{x_t}\boldsymbol{\pi}] - \delta_t \zeta_t}{\omega_t}, 1-\delta_t^2 \right)_{[0,\infty)}$

The full conditional distribution of $\boldsymbol{\pi}$ is Normal:

$$f(\boldsymbol{\pi}|.) \sim N(\overline{\mu}_\pi, \overline{\Sigma}_\pi) \tag{36}$$

$$\begin{aligned}\overline{\mu}_\pi &= \overline{\Sigma}_\pi \left(\sum_{t=1}^{T} \tfrac{1}{\sigma_t^2} \boldsymbol{x_t'} \tilde{y}_t + \underline{\Sigma_\pi}^{-1} \underline{\mu_\pi} \right) \\ \overline{\Sigma}_\pi^{-1} &= \underline{\Sigma_\pi}^{-1} + \sum_{t=1}^{T} \tfrac{1}{\sigma_t^2} \boldsymbol{x_t'} \boldsymbol{x_t}\end{aligned}$$

where

$\tilde{y}_t \equiv y_t - \sqrt{h_t}\zeta_t - \sqrt{h_t}\omega_t\delta_t v_t$

$\sigma_t^2 \equiv h_t\omega_t^2(1-\delta_t^2)$

while $\underline{\mu_\pi}$ and $\underline{\Sigma_\pi}$ are the prior mean and variance covariance matrix.

The full conditional distribution of ϕ_h is a *Normal* :

$$f(\phi_h|.) \sim \mathcal{N}(\bar{\mu}_{\phi_h}, \bar{\sigma}^2_{\phi_h}) \tag{37}$$

$$\bar{\sigma}^2_{\phi_h} = \left(\sum_{t=1}^{T} \frac{log(h_{t-1})^2}{\sigma_\eta^2} + \frac{1}{\underline{\sigma^2_{\phi_h}}} \right)^{-1} \tag{38}$$

$$\bar{\mu}_{\phi_h} = \bar{\sigma}^2_{\phi_h} \left(\frac{\underline{\mu}_{\phi_h}}{\underline{\sigma^2_{\phi_h}}} + \sum_{t=1}^{T} \frac{log(h_{t-1})log(h_t)}{\sigma_\eta^2} \right) \tag{39}$$

where $\underline{\mu}_{\phi_h}$ and $\underline{\sigma}^2_{\phi_h}$ are prior mean and variance.

The full conditional distribution of σ_η^2 is an *Inverse Gamma* :

$$p(\sigma_\eta^2|.) \propto \left(\frac{1}{\sigma_\eta^2}\right)^{\frac{T}{2}} exp\left[\sum_{t=1}^{T} -\frac{1}{2\sigma^2_{\eta,i}}(ln(h_t) - \phi_h ln(h_{t-1}))^2\right] exp\left[-\frac{s_{\sigma_\eta^2}}{\sigma_\eta^2}\right] {\sigma_\eta^2}^{-\nu_{\sigma_\eta^2}-1} \tag{40}$$

where $s_{\sigma_\eta^2}$ and $\nu^2_{\sigma_\eta}$ are the hyper-parameters of the Inverse Gamma prior.

The full conditional distribution of ϕ_λ is a *Normal* :

$$f(\phi_\lambda|.) \sim \mathcal{N}(\bar{\mu}_{\phi_\lambda}, \bar{\sigma}^2_{\phi_\lambda}) \tag{41}$$

$$\bar{\sigma}^2_{\phi_\lambda} = \left(\sum_{t=1}^{T} \frac{\lambda_{t-1}^2}{\sigma_\xi^2} + \frac{1}{\underline{\sigma^2_{\phi_\lambda}}} \right)^{-1} \tag{42}$$

$$\bar{\mu}_{\phi_\lambda} = \bar{\sigma}^2_{\phi_\lambda} \left(\frac{\underline{\mu}_{\phi_\lambda}}{\underline{\sigma^2_{\phi_\lambda}}} + \sum_{t=1}^{T} \frac{\lambda_{t-1}\lambda_t}{\sigma_\xi^2} \right) \tag{43}$$

where $\underline{\mu}_{\phi\lambda}$ and $\underline{\sigma}^2_{\phi_\lambda}$ are prior mean and variance.

The full conditional distribution of σ^2_ξ is an *Inverse Gamma*:

$$p(\sigma^2_\xi|.) \propto \left(\frac{1}{\sigma^2_\xi}\right)^{\frac{T}{2}} exp\left[\sum_{t=1}^{T} -\frac{1}{2\sigma^2_\xi}(\lambda_{it} - \phi_\lambda \lambda_{it-1})^2\right] exp\left[-\frac{s_{\sigma^2_\xi}}{\sigma^2_\xi}\right] \sigma_\xi^{2^- \nu_{\sigma^2_\xi}-1} \tag{44}$$

where $s_{\sigma^2_\xi}$ and $\nu_{\sigma^2_\xi}$ are the hyper-parameters of the Inverse Gamma prior.

The full conditional distribution of h_0 is $\mathcal{N}(\bar{\mu}_{h0}, \bar{\sigma}_{h0})$

$$\bar{\mu}_{h0} = \bar{\sigma}_{h0}\left(\frac{\mu_{h0}}{\sigma^2_{h0}} + \frac{\frac{log(h_1)}{\phi_h}}{\frac{\sigma^2_\eta}{\phi^2_h}}\right) \tag{45}$$

$$\bar{\sigma}_{h0} = \frac{\frac{\sigma^2_{h_0}\sigma^2_\xi}{\phi^2_h}}{\sigma^2_{h0} + \frac{\sigma^2_\eta}{\phi^2_h}} \tag{46}$$

where μ_{λ_0} and $\sigma^2_{\lambda_0}$ are the prior mean and variance.

The full conditional distribution of $\boldsymbol{h}$ is given by:

$$p(\boldsymbol{h}|.) = \prod_{t=1}^{T} p(h_t|h_{t-1}, h_{t+1}, .) \tag{47}$$

$$p(h_t|.) \propto h_t^{-1,5} exp\left[\left(-\frac{1}{2}\left(\frac{y_t - \boldsymbol{x_t\pi} - \sqrt{h_t}\zeta_t - \sqrt{h_t}\omega_t\delta_t v_t}{\sqrt{h_t}\omega_t(1-\delta_t^2)^{0.5}}\right)^2 - \frac{1}{2}\frac{(lnh_t - \mu_{h_t})^2}{\sigma^2_{h_t}}\right)\right] \tag{48}$$

$$\begin{aligned} \mu_{h_t} &= \frac{\phi_h}{\phi^2_h+1}(lnh_{t+1} + lnh_{t-1}) \\ \sigma^2_h &= \frac{\sigma^2_\eta}{\phi^2_h+1} \end{aligned} \tag{49}$$

The full conditional distribution of λ_0 is $\mathcal{N}(\bar{\mu}_{\lambda 0}, \bar{\sigma}_{\lambda 0})$

$$\bar{\mu}_{\lambda 0} = \bar{\sigma}_{\lambda 0}\left(\frac{\mu_{\lambda_0}}{\sigma^2_{\lambda_0}} + \frac{\frac{\lambda_1}{\phi_\lambda}}{\frac{\sigma^2_\xi}{\phi^2_\lambda}}\right) \tag{50}$$

$$\bar{\sigma}_{\lambda 0} = \frac{\frac{\sigma^2_{\lambda_0}\sigma^2_\xi}{\phi^2_\lambda}}{\sigma^2_{\lambda_0} + \frac{\sigma^2_\xi}{\phi^2_\lambda}} \tag{51}$$

where μ_{λ_0} and $\sigma^2_{\lambda_0}$ are the prior mean and variance.

The full conditional distribution of $\boldsymbol{\lambda}$ is given by:

$$p(\boldsymbol{\lambda}|.) = \prod_{t=1}^{T} p(\lambda_t|\lambda_{t-1}, \lambda_{t+1}, .) \tag{52}$$

$$p(\lambda_t|.) \propto \omega_t^{-1}(1-\delta_t^2)^{-0,5} exp\left[-\frac{1}{2}\left(\frac{y_t - \boldsymbol{x_t}\pi - \sqrt{h_t}\zeta_t - \sqrt{h_t}o_t^{-0.5}\omega_t\delta_t v_t}{\sqrt{h_t}o_t^{-0.5}\omega_t(1-\delta_t^2)^{0.5}}\right)^2 - \frac{1}{2}\frac{(\lambda_t - \mu_{\lambda_t})^2}{\sigma_{\lambda_t}^2}\right] \tag{53}$$

$$\begin{aligned} \mu_{\lambda_t} &= \frac{\phi_\lambda}{\phi_\lambda^2+1}(\lambda_{t+1} + \lambda_{t-1}) \\ \sigma_\lambda^2 &= \frac{\sigma_\xi^2}{\phi_\lambda^2+1} \end{aligned} \tag{54}$$

A.2.2 Univariate time varying skewness stochastic volatility model: skew-t shocks

The full conditional distribution of $\{v_t\}_{t=1}^T$ is given by:

$$p(v_t|.) \propto exp\left[-\frac{1}{2}\left(\frac{1}{(1-\delta_t^2)}v_t^2 - \frac{2o_t^{0.5}\delta_t h_t^{-0.5}}{\omega_t(1-\delta_t^2)}(y_t - \boldsymbol{x_t}\boldsymbol{\pi} - \zeta_t\sqrt{h_t})v_t\right)\right]\mathbb{I}(0 \leq v_t < \infty) \tag{55}$$

this is a truncated normal $N\left(\frac{\delta_t o_t^{0,5} h_t^{-0.5}[y_t - \boldsymbol{x_t}\boldsymbol{\pi}] - \delta_t o_t^{0.5}\zeta_t}{\omega_t}, 1-\delta_t^2\right)_{[0,\infty)}$

The full conditional distribution of $\{o_t\}_{t=1}^T$ is given by:

$$p(o_t|.) \propto o_t^{\frac{\nu+1}{2}-1} exp\left[-\frac{o_t}{2}\left(\nu_t + \frac{h_t^{-1}(y_t - \boldsymbol{x_t}\boldsymbol{\pi} - \sqrt{h_t}\zeta_t)^2}{\omega_t^2(1-\delta_t^2)}\right)\right] exp\left[\frac{(y_t - \boldsymbol{x_t}\boldsymbol{\pi} - \sqrt{h_t}\zeta_t)(h_t^{-0.5}o_t^{0.5}\delta_t v_t)}{\omega_t(1-\delta_t^2)}\right] \tag{56}$$

The full conditional distribution of $\boldsymbol{\pi}$ is Normal:

$$f(\boldsymbol{\pi}|.) \sim N(\overline{\mu}_\pi, \overline{\Sigma}_\pi) \tag{57}$$

$$\begin{aligned} \overline{\mu}_\pi &= \overline{\Sigma}_\pi\left(\sum_{t=1}^T \frac{1}{\sigma_t^2}\boldsymbol{x_t'}\tilde{y}_t + \underline{\Sigma_\pi}^{-1}\underline{\mu_\pi}\right) \\ \overline{\Sigma}_\pi^{-1} &= \underline{\Sigma_\pi}^{-1} + \sum_{t=1}^T \frac{1}{\sigma_t^2}\boldsymbol{x_t'}\boldsymbol{x_t} \end{aligned}$$

where:

$\tilde{y}_t \equiv y_t - \sqrt{h_t}\zeta_t - \sqrt{h_t}o_t^{-0.5}\omega_t\delta_t v_t$

$\sigma_t^2 \equiv h_t\omega_t^2 o_t^{-1}(1-\delta_t^2)$

The full conditional distribution of h_t is given by:

$$p(\boldsymbol{h}|.) = \prod_{t=1}^{T} p(h_t|h_{t-1}, h_{t+1}, .) \tag{58}$$

$$p(h_t|.) \propto h_t^{-1,5} exp\left[\left(-\frac{1}{2}\left(\frac{y_t - \boldsymbol{x_t}\boldsymbol{\pi} - \sqrt{h_t}\zeta_t - \sqrt{h_t}o_t^{-0.5}\omega_t\delta_t v_t}{\sqrt{h_t}o_t^{-0.5}\omega_t(1-\delta_t^2)^{0.5}}\right)^2 - \frac{1}{2}\frac{(lnh_t - \mu_{h_t})^2}{\sigma_{h_t}^2}\right)\right] \tag{59}$$

$$
\begin{aligned}
l\mu_{h_t} &= \frac{\phi_h}{\phi_h^2+1}\left(lnh_{t+1}+lnh_{t-1}\right) \\
\sigma_h^2 &= \frac{\sigma_\eta^2}{\phi_h^2+1}
\end{aligned}
\tag{60}
$$

The full conditional distribution of λ_t is given by:

$$
p(\boldsymbol{\lambda}|.) = \prod_{t=1}^{T} p(\lambda_t|\lambda_{t-1},\lambda_{t+1},.) \tag{61}
$$

$$
p(\lambda_t|.) \propto \omega_t^{-1}(1-\delta_t^2)^{-0,5} exp\left[-\frac{1}{2}\left(\frac{y_t-\boldsymbol{x_t}\pi-\sqrt{h_t}\zeta_t-\sqrt{h_t}o_t^{-0.5}\omega_t\delta_t v_t}{\sqrt{h_t}o_t^{-0.5}\omega_t(1-\delta_t^2)^{0.5}}\right)^2-\frac{1}{2}\frac{(\lambda_t-\mu_{\lambda_t})^2}{\sigma_{\lambda_t}^2}\right] \tag{62}
$$

$$
\begin{aligned}
l\mu_{\lambda_t} &= \frac{\phi_\lambda}{\phi_\lambda^2+1}\left(\lambda_{t+1}+\lambda_{t-1}\right) \\
\sigma_\lambda^2 &= \frac{\sigma_\xi^2}{\phi_\lambda^2+1}
\end{aligned}
\tag{63}
$$

The full conditional distribution of $\boldsymbol{v} = v_1,\ldots,v_T$ is given by:

$$
p(\boldsymbol{v}|.) = \prod_{t=1}^{T} p(v_t|.) \tag{64}
$$

$$
p(v_t|.) \propto exp\left[-\frac{1}{2}\left(\frac{1}{(1-\delta_t^2)}v_t^2-\frac{2o_t^{0.5}\delta_t h_t^{-0.5}}{\omega_t(1-\delta_t^2)}(y_t-\boldsymbol{x_t}\boldsymbol{\pi}-\zeta_t\sqrt{h_t})v_t\right)\right]\mathbb{I}(0\leq v_t<\infty) \tag{65}
$$

this is is a *Truncated Normal* $\left(\frac{\delta_t o_t^{0,5}h_t^{-0.5}[y_t-\boldsymbol{x_t}\boldsymbol{\pi}]-\delta_t o_t^{0,5}\zeta_t}{\omega_t}, 1-\delta_t^2\right)_{[0,\infty)}$

A.2.3 VAR with Skew Normal shocks

The full conditional distribution of $vec(\boldsymbol{\Pi})$ is $\mathcal{N}(vec(\bar{\boldsymbol{\mu}}_{\boldsymbol{\Pi}}), \bar{\boldsymbol{V}}_{\boldsymbol{\Pi}})$, where:

$$
\bar{\boldsymbol{\mu}}_{\boldsymbol{\Pi}} = \bar{\boldsymbol{V}}_{\boldsymbol{\Pi}}\left[vec\left(\sum_{t=1}^{T}\boldsymbol{X_t}\tilde{\boldsymbol{y}}_t'\boldsymbol{\Sigma_t^{-1}}\right)+\underline{\boldsymbol{V}_{\boldsymbol{\Pi}}}^{-1}vec(\underline{\boldsymbol{\mu}_{\boldsymbol{\Pi}}})\right] \tag{66}
$$

with $\tilde{\boldsymbol{y}}_t \equiv \boldsymbol{y}_t - \boldsymbol{H_t^{0.5}}\boldsymbol{A^{-1}}\boldsymbol{\zeta}_t - \boldsymbol{H_t^{0.5}}\boldsymbol{A^{-1}}\boldsymbol{\Omega_t}\boldsymbol{\Delta_t}\boldsymbol{v_t}$ and:

$$
\bar{\boldsymbol{V}}_{\boldsymbol{\Pi}} = \underline{\boldsymbol{V}_{\boldsymbol{\Pi}}}^{-1} + \sum_{t=1}^{T}(\boldsymbol{\Sigma_t^{-1}}\otimes\boldsymbol{X_t}\boldsymbol{X_t'}) \tag{67}
$$

where $\boldsymbol{\Sigma}_t \equiv \boldsymbol{A}^{-1}\boldsymbol{H}_t\boldsymbol{\omega}_t^2(\boldsymbol{I}-\boldsymbol{\Delta}_t^2)\boldsymbol{A'}^{-1}$ while $\underline{\boldsymbol{\mu}_{\boldsymbol{\Pi}}}$ and $\underline{\boldsymbol{V}_{\boldsymbol{\Pi}}}$ are the prior mean and variance covariance matrix.

The full conditional distribution of the elements in $\boldsymbol{A}$ is derived adapting our framework to

the approach of Cogley et al. (2005). Considering the system:

$$\boldsymbol{A}\boldsymbol{u_t} = \boldsymbol{H_t^{0.5}}\boldsymbol{\varepsilon_t} \tag{68}$$

since $\varepsilon_{it} = \zeta_{it} + \omega_{it}\delta_{it}v_{it} + \omega_{it}\sqrt{1-\delta_{it}^2}z_{it}$ we have :

$$\begin{aligned}
&u_{1t} = \sqrt{h}_{1t}(\zeta_{1t} + \omega_{1t}\delta_{1t}v_{1t} + \omega_{1t}\sqrt{1-\delta_{1t}^2}z_{1t})\\
&u_{2t} = -a_{21}u_{1t} + \sqrt{h}_{2t}(\zeta_{2t} + \omega_{2t}\delta_{2t}v_{2t} + \omega_{2t}\sqrt{1-\delta_{2t}^2}z_{2t})\\
&u_{3t} = -a_{31}u_{1t} - a_{32}u_{2t} + \sqrt{h}_{3t}(\zeta_{3t} + \omega_{3t}\delta_{3t}v_{3t} + \omega_{3t}\sqrt{1-\delta_{3t}^2}z_{3t})\\
&\quad\vdots \qquad\qquad \vdots \qquad\qquad \vdots\\
&u_{Nt} = -a_{N1}u_{1t} - a_{N2}u_{2t} \ldots - a_{N,N-1}u_{2t} + \sqrt{h}_{Nt}(\zeta_{Nt} + \omega_{Nt}\delta_{Nt}v_{Nt} + \omega_{Nt}\sqrt{1-\delta_{Nt}^2}z_{Nt})
\end{aligned} \tag{69}$$

therefore :

$$\begin{aligned}
&u_{1t} - \sqrt{h}_{1t}(\zeta_{1t} + \omega_{1t}\delta_{1t}v_{1t}) = \sqrt{h}_{1t}\omega_{1t}\sqrt{1-\delta_{1t}^2}z_{1t}\\
&u_{2t} - \sqrt{h}_{2t}(\zeta_{2t} + \omega_{2t}\delta_{2t}v_{2t}) = -a_{21}u_{1t} + \sqrt{h}_{2t}\omega_{2t}\sqrt{1-\delta_{2t}^2}z_{2t}\\
&u_{3t} - \sqrt{h}_{3t}(\zeta_{3t} + \omega_{3t}\delta_{3t}v_{3t}) = -a_{31}u_{1t} - a_{32}u_{2t} + \sqrt{h}_{3t}\omega_{3t}\sqrt{1-\delta_{3t}^2}z_{3t}\\
&\quad\vdots \qquad\qquad \vdots \qquad\qquad \vdots\\
&u_{Nt} - \sqrt{h}_{Nt}(\zeta_{Nt} + \omega_{Nt}\delta_{Nt}v_{Nt}) = -a_{N1}u_{1t} - a_{N2}u_{2t} \ldots - a_{N,N-1}u_{2t} + \sqrt{h}_{Nt}\omega_{Nt}\sqrt{1-\delta_{Nt}^2}z_{Nt}
\end{aligned}$$

Since I condition on the parameters, the mixing variables and the latent states I can define $\tilde{u}_{it} = u_{it} - \sqrt{h}_{it}(\zeta_{it} + \omega_{it}\delta_{it}v_{it})$ for $i = 1, \ldots, N$ and $\tilde{\sigma}_{it}^2 = \sqrt{h_{it}}\omega_{it}\sqrt{1-\delta_{it}^2}$ and derive the full conditional posterior for the elements of $\boldsymbol{A}$ by exploiting the system of equations:

$$\begin{aligned}
&\tilde{u}_{1t} = \tilde{\sigma}_{1t}^2 z_{1t}\\
&\tilde{u}_{2t} = -a_{21}u_{1t} + \tilde{\sigma}_{2t}^2 z_{2t}\\
&\tilde{u}_{3t} = -a_{31}u_{1t} - a_{32}u_{2t} + \tilde{\sigma}_{3t}^2 z_{3t}\\
&\quad\vdots \qquad\qquad \vdots \qquad\qquad \vdots\\
&\tilde{u}_{Nt} = -a_{N1}u_{1t} - a_{N2}u_{2t} \ldots - a_{N,N-1}u_{2t} + \tilde{\sigma}_{Nt}^2 z_{Nt}
\end{aligned} \tag{70}$$

where $z_{it} \sim \mathcal{N}(0,1)$. Assuming a *Normal* prior for the elements in $\boldsymbol{A}$ and defining $\boldsymbol{a}_i$ the vector that collects the free elements in the i^{th} row of the $\boldsymbol{A}$ matrix, I can use standard linear regression results to show that the full conditional posterior of $\boldsymbol{a_i}$ is given by $\boldsymbol{a_i} \sim \mathcal{N}(\bar{\boldsymbol{\mu}}_{\boldsymbol{a,i}}, \bar{\boldsymbol{V}}_{\boldsymbol{a,i}})$ where:

$$\begin{aligned}
&\bar{\boldsymbol{\mu}}_{\boldsymbol{a,i}} = \bar{\boldsymbol{V}}_{\boldsymbol{a,i}}(\underline{\boldsymbol{V}_{\boldsymbol{a,i}}}^{-1}\underline{\boldsymbol{\mu}_{\boldsymbol{a}}} + \textstyle\sum_{t=1}^{T}\tilde{\sigma_{it}}^{2^{-1}}\boldsymbol{u'_{it}}\tilde{u}_{it})\\
&\bar{\boldsymbol{V}}_{\boldsymbol{a,i}} = (\underline{\boldsymbol{V}_{\boldsymbol{a,i}}}^{-1} + \textstyle\sum_{t=1}^{T}\tilde{\sigma_{it}}^{2^{-1}}\boldsymbol{u'_{it}}\boldsymbol{u_{it}})^{-1}
\end{aligned} \tag{71}$$

where $\boldsymbol{u_{it}}$ is the vector colleting the right hand variables of the i^{th} equation in the system above (70) with $i = 2, \ldots, N$ and $\underline{\boldsymbol{\mu}_{\boldsymbol{a}}}$ and $\underline{\boldsymbol{V}_{\boldsymbol{a,i}}}$ are the prior mean and variance covariance matrix

of the free elements of the i^{th} row of $\boldsymbol{A}$.

A.2.4 VAR with Skew-t shocks

The full conditional distribution of $vec(\boldsymbol{\Pi})$ is $\mathcal{N}(vec(\bar{\boldsymbol{\mu}}_{\boldsymbol{\Pi}}), \bar{\boldsymbol{V}}_{\boldsymbol{\Pi}})$, where:

$$\bar{\boldsymbol{\mu}}_{\boldsymbol{\Pi}} = \bar{\boldsymbol{V}}_{\boldsymbol{\Pi}} \left[vec\left(\sum_{t=1}^{T} \boldsymbol{X_t} \tilde{\boldsymbol{y}}_{\boldsymbol{t}}' \boldsymbol{\Sigma_t^{-1}} \right) + \underline{\boldsymbol{V}_{\boldsymbol{\Pi}}}^{-1} vec(\underline{\boldsymbol{\mu}_{\boldsymbol{\Pi}}}) \right] \tag{72}$$

with $\tilde{\boldsymbol{y}}_t \equiv \boldsymbol{y}_t - \boldsymbol{H_t^{0.5}} \boldsymbol{A^{-1}} \boldsymbol{\zeta}_t - \boldsymbol{H_t^{0.5}} \boldsymbol{A^{-1}} \boldsymbol{\Omega_t} \boldsymbol{\Delta_t} \boldsymbol{O_t^{-0.5}} \boldsymbol{v_t}$ and:

$$\bar{\boldsymbol{V}}_{\boldsymbol{\Pi}} = \underline{\boldsymbol{V}_{\boldsymbol{\Pi}}}^{-1} + \sum_{t=1}^{T} (\boldsymbol{\Sigma_t^{-1}} \otimes \boldsymbol{X_t} \boldsymbol{X_t'}) \tag{73}$$

where $\boldsymbol{\Sigma_t} \equiv \boldsymbol{A^{-1}} \boldsymbol{H}_t \boldsymbol{\omega_t^2} (\boldsymbol{I} - \boldsymbol{\Delta_t^2}) \boldsymbol{O_t^{-1}} \boldsymbol{A'^{-1}}$ while $\underline{\boldsymbol{\mu}_{\boldsymbol{\Pi}}}$ and $\underline{\boldsymbol{V}_{\boldsymbol{\Pi}}}$ are the prior mean and variance covariance matrix.

The full conditional for $\boldsymbol{A}$ is derived following the same steps in the VAR with *Skew-normal* shocks just by considering that (69) becomes:

$$\begin{aligned}
u_{1t} &= \sqrt{h}_{1t}(\zeta_{1t} + \omega_{1t}\delta_{1t} o_{1t}^{-0.5} v_{1t} + \omega_{1t}\sqrt{1-\delta_{1t}^2}\, o_{1t}^{-0.5} z_{1t}) \\
u_{2t} &= -a_{21}u_{1t} + \sqrt{h}_{2t}(\zeta_{2t} + \omega_{2t}\delta_{2t} o_{2t}^{-0.5} v_{2t} + \omega_{2t}\sqrt{1-\delta_{2t}^2 o_{2t}^{-0.5}}\, z_{2t}) \\
u_{3t} &= -a_{31}u_{1t} - a_{32}u_{2t} + \sqrt{h}_{3t}(\zeta_{3t} + \omega_{3t}\delta_{3t} o_{3t}^{-0.5} v_{3t} + \omega_{3t}\sqrt{1-\delta_{3t}^2}\, o_{3t}^{-0.5} z_{3t}) \\
&\vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
u_{Nt} &= -a_{N1}u_{1t} - a_{N2}u_{2t} \ldots - a_{N,N-1}u_{2t} + \sqrt{h}_{Nt}(\zeta_{Nt} + \omega_{Nt}\delta_{Nt} o_{Nt}^{-0.5} v_{Nt} + \omega_{Nt}\sqrt{1-\delta_{Nt}^2}\, o_{Nt}^{-0.5} z_{Nt})
\end{aligned} \tag{74}$$

A.3 Metropolis Hastings Step to draw the mixing variable $\boldsymbol{o}$

In the time varying skewness stochastic volatility models with *Skew-t* shocks, the full conditional distribution of $\{o_t\}_{t=1}^T$ is given by:

$$p(o_t|.) \propto o_t^{\frac{\nu+1}{2}-1} exp\left[-\frac{o_t}{2} \left(\nu_t + \frac{h_t^{-1}(y_t - \boldsymbol{x_t}\boldsymbol{\pi} - \sqrt{h_t}\zeta_t)^2}{\omega_t^2(1-\delta_t^2)} \right) \right] exp\left[\frac{(y_t - \boldsymbol{x_t}\boldsymbol{\pi} - \sqrt{h_t}\zeta_t)(h_t^{-0.5} o_t^{0.5} \delta_t v_t)}{\omega_t(1-\delta_t^2)} \right] \tag{75}$$

Since it is not possible to directly sample from this full conditional distribution, I use Metropolis Hastingss step to draw from this conditional distribution. I use as proposal $Gamma\left(\frac{\nu+1}{2}, \frac{1}{2}\left[\nu + \frac{h_t^{-1}(y_t - \boldsymbol{x_t}\boldsymbol{\pi} - \sqrt{h_t}\zeta)^2}{\omega_t^2(1-\delta_t^2)}\right]\right)$. The acceptance probability in the Metropolis

Hastings step is:

$$p = exp\left[\frac{(y_t - \boldsymbol{x_t\pi} - \sqrt{h_t}\zeta_t)h_t^{-0.5}{o_t^*}^{0.5}\delta_t v_t}{\omega_t(1-\delta_t^2)} - \frac{(y_t - \boldsymbol{x_t\pi} - \sqrt{h_t}\zeta_t)h_t^{-0.5}{o^m}_t^{0.5}\delta_t v_t}{\omega_t(1-\delta_t^2)}\right] \tag{76}$$

where o^* is a new draw from the proposal and o^m is the previous draw.

A.4 Particle Step in the Gibbs Sampler

Table 4 presents the details on the *Particle Step* used in the *Gibbs Sampler* to draw the volatilities and the shape parameters. s_t stands for the generic unobserved latent state being $log(h_t)$ in Step 5) and λ_t in Step 7) of the Gibbs Sampler in Table 1. As anticipated above a valid particle approximation to the Gibbs Sampler requires a Conditional Sequential Monte Carlo update which guarantees that a pre-specified path of the state variables is ensured to survive all the resampling steps (Andrieu et al. 2010). Hence, if I consider a generic iteration $m+1$ of the Gibbs Sampler, when using K particles to approximate $p(h_1, \ldots, h_T|\boldsymbol{\Theta}, \boldsymbol{v}, \boldsymbol{\lambda})$ and $p(\lambda_{i1}, \ldots, \lambda_{iT}|\boldsymbol{\Theta}, \boldsymbol{v}, \boldsymbol{h})$, only $K-1$ particles are generated while the K^{th} particle is set to the pre-specified path $h_{1:T}^{(m)}$ and $\lambda_{1:T}^{(m)}$. In the particle approximation I use the transition equations (2) and (3) as importance densities $g_\theta(s_t)$ and compute the weights accordingly. I refer to the original paper, Lindsten et al. (2014) for the details on the *Ancestral Sampling* step, that for $t > 2$ artificially assign a history to the partial pre-specified path $s_{t:T}^{(m)}$.

Table 4: Particle Step in the Gibbs Sampler

Particle Step with Ancestor Sampling	
Draw $s_1^k \sim g_\theta(s_1)$	for $k = 1, \ldots, K-1$
Set $s_1^K = s_1^{(m)}$	
Compute $w_1^k = W_1(s_1^k)$ and normalize the weights	for $k = 1, \ldots, K$
for $t = 2 : T$	
Re-sampling step: sample $\{s_{t-1}^k\}_{k=1}^K$ with probabilities given by $\{w_{t-1}^k\}_{k=1}^K$	
Draw $s_t^k \sim g_\theta(s_t)$	for $k = 1, \ldots, K-1$
Set $s_t^K = s_t^{(m)}$	
Ancestral sampling step	
Compute $w_t^k = W_t(s_t^k)$ and normalize the weights	for $k = 1, \ldots, K$
end	
Draw j with $Pr(j = k) \propto w_T^k$	

In alternative to the particle step, it can also be considered an independence Metropolis Hastings step to draw the log-volatilities and the shape parameters. In particular, I considered

a log-normal proposal density for the volatility (on the lines of Cogley et al. (2005)) as:

$$q(h_t) \propto h_t^{-1} exp\left[-\frac{(lnh_t - \mu_{h_t})^2}{2\sigma_h^2}\right] \tag{77}$$

with μ_{h_t} and σ_h defined in equations (49) for the *Skew-Normal* case and (60) for the *Skew-t*. The acceptance probabilities in the model with *Skew-Normal* and *Skew-t* shocks are respectively given by:

$$p = \frac{h_t^{*-0,5} exp\left[-\frac{(y_t - \boldsymbol{x_t}\boldsymbol{\pi} - \sqrt{h_t^*}\zeta_t - \sqrt{h_t^*}\omega_t\delta_t v_t)^2}{2h_t^*\omega_t^2(1-\delta_t^2)}\right]}{h^{m_t^{-0,5}} exp\left[-\frac{(y_t - \boldsymbol{x_t}\boldsymbol{\pi} - \sqrt{h_t^m}\zeta_t - \sqrt{h_t^m}\omega_t\delta_t v_t)^2}{2h_t^m\omega_t^2(1-\delta_t^2)}\right]} \tag{78}$$

$$p = \frac{h_t^{*-0,5} exp\left[-\frac{(y_t - \boldsymbol{x_t}\boldsymbol{\pi} - \sqrt{h_t^*}\zeta_t - \sqrt{h_t^*}o_t^{-0.5}\omega_t\delta_t v_t)^2}{2h_t^* o_t^{-1}\omega_t^2(1-\delta_t^2)}\right]}{h^{m_t^{-0,5}} exp\left[-\frac{(y_t - \boldsymbol{x_t}\boldsymbol{\pi} - \sqrt{h_t^m}\zeta_t - \sqrt{h_t^m}o_t^{-0.5}\omega_t\delta_t v_t)^2}{2h_t^m o_t^{-1}\omega_t^2(1-\delta_t^2)}\right]} \tag{79}$$

where h_t^* is the new draw from the proposal distribution, while h_t^m is the previous draw. Instead, for the shape parameters, I considered a Normal proposal:

$$q(\lambda_t) \sim N(\mu_{\lambda_t}, \sigma_\lambda^2) \tag{80}$$

with μ_{h_t} and σ_h defined in equations (54) for the *Skew-Normal* case and (63) for the *Skew-t*. The acceptance probabilities in the model with *Skew-Normal* and *Skew-t* shocks are respectively given by:

$$p = \frac{\omega_t^{*^{-1}}(1-\delta_t^{*^2})^{-0,5} exp\left[-\frac{(y_t - \boldsymbol{x_t}\boldsymbol{\pi} - \sqrt{h_t}\zeta_t^* - \sqrt{h_t}\omega_t^*\delta_t^* v_t)^2}{2h_t\omega_t^{*^2}(1-\delta_t^{*^2})}\right]}{(\omega_t^m)^{-1}(1-\delta_t^{m^2})^{-0,5} exp\left[-\frac{(y_t - \boldsymbol{x_t}\boldsymbol{\pi} - \sqrt{h_t}\zeta_t^m - \sqrt{h_t}\omega_t^m\delta_t^m v_t)^2}{2h_t\omega_t^{m^2}(1-\delta_t^{m^2})}\right]} \tag{81}$$

$$p = \frac{\omega_t^{*^{-1}}(1-\delta_t^{*^2})^{-0,5} exp\left[-\frac{(y_t - \boldsymbol{x_t}\boldsymbol{\pi} - \sqrt{h_t}\zeta_t^* - \sqrt{h_t}o_t^{-0.5}\omega_t^*\delta_t^* v_t)^2}{2h_t o_t^{-1}\omega_t^{*^2}(1-\delta_t^{*^2})}\right]}{(\omega_t^m)^{-1}(1-\delta_t^{m^2})^{-0,5} exp\left[-\frac{(y_t - \boldsymbol{x_t}\boldsymbol{\pi} - \sqrt{h_t}\zeta_t^m - \sqrt{h_t}o_t^{-0.5}\omega_t^m\delta_t^m v_t)^2}{2h_t o_t^{-1}\omega_t^{m^2}(1-\delta_t^{m^2})}\right]} \tag{82}$$

where $\omega_t^*, \zeta_t^*, \delta_t^*$ are functions of the new draw from the proposal λ_t^*, while $\omega_t^m, \zeta_t^m, \delta_t^m$ are λ_t^m are functions of the previous draw λ_t^m.

B Appendix

B.1 Priors and hyper-parameters

Table 5 and Table 6 report the specification of the priors and the choice of the hyper-parameters used for the estimation of the models in the empirical application.

Table 5: Priors for the parameters of the TVSSV model

Parameter	Prior
σ_{ξ}^2	$InverseGamma(5, 0.16)$
σ_{η}^2	$InverseGamma(5, 0.16)$
$\phi_{h,\lambda}$	$\mathcal{N}(1, 0.01)$
β_1	$\mathcal{N}(0, 10)$
π_i	$\mathcal{N}(\underline{\mu_\pi}, \underline{\sigma_{\pi,i}})$
$log(h_0)$	$\mathcal{N}\left(\hat{h}_0, 100\right)$
λ_0	$\mathcal{N}(0, 10)$

$\hat{h}_{i,0}$ is the estimated variance from an AR(4) model to each series using an initial sample of 40 observations. In the application in Section 3 I assume that the elements of $\boldsymbol{\pi}$ namely π_i are centered in zero, namely $\mu_{\pi,i} = 0$ and the variances $\sigma_{\pi,i}$ are set following Carriero et al. (2015). For the VAR I consider the following priors:

Table 6: Priors for the parameters of the VAR TVSSV model

Parameter	Prior
$vec(\mathbf{\Pi})$	$\mathcal{N}(vec(\underline{\boldsymbol{\mu}_{\boldsymbol{\Pi}}}), \underline{\boldsymbol{V}_{\boldsymbol{\Pi}}})$
a_{ij}	$\mathcal{N}(0, 100)$

where the elements of $vec(\underline{\boldsymbol{\mu}_{\boldsymbol{\Pi}}})$ are equal to zero for the coefficients on the cross-equation lags and for the intercept. The coefficients of the own lags are centered in 0 for stationary variables and on 1 for non-stationary variables.

$\underline{V_\Pi}$ has the Minnesota type prior:

$$v_{ij,l} = \begin{cases} \frac{\theta_1}{l^{\theta_4}} & \text{if} \quad i = j \\ \frac{\sigma_i^2 \theta_1 \theta_2}{\sigma_j^2 l^{\theta_4}} & \text{if} \quad i \neq j \end{cases} \tag{83}$$

where I set $\theta_1 = 0.04 \quad \theta_2 = 0.025 \quad \theta_3 = 100 \quad \theta_4 = 2$. We estimate σ_i^2 from univariate AR(12) regressions.

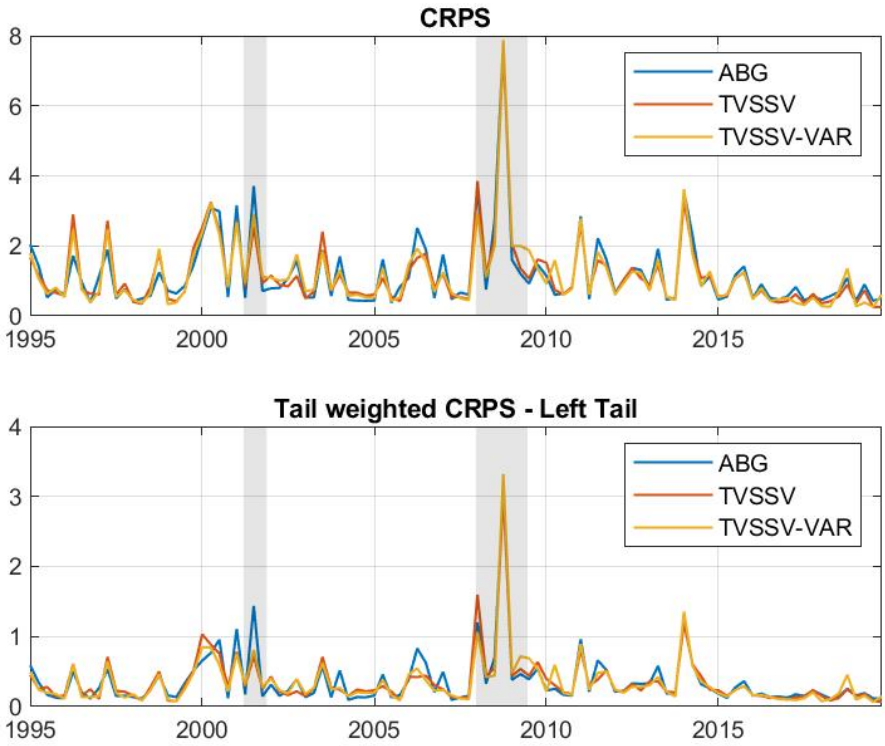
B.2 Variables in the medium scale VAR

Table 7: Variable transformations

Variable	**Transformation**
Real personal consumption expenditures	*log*
Industrial Production	*log*
Unemployment Rate	*level*
Avg Weekly Hours Worked	*log*
Consumer Price Index	*log*
Fed Funds Rate	*level*
10-Year Treasury Yield - Fed Funds Rate	*level*
Moody's Baa Corporate Bond Yield - the Fed Funds Rate	*level*
Standard and Poors index	*log*

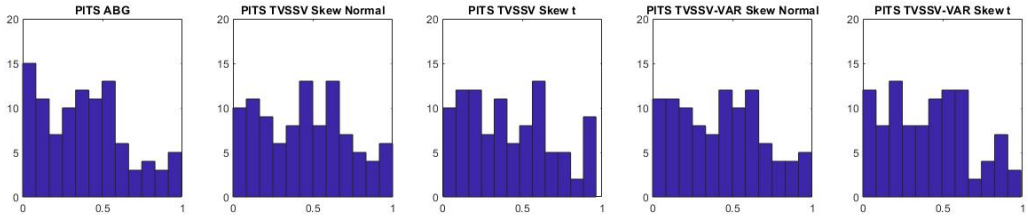
B.3 Other Figures

Figure 6: CRPS and Tail Weighted CRPS (left tail)



Note: The figure above shows the time series of the Cumulative Ranked Probability Scores (CRPS), while the figure below shows the time series of the Left Tail Weighted CRPS (Gneiting et al. 2011). In blue estimates from the two step quantile regression based method by (Adrian et al. 2019), in red from the TVSSV univariate model with *Skew-t* shocks and in yellow the estimates from the TVSSV VAR model.

Figure 7: Probability Integral Transforms



Note: Probability Integral Transforms of the forecasts from the quantile regression based method, the univariate time varying skewness stochastic volatility models with *Skew-Normal* and *Skew-t* shocks, and the VARs with varying skewness and stochastic volatility with *Skew-Normal* and *Skew-t* shocks

Labour at risk*

Vasco Botelho † Claudia Foroni ‡ Andrea Renzetti §

January 22, 2024

Abstract

We propose a Bayesian VAR model with stochastic volatility and time varying skewness to estimate the degree of labour at risk in the euro area and in the United States. We model the asymmetry of the shocks to changes in the unemployment rate as a function of real activity and financial risk factors. We find that the conditional distribution of the changes in the unemployment rate displays time-varying volatility and skewness, with peaks coinciding with the Global Financial Crisis and the COVID-19 pandemic. We take advantage of the multivariate nature of our parametric model to measure stagflation risk defined as the possible joint event of large increases in the unemployment rate and large annual rates of inflation. We find an increasing risk of stagflation for the euro area in 2022 while in the United States stagflation risk increased earlier in 2021 and started decreasing more recently. Notwithstanding the significantly high levels of inflation, stagflation risks have been contained by the resilient performance of the labour market in both areas. Labour at risk is therefore important for the assessment of the inflation-unemployment trade-off.

Keywords: Unemployment risk, Stagflation risk, Labour Market, Bayesian Econometrics.
JEL Codes: C32, C53, E24, E27.

*The views expressed in this paper are those of the authors and do not necessarily reflect the views of the European Central Bank or any other Eurosystem Central Bank. We would like to thank Michele Lenza, Carlos Montes-Galdón, João Tovar Jalles and the seminar participants at the conference "The European Unemployment Problem: Past Trajectory, Present Dilemmas and Future Policies" and at the European Central Bank for the useful comments and suggestions provided. Any errors and omissions are the sole responsibility of the authors.

†European Central Bank.
‡European Central Bank.
§University of Bologna, corresponding author: andrea.renzetti2@unibo.it.

1 Introduction

The cyclical asymmetry of business cycles is a longstanding topic in the economic literature, which can be traced back to Mitchell (1927) and Keynes (1936). Key to this idea is the fact that, on average, contractions in economic activity are briefer and more violent than economic expansions. The asymmetry of cyclical developments is historically salient in the labour market, being identified using both the unemployment rate (Neftçi (1984), DeLong et al. (1986), Falk (1986)) and total employment (McKay et al. (2008), Ferraro et al. (2022)). The number of persons losing their jobs and becoming unemployed rises abruptly during recessions. Conversely, unemployed workers take their time to slowly get back into employment, with the unemployment rate decreasing (or employment increasing) at a slower pace during economic expansions.

More generally, the cyclical asymmetry in the labour market can be related to the assessment of tail risks. These tail risks can be estimated to account for the possible worst case scenarios that could occur in case of economic downturns. Hence, the assessment of tail risks merits the attention of policymakers, who attempt via their policy actions to mitigate some of the welfare losses arising in case a recession occurs and the tail risks are realised.

In this paper, we tackle this topic from a quantitative perspective and estimate the degree of tail risks in the labour market at any given point in time both in the euro area and in the United States. To do so, we propose a fully parametric econometric model with skew-normal shocks featuring both stochastic volatility and stochastic skewness. Our formulation is part of the more general class of Bayesian VAR (BVAR) models with stochastic volatility and time varying skewness discussed in Renzetti (2023). We use this model to extract out-of-sample forecasts that allow us to monitor the degree of "labour-at-risk" and the probability of large increases in the unemployment rate over time. Moreover, we take advantage of the multivariate nature of our parametric model to measure stagflation-risk both the euro area and the US, which we define as the possible joint event of large increases in the unemployment rate and large annual rates of inflation.

There are several channels that could be behind the cyclical asymmetry of the labour market. On the one hand, it is important to quantify the relationship between the cyclical developments in real activity, which can be symmetric or not, and those in the labour market. For example, in the standard "DMP" model with search and matching frictions (Diamond (1982), Mortensen (1982), Pissarides (1985)), shocks to labour productivity can drive the asymmetry in the unemployment

rate via asymmetric fluctuations in the rate of job destruction (Andolfatto (1997)). This implies that even when the business cycles are symmetric, the economy can be faced with sudden and large increases in unemployment during recessions. This could indicate that the Okun's law breaks down during recessions or is at the very least non-linear and state-dependent.

On the other hand, external frictions can simultaneously make both the real output and the labour market to be asymmetric. For example, downward nominal wage rigidities in New Keynesian models inhibit the necessary real wage cuts needed during recessions, thus leading to stronger declines in vacancy posting and employment during downturns than in models featuring symmetric wage adjustment costs (Abbritti et al. (2013)). Also, financial frictions could also induce skewed business cycles, by magnifying the impact of a downturn while leading to a more gradual recovery in a learning model by restricting information after the crisis (Ordoñez (2013)).[1]

We embed these mechanisms in our BVAR model by considering a specification at the monthly time frequency with the changes in the unemployment rate, a variable proxying for changes in real activity (the PMI output for the euro area and the CFNAI for the US), and a measure proxying for changes in financial conditions (the CISS indicator for the euro area and the NFCI for the US). By doing so, we allow the unemployment rate to move either driven by shocks specific to the labour market, as a response to shocks to real activity in a standard Okun's law framework, or as a response to shocks to financial conditions.

We allow for the skewness of the shocks specific to the labour market to be state dependent and change as a function of real activity and financial conditions. In simple terms, the labour market is more likely to be faced with strong adverse shocks that increase the unemployment rate substantially when the economy is in a bad state of the world. This bad state of the world can vary between a very sharp slow down of the economy or a prolonged recession for real activity, or instead a strong tightening of financial conditions. This implies that during bad states of the world, shocks to the monthly changes in the unemployment rate become asymmetric and right-skewed, meaning that large increases in the unemployment rate become more likely in these periods than in good times. By allowing the distribution of the shocks to the unemployment rate to shift as a function of (lagged) real activity and financial risk factors, the model is in this way able to capture possible non-linear relationships between real activity and the labour

1. Regarding alternative mechanisms we highlight, for example, the role of worker-level heterogeneity in models with search and matching frictions in matching the asymmetric response of the employment rate to shocks over the business cycle (Ferraro (2018)). We notice also that micro-level evidence points towards firms following concave hiring rules, making them slower to hire after good shocks and quicker to fire after bad shocks (Ilut et al. (2018)).

market, and to cater for the role of financial frictions. We find that the skewness of the shocks to the labour market is on average positive in both economies and more time-varying in the euro area, increasing further and often peaking in economic downturns as financial conditions deteriorate and economic activity plunges.

Our paper is related to the growing strand of literature aimed at assessing and quantifying tail risk to macroeconomic outcomes. Methodologically, this literature traditionally relies on quantile regression based methods, such as those used by Giglio et al. (2016) and Adrian et al. (2019). Most of the literature on the assessment of tail risks to macroeconomic outcomes focus uniquely on output growth and inflation. Considerably less attention has been devoted to the analysis of tail risks in the labour market. The exception is Kiley (2022), who provides an assessment of tail risks to the US unemployment rate using quantile regressions.

We depart away from quantile regressions and propose a fully parametric model to assess tail risks in the labour market. In this sense, our work is more closely linked to a recent strand in the literature that proposes the usage of fully parametric models to assess and predict tail risks, as in López-Salido et al. (2020), Plagborg-Møller et al. (2020), Carriero et al. (2020a), Carriero et al. (2020b), Delle Monache et al. (2021), Brownlees et al. (2021), Wolf (2021), Iseringhausen (2023), or Montes-Galdón et al. (2022), to provide some examples. Similarly to our own contribution, this strand of literature argues that fully parametric models are more flexible than quantile regression based methods, while simultaneously achieving a similar forecasting performance.[2] We exploit our model to further extend our analysis and incorporate also information on inflation rates to analyze labour market risk and inflation risk jointly. In the specific, we use the augmented setup to assess the joint probability for the economy to experience simultaneously large increases in the unemployment rate and large annual inflation rates. We find an increasing risk of stagflation for the euro area in 2022 while in the United States stagflation risk increased earlier in 2021 and started decreasing more recently. Notwithstanding the significantly high levels of inflation, stagflation risks have been contained by the resilient performance of the

2. It should be noted that both Wolf (2021) and Iseringhausen (2023) employ also stochastic volatility models with asymmetric shocks and stochastic skewness. However, both models are univariate. Our model is multivariate and allows to jointly model the dynamic relationship between the risk factors and the target variables. As well the multivariate nature of our model allows us to assess tail risk to multiple target variables (unemployment rate and inflation). Relatedly, Chavleishvili et al. (2019) extend the quantile regression based methods to a quantile VAR (QVAR) model framework. This methodology allows to model the interaction between any quantiles of the endogenous variables. The increased flexibility comes at the cost of highly complex modelling assumptions and estimation algorithm. For this reason, we leave the comparison of the forecasting performance and labour-at-risk measurement between our application and a QVAR model for the labour market for future research.

labour market in both areas. When using the model for forecasting we find a similar result in our application to the euro area or US labour markets. In particular, our BVAR with stochastic volatility and time varying skewness displays a forecasting performance that is at least as good as, but often superior in terms of density forecast accuracy to other benchmarks including the quantile regression based methods from Adrian et al. (2019), a standard BVAR model, and a BVAR model with stochastic volatility.

This paper is organised as follows. Section 2 describes the data we use and the main features of the unemployment rate, which motivate our modelling choice to assess tail risks. Section 3 describes our BVAR model with stochastic volatility and time varying skewness, and our application of this model to tackle the cyclical asymmetry in the labour market. Section 4 presents most of our empirical results, focusing first on the in-sample assessment of the model, then on the construction of our "labour-at-risk" measure, and finally on the out-of-sample forecasting performance of the model. Section 5 extends our BVAR model application to assess probabilities and risk of joint events, applied to measuring the risk of stagflation. Section 6 concludes.

2 Data description

We work with monthly data from January 1999 to September 2022 for the euro area and from January 1971 to September 2022 for the United States. For both regions, we use the unemployment rate to account for the conditions of the labour market and two risk factors: (i) a monthly indicator for real economic activity, which allows us to account for the reaction of the labour market as a response to changes in economic activity; and (ii) a monthly indicator for financial conditions, which we use to cater for the role of financial frictions in shaping movements in real activity and in the labour market.

For the euro area, the unemployment rate is obtained from the ECB Statistical Data Warehouse. We use the Purchasing Manager Index (PMI) Output from S&P Global to proxy for the developments in real activity.[3] This indicator tracks the assessments by corporate executives regarding the immediate reactions of their firms to idiosyncratic events over time, and contain in this way information on the cyclical developments in real activity. As proxy for financial

3. As a robustness check, we consider instead the Eurocoin indicator provided by the Bank of Italy as a proxy for real activity. From a qualitatively point of view our results do not change. For space considerations, we do not report the results in the text, although they can be provided upon request. We decided to use the PMI output as it allows for a better forecasting performance to our empirical application.

conditions we use the Composite Indicator of Systemic Stress (CISS), which is the most widely used indicator for financial conditions in the euro area in the literature assessing tail risks to macroeconomic outcomes (Kremer et al. (2012), Figueres et al. (2020)). This indicator aims to track the current level of frictions, stresses and strains in the financial system. It is available at daily frequency and we obtain a monthly series by averaging over the daily observations.

For the United States, we calculate the unemployment rate in two steps. First, we construct the number of unemployed workers by obtaining the difference between the civilian labour force and the civilian employment. We then construct the unemployment rate as the ratio between this measure of unemployment and the civilian labour force.[4] To proxy for the developments in real economic activity, we use the National Activity Index provided by the Federal Reserve of Chicago (CFNAI), which compares the growth rate of the economy to its historical trend. To gauge the state of financial conditions in the US, we borrow from the growth-at-risk literature, and use the National Financial Conditions Index (NFCI) from the Federal Reserve of Chicago, which is comparable in scope to the CISS indicator for the euro area. Both the CFNAI and the NFCI indicators are available at a weekly time frequency. Therefore, we aggregate the weekly observations each month by taking their average over the month.

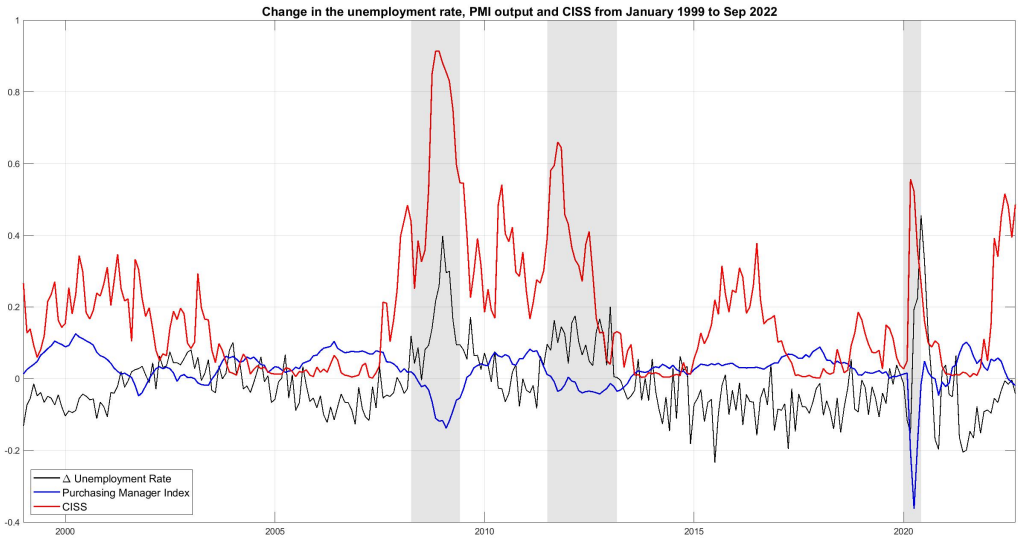
Figure 1 for the euro area and Figure 2 for the US show the time series of the month on month changes in the unemployment rate together with the selected real activity and financial conditions indicators. Both in the euro area and in the US, as the economy enters into the recession, financial conditions indicators rise first and the real activity indicators plunge, with the unemployment rate displaying sharp and fast increases. In the next paragraph we describe in more detail the (a)symmetry of the monthly changes in the unemployment rate over the business cycle in both areas.

2.1 Asymmetry of changes in the unemployment rate

The unemployment rate in the euro area decreased by almost four percentage points between April 1998 and September 2022. However, the long-term decline in the unemployment rate did not happen at a constant pace over time. The unemployment rate tends to decrease during expansions and to increase during recessions with the speed of changes in the unemployment rate depending on the state of the business cycles. In Table 1, we show the peak-to-trough

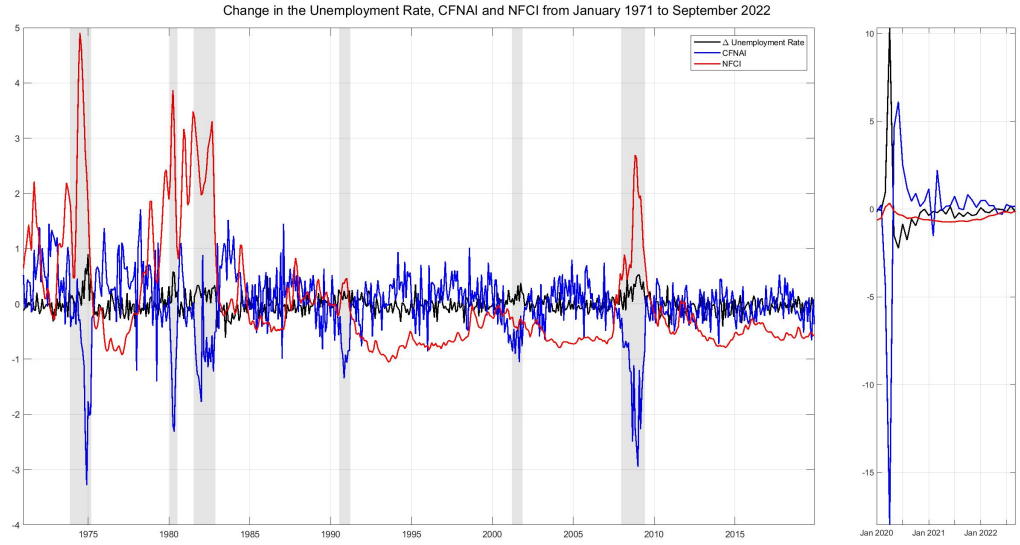
4. We obtain data on civilian employment and labor force from the FRED-MD database, provided by the Federal Reserve Bank of St. Louis. The source code for the civilian employment is LNS12000000 (or alternatively CE16OV), while for the civilian labour force is LNS11000000 (or alternatively CLF16OV).

Figure 1: MoM changes in the unemployment rate and PMI output and CISS



Notes: The figure shows the time series of the month-on-month changes in the unemployment rate, toghether with the PMI Output (minus 50 divided by 100) and the CISS. The shadow bands are for the EACN recessions periods.

Figure 2: MoM changes in the unemployment rate and CFNAI and NFCI



Notes: The figure shows the time series of the month-on-month changes in the unemployment rate, toghether with the CFNAI and the NFCI. The shadow bands are for the NBER recessions periods.

Table 1: Unemployment rate in the EA and the US - peak-to-trough and trough-to-peak changes

Region	Peak	Trough	Months	Peak-to-Trough *(per month)*	Trough	Peak	Months	Trough-to-Peak *(per month)*
EA					Apr-98 (*)	Mar-08	119	-0.03
	Mar-08	Jun-09	15	0.16	Jun-09	Sep-11	27	0.03
	Sep-11	Mar-13	18	0.10	Mar-13	Dec-19	81	-0.06
	Dec-19	Jun-20	6	0.09	Jun-20	Sep-22 (**)	27	-0.05
US	Nov-48	Oct-49	11	0.38	Oct-49	Jul-53	45	-0.12
	Jul-53	May-54	10	0.33	May-54	Aug-57	39	-0.05
	Aug-57	Apr-58	8	0.41	Apr-58	Apr-60	24	-0.09
	Apr-60	Feb-61	10	0.17	Feb-61	Dec-69	106	-0.03
	Dec-69	Nov-70	11	0.21	Nov-70	Nov-73	36	-0.03
	Nov-73	Mar-75	16	0.23	Mar-75	Jan-80	58	-0.04
	Jan-80	Jul-80	6	0.26	Jul-80	Jul-81	12	-0.05
	Jul-81	Nov-82	16	0.22	Nov-82	Jul-90	92	-0.06
	Jul-90	Mar-91	8	0.16	Mar-91	Mar-01	120	-0.02
	Mar-01	Nov-01	8	0.16	Nov-01	Dec-07	73	-0.01
	Dec-07	Jun-09	18	0.25	Jun-09	Feb-20	128	-0.05
	Feb-20	Apr-20	2	5.64	Apr-20	Sep-22 (**)	29	-0.39

Notes: Peaks and trough dates are taken from the CEPR for the euro area, and the months are adjusted to the end of the quarter announced by the CEPR as the relevant business cycle date, and from the NBER for the US. (*) our data for the unemployment rate in the euro area starts in April 1998. It is therefore not a trough, but our earliest comparison to the March 2008 peak. (**) our data stops for both the euro area and the US in September 2022. This is the latest comparison we have for the latest trough in each region.

(downturns) and trough-to-peak (expansions) changes in the unemployment rate in the euro area. In expansions, the unemployment rate decreased mildly by between -0.03 and -0.06 percentage points per month. The euro area unemployment rate even increased slightly in the expansion between the end of the Global Financial Crisis in 2009 Q2 and the start of the Sovereign Debt Crisis in 2011 Q3. By contrast, the unemployment rate increased at a more pronounced pace during recessions, rising on average by between 0.09 and 0.16 percentage points per month. This pattern also implies that the unconditional distribution of the changes in the euro area unemployment rate is skewed to the right (see Figure 3).

The longer time series data for the United States shows a more cyclical unemployment rate than that observed for the euro area. By September 2022, the US unemployment rate was at levels broadly comparable to those observed in 1948. However, the US unemployment rate displays a similar cyclical asymmetry to that of the euro area. Economic expansions are linked with small average decreases in the unemployment rate on a magnitude of -0.06 percentage points per month. Not considering the current expansion following the COVID-19 pandemic, this average monthly decrease in the unemployment rate would be even milder, at -0.04 percentage

Figure 3: Unconditional distribution of month-on-month changes in the unemployment rate in the euro area

Notes: The figure shows the histogram and the estimated distribution of the month-on-month changes in the unemployment rate from January 1999 to September 2022 in the euro area.

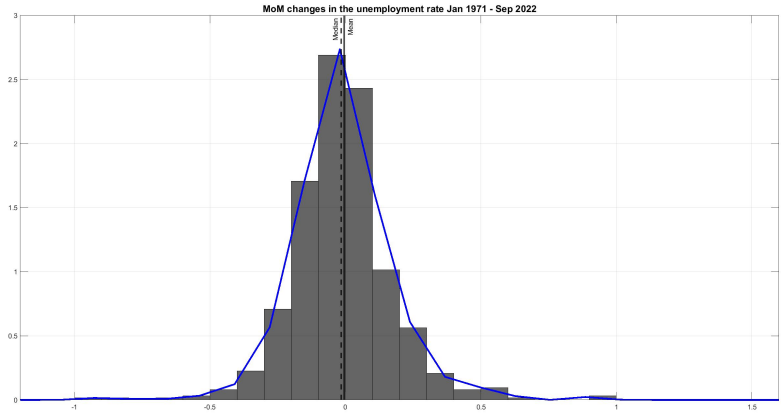
points per month. By contrast, the US unemployment rate increases considerably faster during contractions at an average of 0.34 percentage points per month, or at 0.25 percentage points when we exclude the COVID-19 pandemic contraction. These results also imply that the unconditional distribution of the month-on-month changes in the US unemployment rate is markedly skewed to the right (see Figure 4), similarly to that observed for the euro area.

Hence, in both regions downturns are considerably shorter than upturns. These results taken together confirm that contractions are briefer and more violent than expansions in the euro area and the US labour markets. They also imply that the risks around the unemployment rate outlook are not symmetric, as the likelihood of large increases in the unemployment rate has been historically higher than the likelihood of large decreases. The steepness of the monthly changes in the unemployment rate is also visible when we extract the residuals from an autoregressive model regression to account for potential dynamic or slower movements in the unemployment rate over time. Both in the euro area and in the US, the symmetry of the residuals changes over time, with the residuals more right skewed during recessions than during expansions.[5]

In more technical terms, we use the normality test proposed by Bai et al. (2005) to assess the skewness of the changes in the unemployment rate in both the euro area and in the US. [6] We

5. See Figure 23, 24, 25, 26 in the Appendix A.6.
6. Details on how the statistic is computed can be found in the Appendix A.1

Figure 4: Unconditional distribution of month-on-month changes in the unemployment rate in the US



Notes: The figure shows the histogram and the estimated distribution of the month-on-month changes in the unemployment rate from January to September 2022 in the US.

Table 2: Bai et al. (2005) skewness statistic

	MoM changes	QoQ changes	YoY changes	Level
Euro Area	1.7884*	1.704*	1.5888	1.3458
US	0.9908	1.0261	1.0165	2.1551**

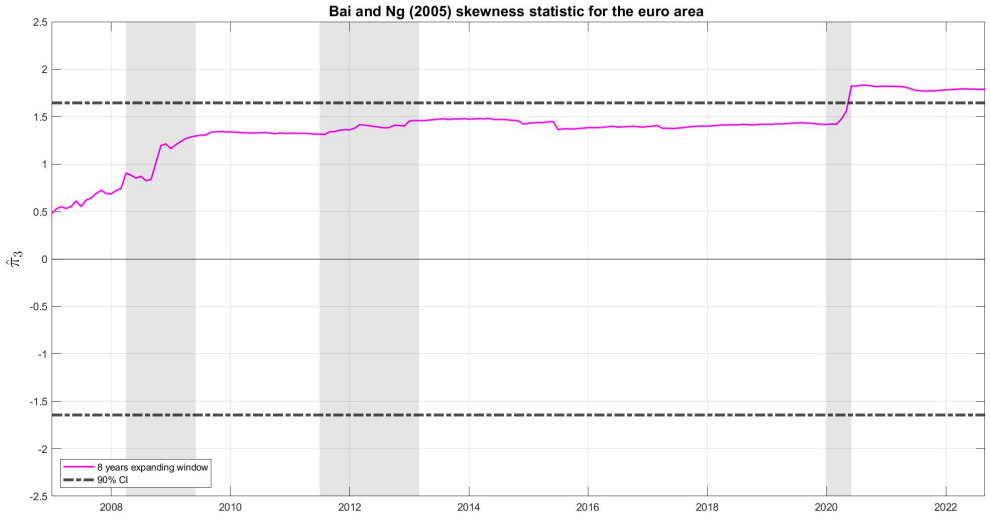
Notes: Bai et al. (2005) skewness statistic. * indicates rejections of the null hypothesis at 90% CI and ** indicates rejections of the null hypothesis at 95% CI.

calculate this statistic for the month-on-month, quarter-on-quarter, and year-on-year changes in the unemployment rate, and also to the level of the unemployment rate. We report the results of this test in Table 2 for our final sample from January 1999 to September 2022 for the euro area and from January 1971 to September 2022 for the United States.

The skewness statistic from Bai et al. (2005) rejects the normality of both the month-on-month and quarter-on-quarter changes in the unemployment rate in the euro area, and it rejects the normality of the unemployment rate levels in the US. Both results highlight that there is cyclical asymmetry in the unemployment rate. For the United States, the adjustment of the unemployment rate during the COVID-19 pandemic and ensuing recovery exhibited a set of outliers that shifted the skewness statistic. When applied to the sample ending in December 2019, we find also a statistically significant skewness for the month-on-month changes in the US unemployment rate.

More generally, we find that the skewness of the changes in the unemployment rate is time-

Figure 5: Bai et al. (2005) skewness statistic in the euro area

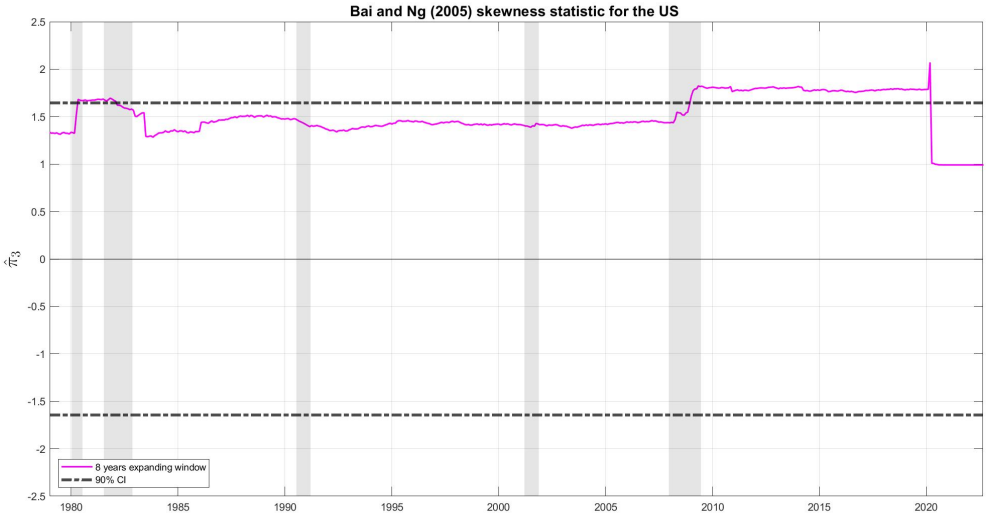


Notes: The figure shows the time series of the Bai et al. (2005) skewness statistic for the month-on-month changes in the unemployment rate in the euro area computed using expanding recursive windows of 8 years. The dashed horizontal black lines shows the 90% confidence interval.

varying in both the euro area and the US. To showcase this, we compute the skewness statistic from Bai et al. (2005) for the distribution of the month-on-month changes in the unemployment rate on an expanding recursive window that comprises the first eight years of data for each region and adds one month of data at the time before calculating again the skewness statistic. The skewness statistics for the expanding windows are reported for the euro area in Figure 5 and for the United States in Figure 6. In September 2022, the expanding window covers the full sample and mimics the results in Table 2. For the euro area, the skewness statistic started very low and increases to a high level close to the 90% confidence band over the entire sample. Skewness becomes statistically significant in the euro area with the onset of the COVID-19 pandemic, with the skewness statistics increasing outside the 90% confidence interval.

The time variation in the skewness of the month-on-month changes in the unemployment rate is even more prevalent in the US. The skewness statistic is statistically significant during the period comprising the oil prices shocks and the Volcker's recession in the late 1970s and early 1980s, and then again following the Global Financial Crisis. The COVID-19 pandemic shifted the skewness statistic because of the unique timing and magnitudes of the monthly changes in the unemployment rate. As shown in Table 1, the unemployment rate increased by 5.6 percentage points per month between February 2020 and April 2020 and it decreased at a

Figure 6: Bai et al. (2005) skewness statistic in the US



Notes: The figure shows the time series of the Bai et al. (2005) skewness statistic for the MoM changes in the unemployment rate in the United States computed using expanding recursive windows of 8 years. The dashed horizontal black lines shows the 90% confidence interval.

faster pace from May 2020 onwards over a relatively longer period of time. These results mark the importance of catering for the time-variation in the skewness of the unemployment rate in a time series modelling approach and provides the main motivation for our empirical model in the next section.

3 Model

3.1 Time varying skewness stochastic volatility VAR model

The time varying asymmetry of the unemployment rate leads to time varying tail risks that should be considered by any model that attempts to predict any future dynamics in the labour market. In this section we propose a fully parametric econometric model to provide a meaningful characterisation of tail risk in the labour market both the euro area and the United States.

Our model is a standard VAR model equipped with stochastic volatility and skew normal shocks.[7] These additional features are designed to capture both shifts in the volatility and in the skewness of the shocks to the endogenous variables in the VAR. The exact specification is:

7. This class of models is comprehensively covered in Renzetti (2023), with an application for the "growth-at-risk" literature.

$$\boldsymbol{y_t} = \boldsymbol{\Pi_0} + \boldsymbol{\Pi_1 y_{t-1}} + \ldots + \boldsymbol{\Pi_p y_{t-p}} + \boldsymbol{A}^{-1}\boldsymbol{H}_t^{0.5}\boldsymbol{\varepsilon_t} \tag{1}$$

$$\varepsilon_{it} \sim \textit{Skew normal}(\zeta_{it}, \omega_{it}, \lambda_{it})$$

where $\boldsymbol{y}_t$ is the vector of endogenous variables, observed for $t = 1, \ldots, T$ periods (months) and $i = 1, \ldots, N$ with N being the number of endogenous variables. As endogenous variables, we consider the month-on-month changes in the unemployment rate, a monthly real activity indicator, the PMI output for the euro area and the CFNAI for the US, and a monthly indicator for financial conditions, the CISS indicator for the euro area and the NFCI for the US.

The matrix $\boldsymbol{A^{-1}}$ is a lower triangular matrix with ones on the main diagonal, $\boldsymbol{H_t} = diag(h_{1,t}, \ldots, h_{N,t})$ is the diagonal matrix collecting the volatilities of the shocks and $\boldsymbol{\varepsilon_t}$ is a column vector collecting the skew normal (as in Azzalini (1986)) shocks ε_{it}.

In general the shape parameter λ_{it} shifts both the mean and the variance of skew normal distribution. To interpret ε_{it} as structural shocks in the VAR model, we then re-parametrize the skew normal distribution parameters ζ_{it} and ω_{it} such that $\mathbb{E}[\boldsymbol{\varepsilon_t}] = \boldsymbol{0_N}$ and $\mathbb{E}[\boldsymbol{\varepsilon_t}\boldsymbol{\varepsilon_t'}] = \boldsymbol{I}_N$. That is, we ensure that the elements in $\boldsymbol{\varepsilon_t}$ are unpredictable in terms of their mean and that they have unit variance. This ensures as well that the elements on the diagonal matrix $\boldsymbol{H_t}$ provide the sufficient information on the variances of the shocks while the shape parameters λ_{it} carry sufficient information on the skewness of the shocks. In more detail, this re-parameterization implies the following constraints on the location and scale parameters of the shocks:

$$\zeta_{i,t} = -\omega_{i,t}\delta_{i,t}\sqrt{\frac{2}{\pi}} \qquad \forall i, t \tag{2}$$

$$\omega_{i,t}^2 = \left[1 - \frac{2}{\pi}\delta_{i,t}^2\right]^{-1} \qquad \forall i, t \tag{3}$$

where $\delta_{i,t} = \frac{\lambda_{i,t}}{\sqrt{1+\lambda_{i,t}^2}}$, with $-1 < \delta_{i,t} < 1$. The re-parameterized skew normal shocks are identified and interpreted as structural shocks assuming the short run restrictions implied by the ordering of the variables in the model (Cholesky identification).[8] We order the financial conditions indicator last, so as to allow financial markets to adjust within the month to shocks to real activity and to the labour market. The unemployment rate is ordered second and it is allowed

8. It is worth to mention that as discussed in Primiceri (2005) and recently outlined in Arias et al. (2021) the ordering of the variables in this model matters not only for the identification of the shocks but also for estimation. This occurs because the Normal prior on the free elements of the lower triangular matrix $\boldsymbol{A}$ induces *a priori* an asymmetric prior for the variance-covariance matrix of the reduced-form residuals.

to adjust within the month to shocks to the labour market and real activity, but not to shocks to financial conditions. This follows the standard Okun's law, which relates real activity and the labour market. Finally, we order the real activity indicator first, assuming that shocks to the labour market and to financial conditions affect the real activity indicator with one month of lag.

To capture changes over time in the size of the shocks, the log-volatilities are assumed to be independent stochastic processes which evolve over time according:

$$log(h_{i,t}) = log(h_{i,t-1}) + \eta_{i,t} \qquad \eta_{i,t} \sim \mathcal{N}(0, \sigma^2_{i,\eta}) \tag{4}$$

where $i = \{PMI, \Delta U, CISS\}$ for the euro area and $i = \{CFNAI, \Delta U, NFCI\}$ for the US, and where ΔU stands for the month-on-month changes in the unemployment rate. [9]

On the other hand, to capture changes in the symmetry of the shocks, it is assumed that the shape parameters λ_{it} are another set of independent stochastic processes with their own dynamics. In general, positive (negative) values of $\lambda_{i,t}$ are associated to right (left) skewed shocks and right (left) skewed shocks decrease the likelihood of left (right) tail events while correspondingly increase the likelihood of right (left) tail events. For example, when the shape parameters of the shocks to the labour market is positive, namely $\lambda_{\Delta U,t} > 0$, the labour market shock is skewed to the right, and large increases in the unemployment rate become more likely as a consequence of this shock. We consider a specification in which the endogenous variables on real activity and financial conditions can be thought as risk factors affecting the skewness of the shocks to the labour market. In intuitive terms, this implies that the monthly changes to the unemployment rate are more likely to be hit by adverse right skewed shocks when the state of the economy is weak, either via a bad performance of real activity or a strong tightening of financial conditions. In practice, we assume that these risk factors provide information on the evolution of the shape parameter of shocks to the unemployment rate over time, following:

$$\lambda_{\Delta U,t} = \phi_{1,\Delta U}\lambda_{\Delta U,t-1} + \boldsymbol{\phi_2 x_{t-1}} + \xi_{\Delta U,t} \qquad \xi_{\Delta U,t} \sim \mathcal{N}(0, \sigma^2_{\xi,\Delta U}) \tag{5}$$

where $\boldsymbol{x_{t-1}}$ is a vector that includes a constant and the lagged risk factors, being $\{PMI_{t-1}, CISS_{t-1}\}$ for the euro area and $\{CFNAI_{t-1}, NFCI_{t-1}\}$ for the US. This setup makes the monthly changes in the unemployment rate to be our target variable and enables to capture persistency and state

9. CISS, NFCI and CFNAI are demeaned, while the PMI is rescaled subtracting 50 from the original value.

dependence in the shape of the shocks to the unemployment rate in connection to past developments in real activity and in the financial conditions.[10] The coefficients in the vector $\boldsymbol{\phi_2}$ determine the relationship between the risk factors and the shape of the shocks. When this coefficient is positive, an increase in the risk factors in associated to an increase of the skewness of the shock, that is to a shift of the conditional quantiles of the change in the unemployment rate to the right; conversely, when this coefficient is negative, increases of the risk factors are associated to a decrease in the skewness of the shocks, that is, to a shift of the conditional quantiles of the change in the unemployment rate to the left.

In their turn, in the model, shocks to real activity and to financial conditions are allowed to be potentially asymmetric with the degree of asymmetry changing over time. In particular, we assume that the shape parameters of the shocks to the real activity and financial conditions indicators follow independent AR(1) stochastic processes

$$\lambda_{i,t} = \phi_{1,i}\lambda_{i,t-1} + \xi_{i,t} \qquad \xi_{it} \sim \mathcal{N}(0, \sigma^2_{\xi,i}) \tag{6}$$

for $i = \{PMI, CISS\}$ in the euro area and $i = \{CFNAI, NFCI\}$ for the US.

Given that the risk factors influence not only the conditional mean of the changes in the unemployment rate through equation (1) but also the conditional skewness via equation (5), this model can capture the potential non-linear effects of real activity and financial conditions on the unemployment rate, similarly to the quantile regression framework by Kiley (2022). However our approach displays some advantages with respect to the methods based on univariate quantile regressions. First, our Bayesian VAR model allows to properly address the rich autocorrelation structure of macroeconomic time series and to exploit a potentially richer information set that is coherent over time, while this is in general more difficult in the univariate quantile regression framework. Second, we can obtain the entire predictive distribution for the changes in the unemployment rate in a single step, without the need of relying on quantile interpolation methods as it is done in the "at-risk" literature following the paper of Adrian et al. (2019). Third, we can jointly model the multivariate dynamics of all endogenous variables in our model, which in the case of our application are the changes the unemployment rate, the indicator for real activity, and the indicator for financial conditions. Quantile regression based methods focus instead on one target variable in a partial equilibrium without accounting explicitly for general

10. It is important to notice that in our specification the risk factors do not enter the variance of the shocks, thus only moving the asymmetric shape of the distribution of shocks.

equilibrium feedback effects over time. Fourth, our flexible structure allows us to assess tail risks simultaneously in multiple target variables. We will develop on this in Section 5, where we use an augmented version of our model to quantitatively assess the risk of stagflation in the euro area and in the US. Finally, a more general advantage of our BVAR methodology is that even by adding time varying skewness to the model we are still able to use all the standard tools used for policy evaluation and scenario analysis that are traditionally used in the VAR literature, which are not readily available to univariate quantile regression based methods.[11]

3.2 Priors and estimation

The equations of the time varying skewness stochastic volatility VAR model form a non-linear and non-Gaussian state space model. The model is estimated using the Gibbs sampler algorithm described by Renzetti (2023). The estimation strategy goes around the potential difficulties arising because of non-Gaussianity of shocks by leveraging on the normal mixture representation of the skew normal distribution with a truncated normal used as the mixing distribution and by exploiting the triangularization of the VAR system as in Carriero et al. (2019) to draw the parameters and the mixing variables equation by equation.[12] The Gibbs sampler algorithm includes particle steps to approximate the full conditional posterior distributions of the log-volatilities and the shape parameters.[13] The transition equations of the latent states are used as importance densities in the particle steps. [14]

As for the specification of the prior distribution for the parameters of the model, we specify a *Normal* prior for the autoregressive coefficients stored in the matrices $\mathbf{\Pi_j}$ with $j = 0, \ldots, p$ with Minnesota type variance covariance matrix (Litterman (1986)). Following Cogley et al. (2005), we specify a Normal prior for the free elements in the matrix $\boldsymbol{A}$. We also assume a Normal

11. Multivariate quantile regression based methods (as in Chavleishvili et al. (2019)) mitigate a large share of the disadvantages faced by univariate quantile regressions. The main difference between our model and a QVAR is that we use a simpler and fully parametric structure that allows us to extend the model in a more straightforward way that is also more efficient from a computational perspective.

12. See Appendix A.2 and in particular equation (13) for further details on the mixture representation.

13. It is worth to remark that in addition to referring to a multivariate model, the estimation algorithm is conceptually different from Wolf (2021) since the algorithm used in this paper leverages on the normal mixture representation of the skew normal random variable as detailed in the Appendix A.2. In particular, the estimation procedure exploits the fact that conditionally on the vector of mixing variables $\boldsymbol{v}_t$, on the elements of the diagonal matrix $\mathbf{\Delta_t}$ (which are one to one map to the shape parameters $\lambda_{1,t}, \ldots, \lambda_{N,t}$) and on the log-volatilities the state space is Gaussian. Further details on this difference can be found in Renzetti (2023).

14. In order to alleviate path degeneracy in the underlying conditional sequential Monte Carlo sampler, we exploit the ancestor sampling procedure that enables a fast mixing even when using seemingly few particles. The ancestor sampling procedure was developed by Lindsten et al. (2014), who provide a formal proof for the convergence of the algorithm and a comprehensive study on the properties of the sampler.

prior for the initial state of the log-volatilities $log(h_{i,0})$, for the shape parameters $\lambda_{i,0}$, for the coefficients in the state equations of the shape parameters $\phi_{1,i}$ and for the elements of $\boldsymbol{\phi_2}$. Finally, we assume an Inverse Gamma prior for the variance of the innovations in the state equations of the the log-volatilities and the shape parameters . Table 6 in Appendix A.3 summarises our choices on the priors and relevant hyperparameters.

The steps of the Gibbs sampler to simulate draws from the joint posterior distribution of the parameters $\boldsymbol{\Theta}$, the latent states $\boldsymbol{s}$ (i.e., the log-volatilities and the shape parameters), and the mixing variables $\boldsymbol{v}$, are as follows:

1. Draw the path for the mixing variables $\{\boldsymbol{v_{it}}\}_{t=1}^{T}$ from $p(\boldsymbol{v_{i1}} \ldots, v_{iT}|\boldsymbol{\Theta}, \boldsymbol{s}, \boldsymbol{Y})$ for $i = 1, \ldots, N$.

2. Draw the VAR coefficients $\boldsymbol{\Pi}$ from $p(\boldsymbol{\Pi}|\boldsymbol{\Theta}, \boldsymbol{v}, \boldsymbol{s}, \boldsymbol{Y})$. The coefficients are drawn equation by equation exploiting the triangular algorithm developed in Carriero et al. (2019).

3. Draw the free elements in the lower triangular matrix $\boldsymbol{A}$ from $p(\boldsymbol{A}|\boldsymbol{\Theta}, \boldsymbol{v}, \boldsymbol{s}, \boldsymbol{Y})$.

4. Draw the variances in the state equations of the shape parameters $\sigma^2_{\xi,i}$ from $p(\sigma^2_{\xi,i}|\boldsymbol{\Theta}, \boldsymbol{s}, \boldsymbol{v}, \boldsymbol{Y})$ for $i = 1, \ldots, N$.

5. Draw the autoregressive coefficients in the state equations of the shape parameters $\phi_{1,i}$ from $p(\phi_{1,i}|\boldsymbol{\Theta}, \boldsymbol{s}, \boldsymbol{v}, \boldsymbol{y})$ for $i = 1, \ldots, N$. and the coefficients of the risk factors in the state equation of the shape parameter of the target variable from from $p(\boldsymbol{\phi_2}|\boldsymbol{\Theta}, \boldsymbol{s}, \boldsymbol{v}, \boldsymbol{Y})$.

6. Draw the variances in the state equations of the log-volatilities $\sigma^2_{\eta,i}$ from $p(\sigma^2_{\eta,i}|\boldsymbol{\Theta}, \boldsymbol{s}, \boldsymbol{v}, \boldsymbol{Y})$ for $i = 1, \ldots, N$.

7. Draw the initial states for the volatilities $h_{i,0}$ from $p(h_{i,0}|\boldsymbol{\Theta}, \boldsymbol{v}, \boldsymbol{s}, \boldsymbol{Y})$ for $i = 1, \ldots, N$.

8. Draw the initial states for the shape parameters $\lambda_{i,0}$ from $p(\lambda_{i,0}|\boldsymbol{\Theta}, \boldsymbol{v}, \boldsymbol{s}, \boldsymbol{Y})$ for $i = 1, \ldots, N$.

9. Draw the path of the shape parameters $\{\lambda_{it}\}_{t=1}^{T}$ from $p(\lambda_{i1}, \ldots, \lambda_{iT}|\boldsymbol{\Theta}, \boldsymbol{v}, \boldsymbol{s}, \boldsymbol{Y})$ for $i = 1, \ldots, N$ using the particle approximation.

10. Draw the path of the volatilities $\{h_{it}\}_{t=1}^{T}$ from $p(h_{i1}, \ldots, h_{iT}|\boldsymbol{\Theta}, \boldsymbol{v}, \boldsymbol{s}, \boldsymbol{Y})$ for $i = 1, \ldots, N$. using the particle approximation.

The Markov Chain Monte Carlo (MCMC) algorithm consists of 50,000 draws, with the initial 30,000 draws discarded as burn-in. In the particle steps, we use 100 particles to approximate the full conditional posterior distribution of the volatilities and 150 particles to approximate the full conditional posterior distribution of the shape parameters.

4 Results

4.1 In-sample analysis

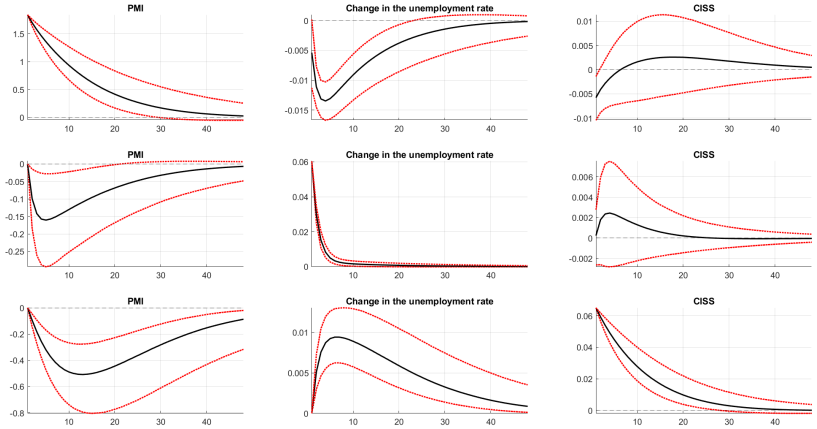
We start by examining the in-sample performance of our time varying skewness and stochastic volatility (TVSSV) VAR model. The starting point is to assess the estimated impulse response functions (IRFs) of our model. In particular, for both the euro area and the US, the IRFs identify that an expansionary shock to real activity implies a decrease in the unemployment rate, while not having a strong impact on the financial conditions. Similarly, a shock to the labour market that increases the unemployment rate has no impact on financial conditions but induces instead a feedback loop that decreases real activity. Finally, a shock that tightens the financial conditions induces a slow down in real activity and increases the unemployment rate. The IRFs unveil that the model is able to capture important relationships such as the Okun's law, in which improvements in real activity are associated with declines in the unemployment rate, and the fact that the tightening of financial conditions induces on average a slowdown in real activity and an increase in the unemployment rate.

In the model the likelihood of good and large shocks versus bad and large shocks is state dependent and changes over time as a function of real activity and financial conditions. This reflects the idea that large adverse shocks are more likely to occur during recessions, as these are periods when real activity is depressed and financial conditions are tight. We include this potential non-linearity via equation (5), which allows for past developments in real activity and financial conditions to act as risk factors affecting the shape of the unemployment rate shocks. Table 3 presents the estimated posterior median of these coefficients and their 15^{th}-85^{th} credible sets. For the US we report the estimates both from the model estimated using observations up to February 2020 and the full sample in which we include time fixed effects in equation (5) to account for the Covid period.[15]

As it would be expected, the shape parameter of shocks to the unemployment rate is esti-

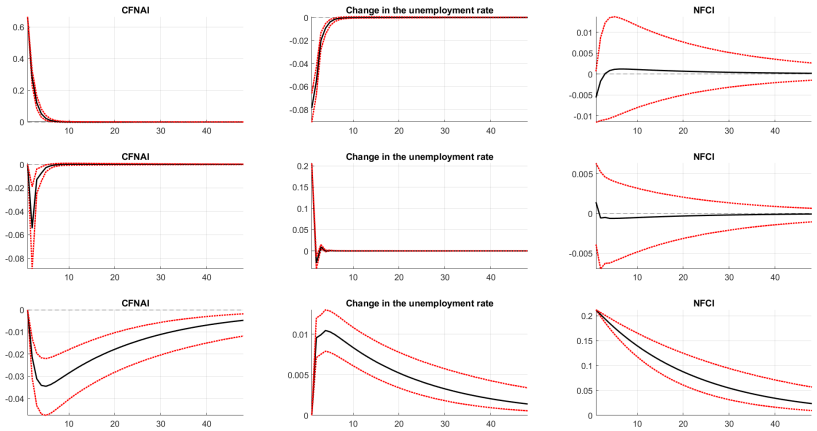
15. The time fixed effects cover the period from March 2020 up to July 2020.

Figure 7: Estimated impulse response functions from VAR model for the euro area



Notes: In the first row we report the impulse response functions (IRFs) to a one standard deviation shock to the real activity indicator (PMI Output). In the second row we report instead the IRFs to a one standard deviation shock to the monthly changes in the unemployment rate, and in the third row the IRFs to a one standard deviation shock to the financial conditions indicator (CISS).

Figure 8: Estimated impulse response functions from VAR model for the US



Notes: In the first row we report the impulse response functions (IRFs) to a one standard deviation shock to the real activity indicator (CFNAI). In the second row we report instead the IRFs to a one standard deviation shock to the monthly changes in the unemployment rate, and in the third row the IRFs to a one standard deviation shock to the financial conditions indicator (NFCI).

Table 3: Coefficients in the state equation of the shape parameter λ_{Δ_u}

Euro area	$\phi_{\lambda_{t-1}}$	$\phi_{PMI_{t-1}}$	$\phi_{CISS_{t-1}}$	c_λ
	0.7314	-0.0200	0.3120	0.3669
	[0.5845 0.8526]	[−0.0438 0.0043]	[−0.3336 1.0671]	[0.7180 0.1345]
US	$\phi_{\lambda_{t-1}}$	$\phi_{CFNAI_{t-1}}$	$\phi_{NFCI_{t-1}}$	c_λ
Full sample	0.7536	-0.0061	0.0046	0.2576
	[0.5842 0.8782]	[−0.0688 0.0613]	[−0.1824 0.2110]	[0.0635 0.5155]
Pre-Covid	0.6880	-0.0230	0.0123	0.1753
	[0.5367 0.8256]	[−0.1605 0.0973]	[−0.1824 0.2110]	[−0.0116 0.4277]

Notes: The Table reports the posterior median estimates with 85^{th} and 15^{th} credible sets in brackets of the coefficients of the risk factor in the state equation for the shape parameters of the shocks to the change in the unemployment rate. For the US we report estimates based both on full sample and pre-Covid.

mated to increase – thus leading to extreme adverse shocks becoming more likely to materialise – following weaker developments in real activity and/or tighter financial conditions, both for the euro area and for the US. This is reflected by the sign of the estimated posterior median coefficients. However, for both areas the 15^{th}-85^{th} credible sets for these coefficients are wide and include zero. Shocks to the changes in the unemployment rate exhibits both time varying volatility and time varying skewness. In particular, Figure 9 and Figure 10 shows the estimated paths for the volatility and the shape parameters in the euro area and in the United States.[16] For the euro area, the volatility of unemployment rate shocks increased gradually over time from 1999 to 2015. The same pattern occurred in the United States over the same period. The volatility of unemployment rate shocks is however more time-varying in the US over the same period than in the euro area, although this is partially due to the longer data availability. In general, the volatility of unemployment rate shocks is broadly countercyclical in both areas, usually increasing during recessions and decreasing during expansions. However, it did not move similarly across all business cycles. For example, the increase in volatility during the Global Financial crisis and Sovereign Debt crisis were quite limited in the euro area. For the United States, the volatility of the unemployment rate shocks did not increase during the recessions in the early 1980s and increased instead right after the start of the economic recovery. Both areas recorded a strong increase in the volatility of unemployment rate shocks during the COVID pandemic. This strong increase in the unemployment rate shocks volatility was considerably more pronounced for the United States, reaching historical magnitudes as the unemployment rate suddenly increased from 3.5% in February 2020 to 14.7% in April 2020. In the euro area the volatility of unemployment rate shocks also increased but the magnitudes remained moderate, as European countries benefited from the widespread use of job retention schemes, which

16. In the Appendix, Fig 27 we report the estimates based on pre-covid sample for the US.

protected employment relations between firms and employees during the pandemic.

Shocks to the unemployment rate are on average right skewed both in the euro area and the United States. In the euro area the shape parameter ruling the skewness of unemployment rate shocks moves countercyclically over time and increases further during recessions as financial conditions tighten and real activity plunges. This pattern is more pronounced in the euro area than in the US.

Figure 11 and Figure 12 shows the contemporaneous contribution of the real activity and the financial risk factors to the skewness parameter of the labour market shocks over time. At the starting of the financial crisis, first worsening financial conditions and then the plunge in economic activity led to right skewed shocks to the labour market. In the euro area, worsening financial conditions predominantly contributed to the right skewness of the labout market shocks during the period of the Sovereign Debt Crisis, while the plunge in economic activity was the main contributor for the increase in the skewness of the shocks during the Covid crisis. In the US, the financial conditions risk factor played a predominant role in explaining the positive skewness to the labour market shocks both in the 80's and during the financial crisis. Instead, the extraordinary increase in the skewness of labour market shocks experienced during the Covid-period can be attributed almost entirely to the idiosyncratic component.

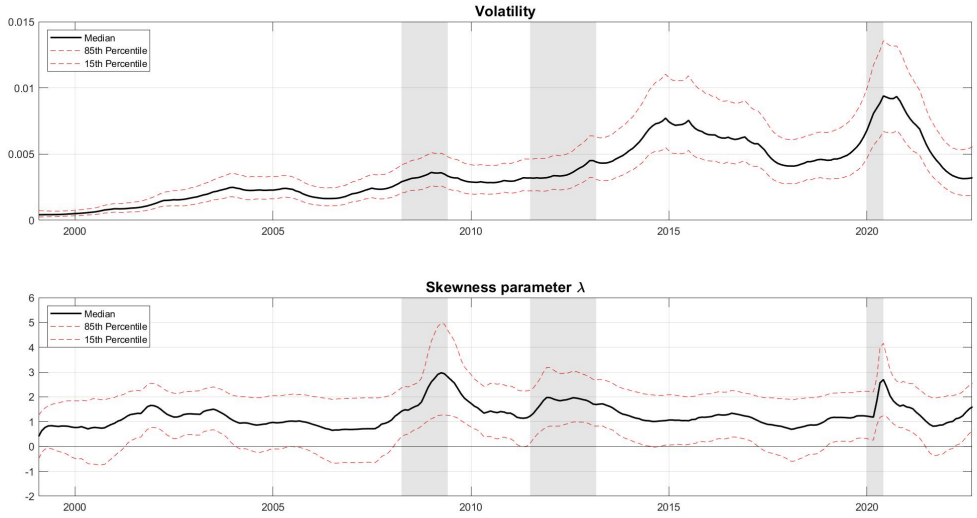
4.2 Labour at risk

The positive skewness of unemployment rate shocks in our VAR model implies that the unemployment rate is more likely to increase at a faster pace when real activity is weaker or when financial conditions are tighter. These "bad" states of the world increase the likelihood of adverse shocks to the unemployment rate to occur. For policymakers this raises two important questions – how many jobs can be at risk in case the economy is suddenly hit by a series of large adverse shocks? And how likely is this to happen over the next year?

We define "labour-at-risk" (LaR) to be the lowest predicted increase in the unemployment rate following a series of shocks, after excluding all the more favourable outcomes that could occur at a given joint probability level. We denote this probability level to be α. Moreover, we estimate our measure of labour-at-risk for a given h-periods ahead horizon. This is,

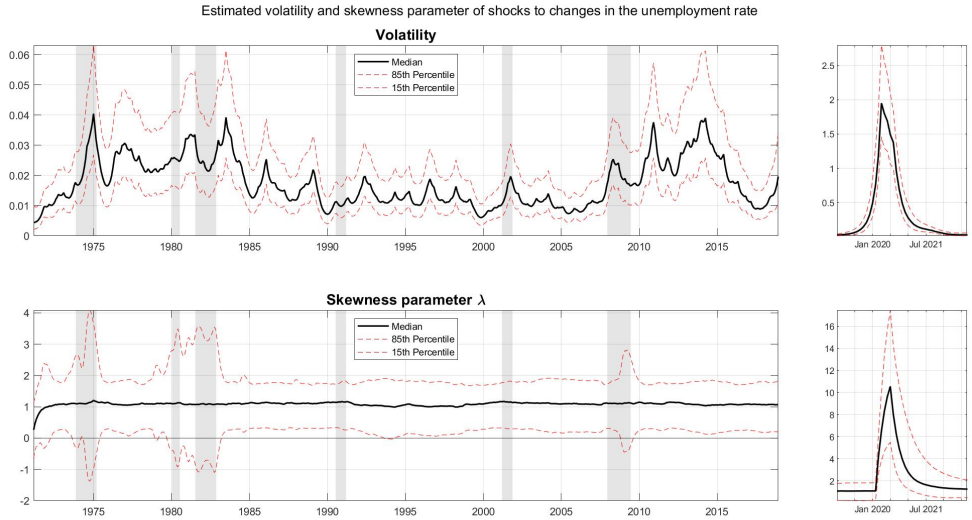
$$LaR(\alpha)_{t+h} = F^{-1}_{\Delta^h U_{t+h}}(\alpha) \qquad \alpha \in (0,1) \tag{7}$$

Figure 9: Estimated volatility and shape parameter of shocks to changes in the unemployment rate



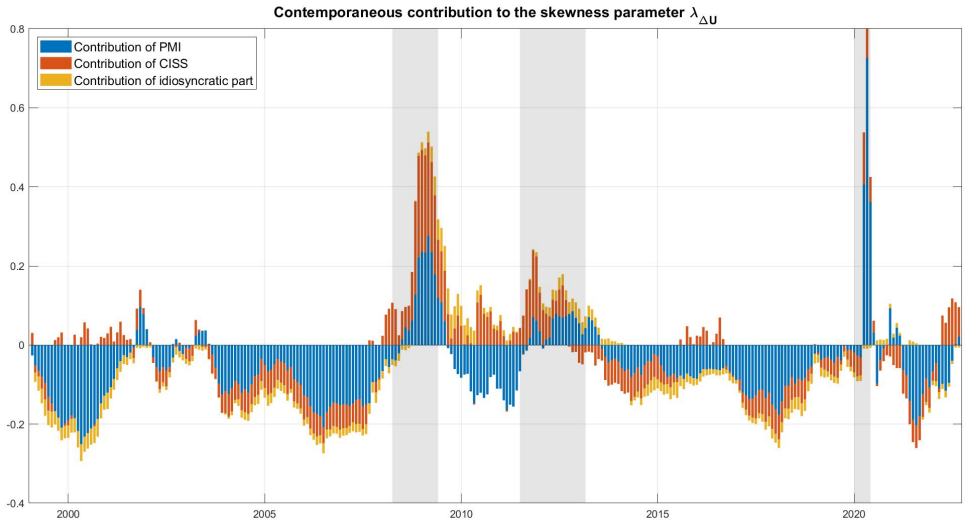
Notes: The figure shows the time series of the estimated volatilities and shape parameters of the shocks to changes in the unemployment rate in the trivariate TVSSV VAR model for the euro area. The shadow bands are for the EACN recessions periods.

Figure 10: Estimated volatility and shape parameter of shocks to changes in the unemployment rate for the US



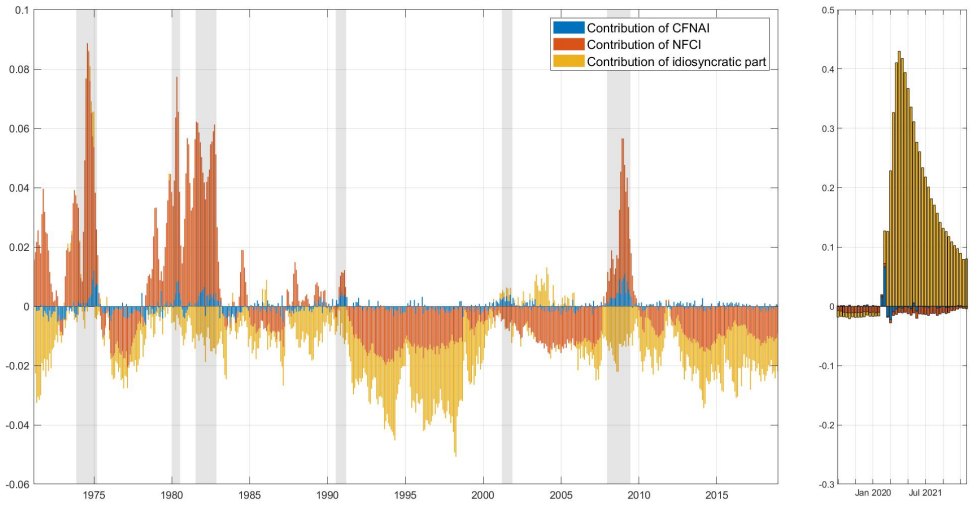
Notes: The figure shows the time series of the estimated volatilities and shape parameters of the shocks to changes in thew unemployment rate in the trivariate TVSSV VAR model for the US. The shadow bands are for the NBER recessions periods.

Figure 11: Contemporaneous contributions to the skewness parameter of the labour market shocks over time in the Euro Area



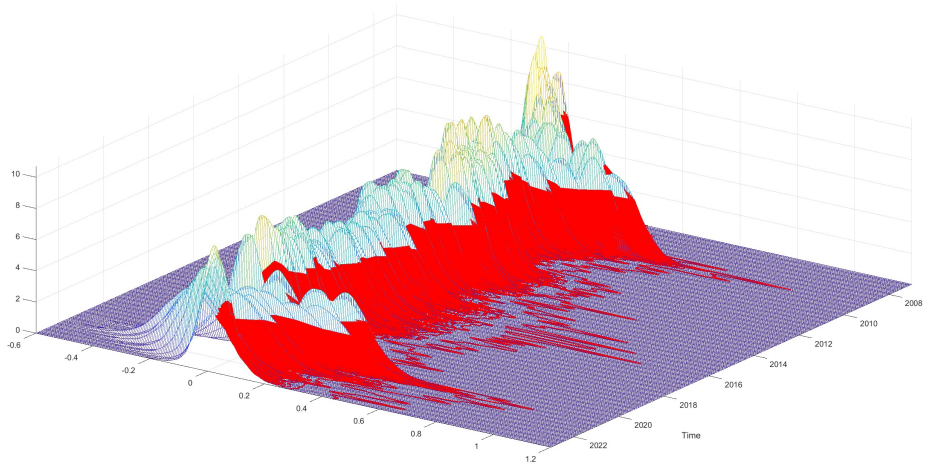
Notes: The bar chart presents the contribution to the skewness parameter of the real activity (PMI) and the financial (CISS) risk factors over time in the euro area.

Figure 12: Contemporaneous contributions to the skewness parameter of the labour market shocks over time in the US



Notes: The bar chart presents the contribution to the skewness parameter of the real activity (CFNAI) and the financial (NFCI) risk factors over time in the US.

Figure 13: One month ahead predictive distribution of changes in the unemployment rate and 80^{th} percentile labour-at-risk for the euro area

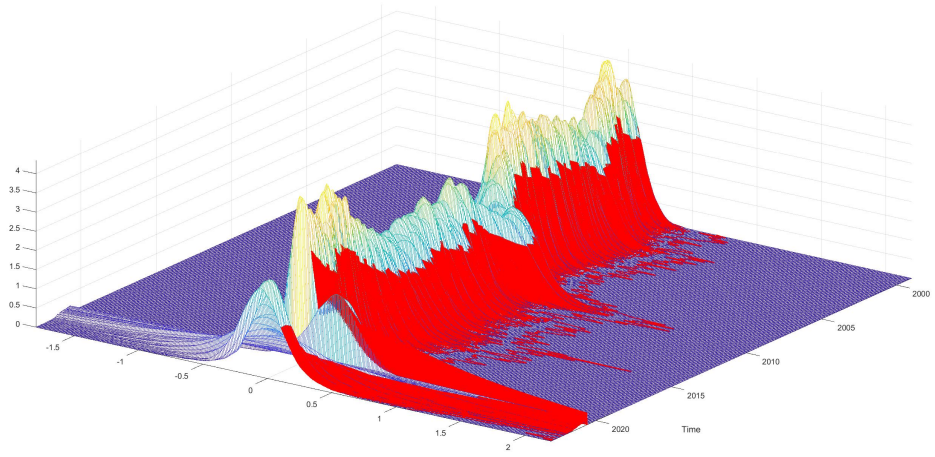


Notes: The figure shows the one month ahead predictive distribution for the month-on-month changes in the unemployment rate from January 2007 to September 2022 in the euro area. In red the part of the distribution on the right of the estimated 80^{th} percentile.

Where F^{-1} is the inverse predictive cdf of the change in the unemployment rate. In simple terms, we define labour-at-risk (α) to be the α percentile of the predictive distribution of changes in the unemployment rate h-periods ahead ($\Delta^h U_{t+h}$). We follow Kiley (2022) and focus on $\alpha = 0.8$, that is, we look to the minimum increase in the unemployment rate that would occur in case the economy was hit by shocks in the set of the 20% most adverse shocks to real activity, labour market, and financial conditions. Figure 13 for the euro area and Figure 14 for the United States show the one month ahead predictive densities for the changes in the unemployment rate in our model. For the euro area, we compute the predictive distribution of changes in the unemployment rate from January 2007 to September 2022. For the United States, we take advantage of the longer time series available and plot the predictive distribution of changes in the unemployment rate from January 1999 to September 2022. In red, we highlight the possible changes in the unemployment rate that would be equal or higher than our labour-at-risk measure at the 80^{th} percentile.

The predictive distribution of the changes in the unemployment rate changes over time reflecting any shift in its conditional mean, in its conditional variance in its conditional skewness. Hence, shifts in the conditional mean, in the conditional variance and in the conditional skewness

Figure 14: One month ahead predictive distribution of changes in the unemployment rate and 80^{th} percentile labour-at-risk for the US

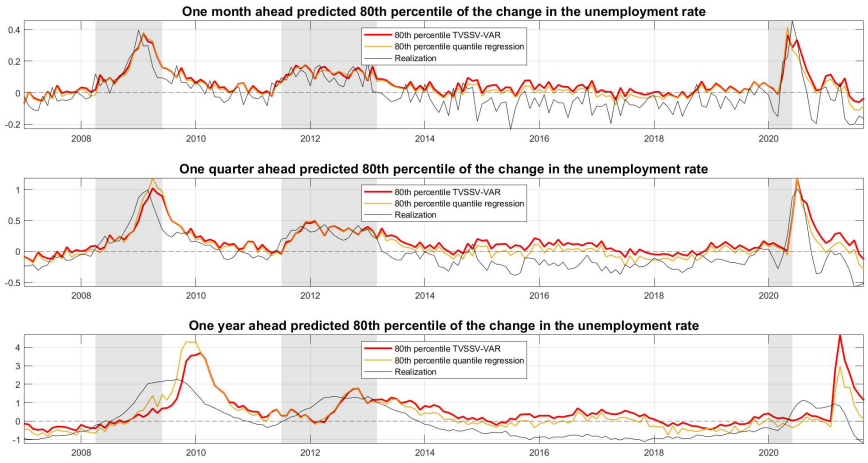


Notes: The figure shows the one month ahead predictive distribution for the month-on-month changes in the unemployment rate from January 1999 to September 2022 in the US. In red the part of the distribution on the right of the estimated 80^{th} percentile.

contribute jointly to the determine the time variation of our labour-at-risk measure. Figure 15 and Figure 16 show the estimated labour-at-risk for the euro area and the United States in a two-dimensional setup. Instead of focusing only on the predicted labour-at-risk one month ahead, we look also at different forecast horizons, and in particular we highlight the predicted labour-at-risk both one quarter ahead and one year ahead. We compare our estimates of labour at risk with those that would arise from the two-step approach from Adrian et al. (2019) that is based on quantile regressions, together with the realised value for the corresponding change in the unemployment in the same period and within the same horizon. The two econometric approaches provide observationally similar estimates for labour-at-risk.

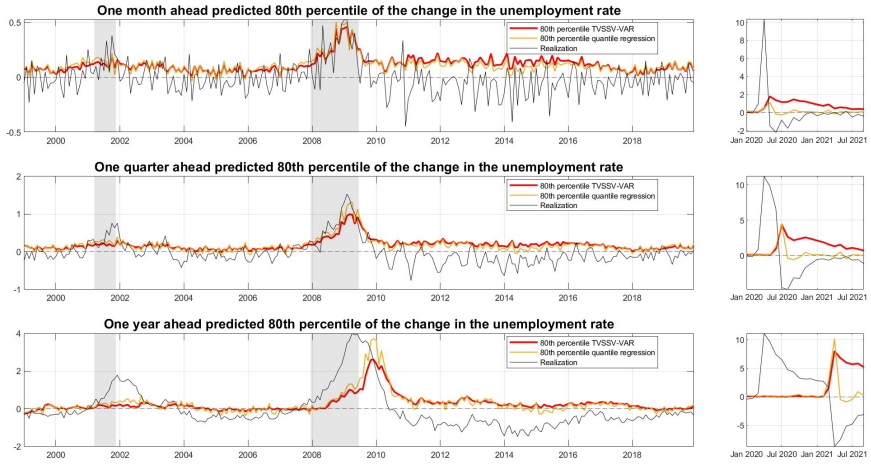
Our labour-at-risk measure targets well on average the realised changes in the unemployment rate during recessions in both the euro area and the United States. These are periods characterised by sudden increases in the unemployment rate, implying that our labour-at-risk measure provides information on the amount of jobs that are at risk in case the economy is hit by recessionary shocks. The temporary layoffs following the COVID pandemic and the associated lockdowns in the United States provided a unique set of shocks that our labour at risk measure was not able to fully cater for. Hence, the increase in the unemployment rate in the early 2020

Figure 15: Estimated labour-at-risk 80^{th} percentile for the euro area



Notes: The figure shows the estimated one month ahead (first panel), one quarter ahead (second panel), and one year ahead (third panel) estimated 80^{th} percentile of the predictive distribution of the month-on-month change in the unemployment rate (labour-at-risk) in the euro area. In red the estimates according to the TVSSV VAR model, in yellow the estimates according to the two-step quantile regression based method by Adrian et al. (2019) and in black the realization.

Figure 16: Estimated labour-at-risk 80^{th} percentile for the US



Notes: The figure shows the estimated one month ahead (first panel), one quarter ahead (second panel), and one year ahead (third panel) estimated 80^{th} percentile of the predictive distribution of the month-on-month change in the unemployment rate (labour-at-risk) in the US. In red the estimates according to the TVSSV VAR model, in yellow the estimates according to the two-step quantile regression based method by Adrian et al. (2019) and in black the realization.

was an outlier considerably stronger than predicted by our labour at risk measure. For the euro area the decrease in real activity and increase in financial tightening were good predictors of the increase in the unemployment rate over the same period as the amount of temporary layoffs was limited via the widespread use of job retention schemes, which mitigated strongly possible increases in the unemployment rate. Hence, for the euro area, the increases in the unemployment rate during the COVID pandemic were considerably closer to our labour-at-risk estimate. By contrast, and as expected, our measure of labour-at-risk is higher and far from the realized change in the unemployment rate during recoveries and expansions, as the unemployment rate usually decreases in those periods.

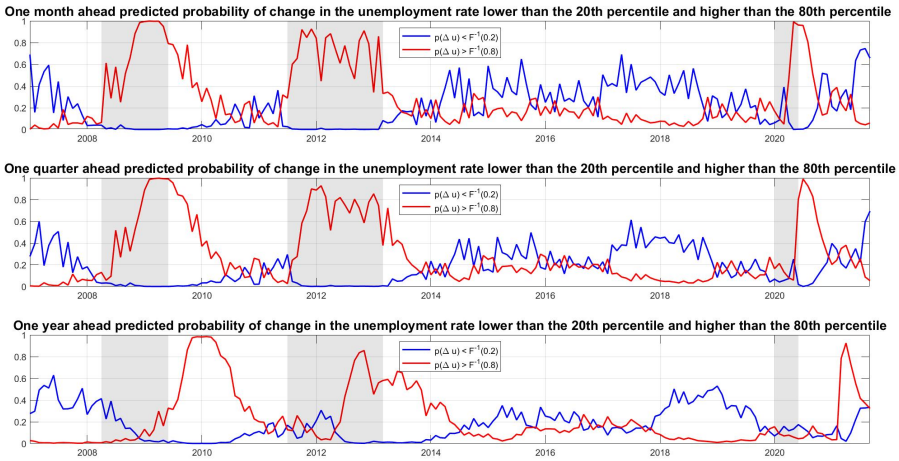
Relatedly, we use our model to estimate the conditional probability of "large" changes in the unemployment rate given the other variables in the model. At any given horizon h-months ahead we compute the probability of changes in the unemployment rate to be larger than a given threshold. This is calculated by computing the ratio between the number of simulated posterior draws in which changes in the unemployment rate exceed the threshold of interest, over the total number of simulated posterior draws. To identify the threshold across the different horizon, we look to the unconditional distribution of monthly, quarterly and yearly changes in the unemployment rate for both the euro area and the US. We denote large swings in the unemployment rate at the 20th percentile (i.e., a large downward swing) and 80th percentile (i.e., a large upward swing) in this distribution.[17]

Figure 17 show these probabilities for the euro area and Figure 18 for the United States. For the euro area, the periods in which the predicted probability of a quarterly change in the unemployment rate is higher coincide with the three recessions observed during our sample, in which real activity plunged and financial conditions tightened substantially.[18] Outside these recessionary periods, the model predicts only low probabilities of large increases in the unemployment rate for the euro area at 20% or below. More recently, the probability of a large yearly

17. For the euro area, large upward swings are identified when the increase in the unemployment rate is larger than 0.05 percentage points for $h = 1$ (one month ahead), 0.13pp for $h = 3$ (one quarter ahead), and 0.51pp for $h = 12$ (one year ahead). Conversely, large downward swings for the euro area are identified for decreases in the unemployment rate stronger than -0.08pp for $h = 1$, -0.23pp for $h = 3$, and -0.84pp for $h = 12$. For the US, large upward swings are identified for increases in the unemployment rate larger than 0.13pp for $h = 1$, 0.20pp for $h = 3$, and 0.83pp for $h = 12$, while large downward swings for decreases in the unemployment rate stronger than -0.15pp for $h = 1$, -0.28pp for $h = 3$, and -0.81pp for $h = 12$. Details on the percentiles of the unconditional distribution can be found in Table 7.

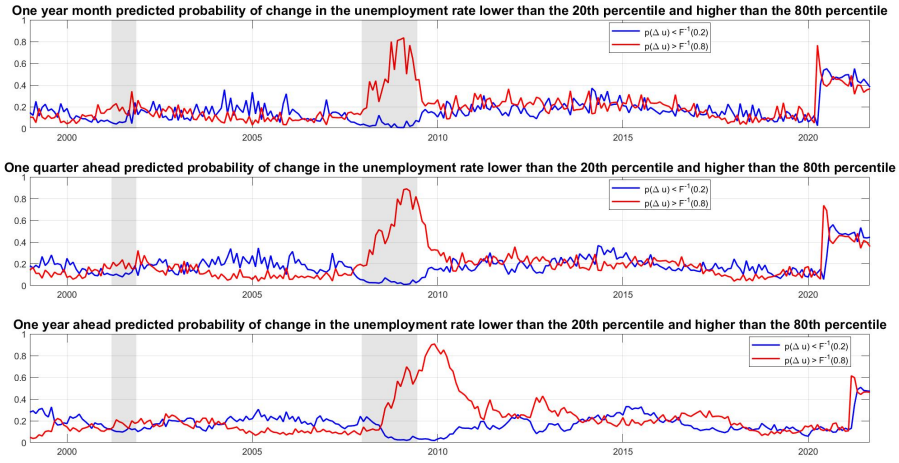
18. We focus on the quarterly probabilities for two reasons. First, the monthly thresholds are relatively low and more prone to short-term corrections. Second, the predicted yearly increases in the unemployment rate have larger uncertainty bands and are usually less timely for policymakers.

Figure 17: Predicted probability of large changes in the unemployment rate for the euro area



Notes: The figure shows the estimated one month ahead (first panel), one quarter ahead (second panel), and one year ahead (third panel) probability of changes in the month-on-month change unemployment rate larger than the unconditional 20^{th} percentile (in blue) and 80^{th} percentile (in red) in the euro area.

Figure 18: Predicted probability of large changes in the unemployment rate for the US



Notes: The figure shows the estimated one month ahead (first panel), one quarter ahead (second panel), and one year ahead (third panel) probability of changes in the month-on-month change unemployment rate larger than the unconditional 20^{th} percentile (in blue) and 80^{th} percentile (in red) in the US.

upward swing in the unemployment rate in the euro area increased during 2022, reaching almost 40% in September 2022 (the last observation in our sample). Conversely, the probability of large downward swings in the unemployment rate is higher during the expansionary periods and practically null during recessions. Similar results are obtained for the United States, although the probabilities of large swings in the unemployment rate are more symmetric during expansion periods than for the euro area. The post-pandemic period in the US was characterised by an increase in both the probabilities of large upward swings and large downward swings in the unemployment rate, as a result of the strong increase in the volatility of the changes in the unemployment rate that stemmed from the large reallocation flows that followed the temporary layoffs and Great Resignation, and the re-entry of these workers back into the US labour market.

4.3 Out-of-sample forecast accuracy

To assess the forecast accuracy of our model, we compare the forecasts from our time varying skewness VAR model with stochastic volatility (BVAR-TVSSV) to the forecasts from other competing models: (i) a Bayesian VAR model with *Independent Normal Inverse-Wishart* prior (BVAR), (ii) a Bayesian VAR model with stochastic volatility (BVAR-SV), and (iii) the quantile regression based method proposed by Adrian et al. (2019). This set of competing models allows us to assess the relative importance of accounting for modelling different features such as stochastic volatility, time varying skewness, and non-linearities among the risk factors and the target variables. On the one hand, the BVAR-SV, BVAR-TVSSV and quantile regression allow to capture time varying conditional volatility, while the simple BVAR cannot. On the other hand, only the BVAR-TVSSV and the quantile regression based method allow to account for time varying conditional skewness and the for the potential nonlinear effect of the real activity and financial risk factors on the labour market.[19]

The forecasting exercise is designed such that we compute the recursive one month, one quarter, and one year ahead forecasts on starting in January 2007 for the euro area and in January 1999 for the US. The forecast accuracy is evaluated using an expanding window over the sample between January 2007 and September 2022 for the euro area, described in Table 4, and over the sample between January 1999 and September 2022 for the United States, showcased in Table 5. We highlight the best performer according to various metrics in bold. These metrics comprise the average Root Mean Squared Error (RMSE) to evaluate point forecast accuracy,

19. Details on the competing models are presented in the Appendix A.4.

Table 4: Forecast accuracy January 2007 - September 2022 in the euro area

	RMSE	CRPS	LTw-CRPS	RTw-CRPS	QS 20th	QS 80th
a) One month ahead change						
BVAR	0.0051	0.0399	0.0117	0.0126	0.0176	0.0202
BVAR SV	0.0050	0.0396	0.0116	0.0126	**0.0175**	**0.0201**
BVAR TVSSV	**0.0049**	**0.0395**	**0.0115**	0.0126	0.0180	0.0212
Quantile regression	0.0049	0.0396	0.0120	**0.0123**	0.0194	0.0206
b) One quarter ahead change						
BVAR	0.0298	0.0964	0.0277	0.0313	0.0418	0.0521
BVAR SV	0.0285	0.0936	0.0263	0.0310	0.0384	0.0514
BVAR TVSSV	0.0276	**0.0926**	**0.0257**	0.0310	**0.0382**	0.0524
Quantile regression	**0.0270**	0.0949	0.0304	**0.0281**	0.0508	**0.0455**
c) One year ahead change						
BVAR	0.8184	0.4726	0.1300	0.1583	0.1948	0.2775
BVAR SV	0.5974	0.4265	0.1085	0.1538	0.1540	0.2735
BVAR TVSSV	0.5797	**0.4212**	**0.1055**	0.1538	**0.1469**	0.2747
Quantile regression	**0.5303**	0.4568	0.1488	**0.1387**	0.2450	**0.2238**

Notes: The Table reports the average Root Mean Squared Error (RMSE), Average Cumulative Ranked Probability Score, and Quantile Scores for the 20^{th} and 80^{th} percentiles. In bold, the best model according to each forecast metric.

the average Cumulative Ranked Probability Scores (CRPS) to evaluate overall density forecast accuracy, the average right and left tail CRPS (Gneiting et al. 2011) to evaluate density forecast accuracy on the tails of the predictive distributions, and the average quantile scores at the 20th and 80th percentiles to evaluate the accuracy for targeted percentiles.[20]

Our BVAR TVSSV model is almost always the best performer both for euro area and for the US in terms of density forecast accuracy, as measured by the CRPS. When it is not the best model, it nevertheless provides accurate and a competing forecasts to the other models. Both quantile regression based model and our BVAR TVSSV model provide the most accurate forecasts according to the RMSE for the euro area, suggesting the importance of accounting for the non-linear effects of real activity and financial conditions on the unemployment. For the US, BVAR TVSSV and BVAR SV provide most accurate point forecasts. For what concerns density forecast accuracy, for both areas we find that our BVAR TVSSV model and the BVAR-SV often provide the most accurate density forecasts out of sample, outperforming both the simple BVAR and the quantile regression based model. This result confirms the importance of modelling changes in the conditional variance in order to obtain accurate density forecasts of the unemployment rate, consistently with Carriero et al. (2020a, 2020b). Regarding the density forecast accuracy on the tails, the best performers are often the BVAR TVSSV for the left tail and the quantile regression based model for the right tail according both to the tail weighted CRPS and the quantile scores metrics.

20. The details about the forecasts metrics can be found in the Appendix A.5

Table 5: Forecast accuracy January 1999 - September 2022 in the US

	RMSE	**CRPS**	**LTw-CRPS**	**RTw-CRPS**	**QS 20th**	**QS 80th**
a) One month ahead change						
BVAR	0.7309	0.1683	0.0520	0.0551	0.0817	0.0868
BVAR SV	0.4176	0.1355	0.0337	0.0512	**0.0461**	0.0835
BVAR TVSSV	**0.4128**	**0.1337**	**0.0333**	0.0505	0.0489	0.0854
Quantile regression	0.4240	0.1432	0.0441	**0.0468**	0.0699	**0.0735**
b) One quarter ahead change						
BVAR	3.4308	0.3669	0.1136	0.1210	0.1800	0.1930
BVAR SV	**1.2565**	**0.2819**	**0.0684**	0.1094	**0.0965**	0.1798
BVAR TVSSV	1.2572	0.2823	0.0687	**0.1091**	0.1010	0.1847
Quantile regression	1.3636	0.3004	0.0935	0.0988	0.1492	**0.1559**
c) One year ahead change						
BVAR	14.8082	0.8996	0.2562	0.3111	0.3965	0.5239
BVAR SV	**3.1297**	0.7385	0.1674	0.2989	0.2268	0.5108
BVAR TVSSV	3.1475	**0.7369**	**0.1670**	0.2979	**0.2233**	0.5088
Quantile regression	3.5671	0.7723	0.2205	**0.2694**	0.3469	**0.4360**

Notes: The Table reports the average Root Mean Squared Error (RMSE), Average Cumulative Ranked Probability Score, and Quantile Scores for the 20^{th} and 80^{th} percentiles. In bold, the best model according to each forecast metric.

As a caveat, the performance of our model is sometimes not as good as that of quantile regression based methods on the right tail, while it is always outperforming quantile regression on the left tail. This comes from the fact that the median estimate of the shape parameter of the shocks to changes in the unemployment rate is estimated to be persistently positive for all the sample in analysis. That is, the model at times efficiently assigns low probability to large decreases in the unemployment rate while assigning sometimes too high probability to large increases. This comes from the flexible but parametric nature of our model which allows the shape of the distribution of shocks to vary over time according to a persistent stochastic process. This feature allows for the model structure to be easily augmented to extend the analysis towards a multivariate setting, which we will explore in the next section where we assess the risk of stagflation.

5 Stagflation risk

The recent sudden increase in inflation rates in both the euro area and the United States gave rise to a discussion on whether these economies would enter into a stagflation period. The term was initially coined by the British politician Iain Macleod in 1965, and was later used to assess the macroeconomic situation in the United States in the early 1970s. Stagflation is loosely defined as periods of low or negative output growth, an increasing or persistently high level of unemployment, and an inflation rate that is high by historical standards (Ha et al. (2022)). We

exploit the multivariate nature of our model to study stagflation risk. In particular we analyze how stagflation risk has changed over time both in the euro area and in the US, by disentangling the risk deriving from an increasing probability of large upswings in the unemployment rate and the risk deriving from an increasing probability of large inflation rates.

To do so, we extend our model to include monthly data on inflation rates. For the euro area we use the Harmonized Index of Consumer Prices (HICP) for all items and for the United States we use the Consumer Price Index (CPI) for all items.[21] We slightly adjust our BVAR TVSSV to account for the year-on-year changes in the unemployment rate and the yearly inflation rate, π_t. We denote $\boldsymbol{y}_t = [\text{PMI}_t, \Delta^{12}U_t, \pi_t, \text{CISS}_t]$ for the euro area and $\boldsymbol{y}_t = [\text{CFNAI}_t, \Delta^{12}U_t, \pi_t, \text{NFCI}_t]$ for the United States, where $\Delta^h U_t = U_t - U_{t-h}$. The risk factors are left unaltered compared to the model in Section 3, but are now allowed to affect the shape of the shocks both to changes in the unemployment rate and to inflation,

$$\begin{aligned} \lambda_{\Delta^{12}U,t} &= \phi_1 \lambda_{\Delta^{12}U,t-1} + \boldsymbol{\phi_2 x_{t-1}} + \xi_{\Delta^{12}U,t} & \xi_{\Delta^{12}U,t} &\sim \mathcal{N}(0, \sigma^2_{\xi,\Delta^{12}U}) \\ \lambda_{\pi,t} &= \rho_1 \lambda_{\pi,t-1} + \boldsymbol{\rho_2 x_{t-1}} + \xi_{\pi,t} & \xi_{\pi,t} &\sim \mathcal{N}(0, \sigma^2_{\xi,\pi}) \end{aligned}$$

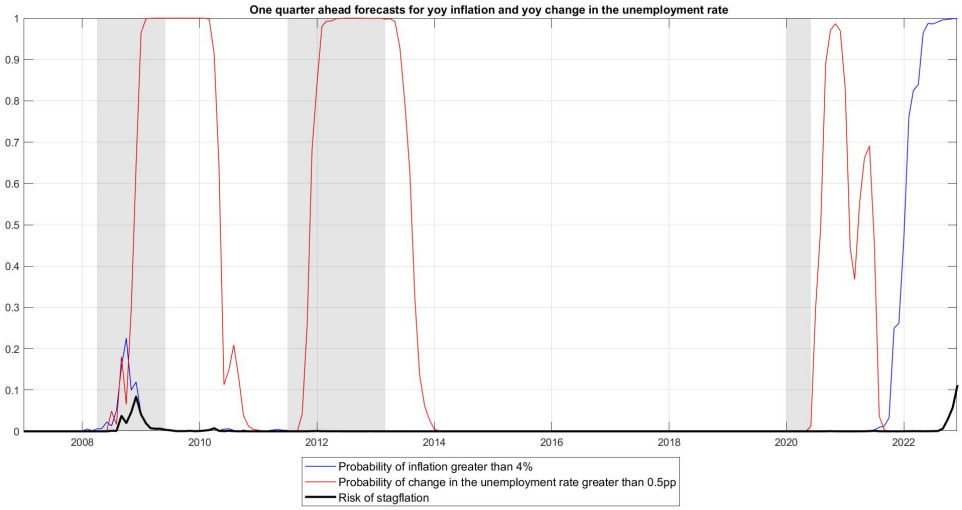
where $\boldsymbol{x_{t-1}}$ is the vector of lagged real activity and financial risk factors. This specification is in line with López-Salido et al. (2020), who document a nonlinear relation between financial conditions and inflation using quantile regression and a Markov switching model.

We define stagflation risk as the joint probability that the yearly changes in the unemployment rate and inflation rates are above their given thresholds at any given point in time. In this way, we consider as stagflation periods those with a large increase in the unemployment rate over a year and with high inflation levels. These thresholds are identified by making use of the information in the unconditional distribution of yearly changes in the unemployment rate and inflation rates for both the euro area and the United States over time. We use the wider sample for the United States to calibrate our threshold for inflation in the two areas with data from 1984 onwards. [22] The 80th percentile of the distribution of US inflation rates stands at 4%. For the yearly changes in the unemployment rate, we set the threshold at 0.5 percentage points.

21. The series for the HICP can be obtained from the ECB Statistical Data Warehouse and the series for the CPI is obtained from the FRED-MD database with the code CPIAUSL.

22. The correlation between the inflation rates in the euro area and in the US stands at above 85% between January 1999 and September 2022, and inflation rates moving in a broadly synchronised way for both areas. Hence, we use the historical distribution of inflation rates in the US as an approximation for the historical distribution of inflation rates in the euro area.

Figure 19: One quarter ahead probability of stagflation in the euro area



Notes: The figure shows the estimated one quarter ahead probability of the change in the year-on-year unemployment rate being greater than 0.5pp (in red), the year-on-year inflation rate being greater than 4% (in blue) and the probability of both events occurring (in black) in the euro area. The shadow bands are for the EACN recessions periods.

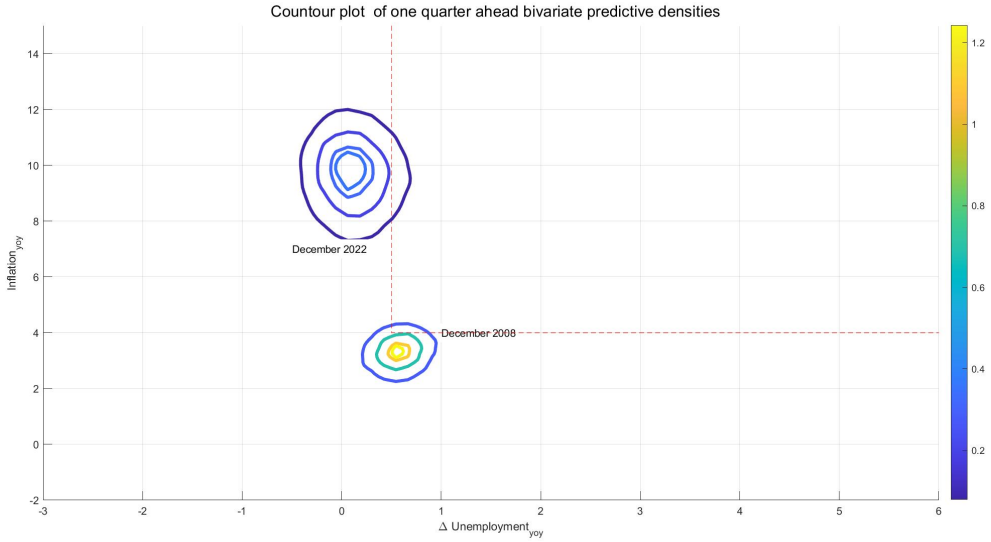
This corresponds to the 80th percentile for the euro area since 1999 and to the 83rd percentile for the US since 1984.

We assess the risks of stagflation in Figure 19 for the euro area. In particular, we show the estimated one quarter ahead joint probabilities on yearly changes in the unemployment rate exceeding the 0.5 percentage points and inflation surpassing 4%. We decompose the stagflation risk by displaying separately the labour at risk channel (in red) and the inflation risk channel (in blue).[23] The models are estimated with data available up to September 2022, implying that we assess the risk of stagflation up to December 2022.

There was only a limited risk of stagflation in the euro area since 2007. The probability of stagflation reached around 10% in December 2008 during the Global Financial crisis, first with an increase in inflation risk and later with a strong increase in the amount of labour at risk. During the Sovereign Debt crisis there was a high degree of labour at risk but no inflation risk. The risk of stagflation decreased and remained virtually null until the second half of 2022, when it started increasing driven by the persistently high levels of inflation and a gradually increasing

23. We compute the joint probability of two variables exceeding a given threshold at a given horizon h by computing the ratio between the number of draws in which the two variables exceed the threshold over the total number of draws.

Figure 20: Contour plot for one quarter ahead joint predictive densities for π_{yoy} and ΔU_{yoy} in the euro area: December 2008 vs December 2022

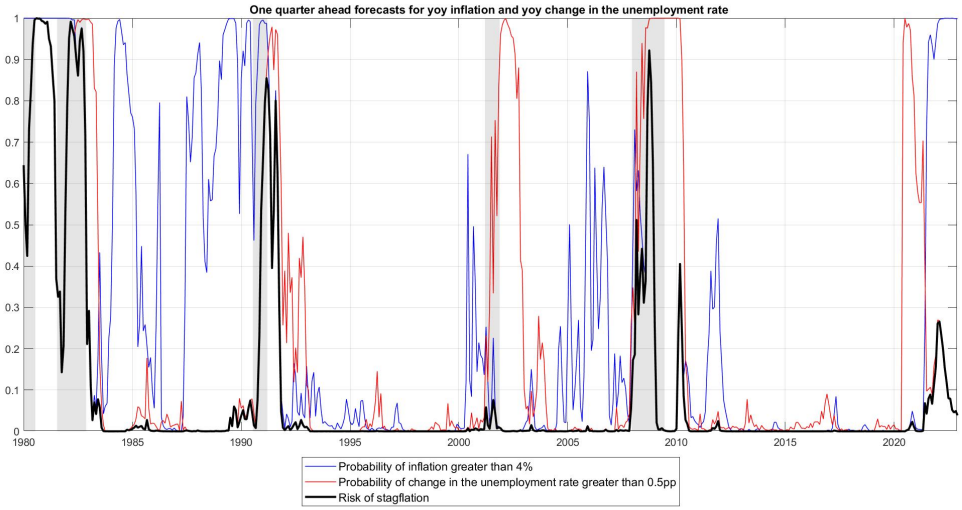


Notes: The figure shows the contours from one quarter ahead bivariate predictive density for the year-on-year change in the unemployment rate and year-on-year inflation rate for December 2008 and September 2023 in the euro area. The contours identify 20% 50%, 80% and 90% of the bivariate predictive density. The area inside the red-dotted rectangle signals year-on-year inflation greater than 4% and changes in the year-on-year unemployment rate greater than 0.5pp.

degree of labour at risk in the economy. Figure 20 provides further information on the risk of stagflation by contrasting the contour plots for the 20th, 50th, 80th and 90th percentiles of the bivariate predictive densities for inflation and yearly changes in the unemployment rate in the euro area between December 2008, using data up to September 2008, and December 2022 using data up to September 2022. While the stagflation risk reached similar magnitudes in both periods, it was closer to both borders of the stagflation area in 2008, with a more limited inflation risk and a slightly higher labour at risk. By contrast, there is a higher uncertainty in the estimate of stagflation risk in 2022 as the volatility of the bivariate predictive density is wider.

In Figure 21 we show the estimated one quarter ahead joint probabilities of yearly changes in the unemployment rate exceeding the 0.5 percentage points and inflation surpassing 4% for the US. The period considered runs from from January 1980 up to December 2022. In the early 80s, during the recession following the Oil Crisis both high predicted probabilities of large increases in the unemployment rate and in the inflation rate contributed to high stagflation risk. Instead, after July 1984 until 1990, stagflation risk has been muted in the US. As a matter of fact, while

Figure 21: One quarter ahead probability of stagflation in the US

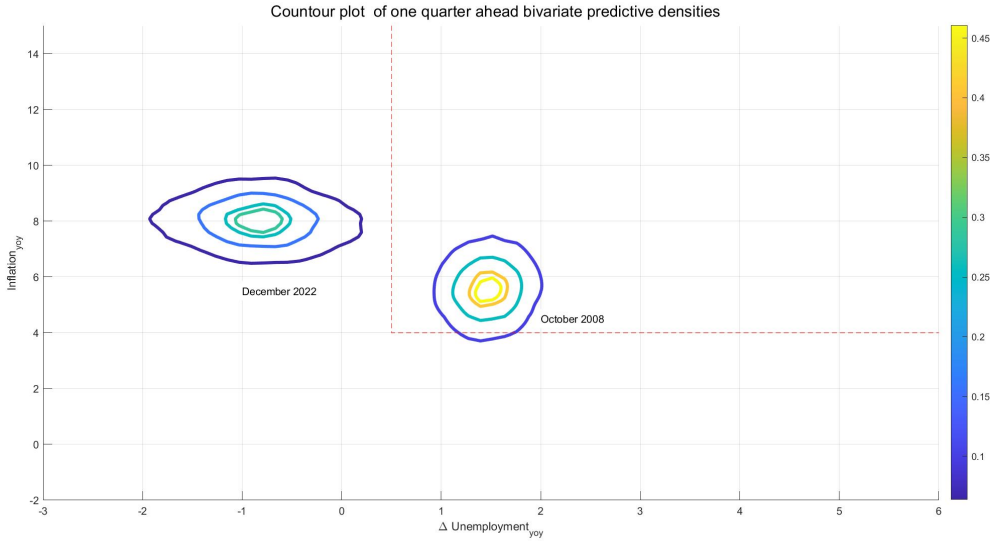


Notes: The figure shows the estimated one quarter ahead probability of the change in the year-on-year unemployment rate being greater than 0.5pp (in red), the year-on-year inflation rate being greater than 4% (in blue) and the probability of both events occurring (in black) in the US. The shadow bands are for the NBER recessions periods.

inflation risk remained relatively high until January 1985 increasing a second time in the late 80s, the estimated probability of large increases in the unemployment rate remained very low in the entire period 1984-1990. As the recession started in July 1990, stagflation probabilities increased and reached almost 85 % in March 1990. At the end of the recession, a decrease in the inflation risk that preceded a decrease in the estimated probability of large increases in the unemployment rate induced a decrease in the stagflation risk. Afterwards, stagflation risk remained muted for almost 19 years, with large predicted increases in the unemployment rate during the dot-com bubble not followed by expected large inflation rates. Stagflation risk reemerged at the burst of the financial crisis. As a matter of fact, our estimated measure for stagflation risk increased above 90% in the second half of 2008 and around 40% in 2010. In 2021, with the increase in energy prices, the probability of stagflation risk increased in the US, but remained contained and decreased in 2022 due to the strong performance of the US labour market, which points in our model to a lower risk of unemployment with the inflation risk remaining instead elevated.

A comparison between the estimated stagflation risk for October 2008, using data up to July 2008, and December 2022 using data up to September 2022 is showcased in Figure 22. It unveils that the current economic juncture is considerably more uncertain than that in 2008, similarly

Figure 22: Contour plot for one quarter ahead joint predictive densities for π_{yoy} and ΔU_{yoy} in the United States: October 2008 vs December 2022



Notes: The figure shows the contours from one quarter ahead bivariate predictive density for the year-on-year change in the unemployment rate and year-on-year inflation rate for July 2009 and September 2023 in the US. The contours identify 20% 50%, 80% and 90% of the bivariate predictive density. The area inside the red-dotted rectangle signals year-on-year inflation greater than 4% and changes in the year-on-year unemployment rate greater than 0.5pp.

to what happened for the euro area. However, and in contrast to what was observed for the euro area, this uncertainty seems to be more prevalent in the predicted degree of labour at risk.

6 Conclusion

We develop a BVAR model with time varying skewness and stochastic volatility that caters for the fact that the unemployment rate changes are asymmetric over the business cycles, declining slowly and on average during economic expansions and rising suddenly and violently during downturns. The model is applied to both the euro area and the United States to capture and quantify the degree of labour-at-risk in the economy, providing policymakers with timely information about possible risks affecting the labour market and showing how much can the unemployment rate increase at any given moment in time if the economy is hit by a persistent series of negative shocks. Movements in the average response of the unemployment rate and in the asymmetry of labour market shocks depend both on the developments in real activity or on the tightening of financial conditions. Further, we use our BVAR to track stagflation risk in the economy, defined as the joint event of both a high degree of labour at risk and a

high inflation risk. The analysis of joint risks could prove important for the assessment of the unemployment-inflation trade-off and of the scope of monetary policy. Our work provides also the foundation for embedding asymmetric shocks as part of the toolkit used for the estimation of medium-scale DSGE models.

References

Abbritti, Mirko, and Stephan Fahr. 2013. "Downward Wage Rigidity and Business Cycle Asymmetries." *Journal of Monetary Economics* 60 (7): 871–886.

Adrian, Tobias, Nina Boyarchenko, and Domenico Giannone. 2019. "Vulnerable Growth." *American Economic Review* 109, no. 4 (April): 1263–89.

Andolfatto, David. 1997. "Evidence and Theory on the Cyclical Asymmetry in Unemployment Rate Fluctuations." *Canadian Journal of Economics* 30 (3): 709–721.

Arias, Jonas E., Juan F. Rubio-Ramirez, and Minchul Shin. 2021. *Macroeconomic Forecasting and Variable Ordering in Multivariate Stochastic Volatility Models.* Working Papers 21-21. Federal Reserve Bank of Philadelphia, June.

Azzalini, Adelchi. 1986. "Further results on a class of distributions which includes the normal ones." *Statistica* 46 (2): 199–208.

Bai, Jushan, and Serena Ng. 2005. "Tests for Skewness, Kurtosis, and Normality for Time Series Data." *Journal of Business Economic Statistics* 23:49–60.

Brownlees, Christian, and André B.M. Souza. 2021. "Backtesting global Growth-at-Risk." *Journal of Monetary Economics* 118:312–330.

Carriero, Andrea, Todd E Clark, and Massimiliano Marcellino. 2019. "Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors." *Journal of Econometrics* 212 (1): 137–154.

———. 2020a. *Capturing Macroeconomic Tail Risks with Bayesian Vector Autoregressions.* Working Papers 20-02R. Federal Reserve Bank of Cleveland, January.

Carriero, Andrea, Todd E. Clark, and Massimiliano Marcellino. 2020b. *Nowcasting Tail Risks to Economic Activity with Many Indicators.* Working Papers 20-13R2. Federal Reserve Bank of Cleveland, May.

Chavleishvili, Sulkhan, and Simone Manganelli. 2019. *Forecasting and stress testing with quantile vector autoregression.* Working Paper Series 2330. European Central Bank.

Cogley, Timothy, and Thomas J. Sargent. 2005. "Drifts and volatilities: monetary policies and outcomes in the post WWII US." Monetary Policy and Learning, *Review of Economic Dynamics* 8 (2): 262–302.

Delle Monache, Davide, Andrea De Polis, and Ivan Petrella. 2021. *Modeling and forecasting macroeconomic downside risk.* Temi di discussione (Economic working papers) 1324. Bank of Italy, Economic Research and International Relations Area, March.

DeLong, Brad, and Larry Summers. 1986. "Are Business Cycles Symmetrical?" In *American Business Cycle: Continuity and Change,* edited by Robert Gordon, 166–79. Chicago: University of Chigago Press.

Diamond, Peter A. 1982. "Aggregate Demand Management in Search Equilibrium." *Journal of Political Economy* 90 (5): 881–894.

Falk, B. 1986. "Further Evidence on the Asymmetric Behaviour of Economic Time Series over the Business Cycle." *Journal of Political Economy* 94:1096–1109.

Ferraro, Domenico. 2018. "The asymmetric cyclical behavior of the U.S. labor market." *Review of Economic Dynamics* 30:145–162.

Ferraro, Domenico, and Giuseppe Fiori. 2022. "Search Frictions, Labor Supply, and the Asymmetric Business Cycle." *Board of Governors of the Federal Reserve System, International Finance Discussion Papers no. 1355.*

Figueres, Juan Manuel, and Marek Jarociński. 2020. "Vulnerable growth in the euro area: Measuring the financial conditions." *Economics Letters* 191 (C).

Giglio, Stefano, Bryan Kelly, and Seth Pruitt. 2016. "Systemic risk and the macroeconomy: An empirical evaluation." *Journal of Financial Economics* 119 (3): 457–471.

Gneiting, Tilmann, and Roopesh Ranjan. 2011. "Comparing Density Forecasts Using Threshold- and Quantile-Weighted Scoring Rules." *Journal of Business Economic Statistics* 29 (3): 411–422.

Ha, Jongrim, M. Ayhan Kose, and Franziska Ohnsorge. 2022. "Global Stagflation." *CEPR Discussion Paper No. DP17381,* 1–48.

Ilut, Cosmin, Matthias Kehrig, and Martin Schneider. 2018. "Slow to Hire, Quick to Fire: Employment Dynamics with Asymmetric Responses to News." *Journal of Political Economy* 126 (5): 1785–2178.

Iseringhausen, Martin. 2023. "A time-varying skewness model for Growth-at-Risk." *International Journal of Forecasting.*

Keynes, John Maynard. 1936. *The General Theory of Employment, Interest, and Money.* Macmillan.

Kiley, Michael T. 2022. "Unemployment Risk." *Journal of Money, Credit and Banking* 54 (5): 1407–1424.

Kremer, Manfred, Marco Lo Duca, and Dániel Holló. 2012. *CISS - a composite indicator of systemic stress in the financial system.* Working Paper Series 1426. European Central Bank, March.

Lindsten, Fredrik, Michael I Jordan, and Thomas B Schon. 2014. "Particle Gibbs with ancestor sampling." *Journal of Machine Learning Research* 15:2145–2184.

Litterman, Robert. 1986. "Forecasting with Bayesian Vector Autoregressions-Five Years of Experience." *Journal of Business Economic Statistics* 4 (1): 25–38.

López-Salido, J. David, and Francesca Loria. 2020. *Inflation at Risk.* Finance and Economics Discussion Series 2020-013. Board of Governors of the Federal Reserve System (U.S.), February.

McKay, Alisdair, and Ricardo Reis. 2008. "The brevity and violence of contractions and expansions." *Journal of Monetary Economics* 55:738–751.

Mitchell, Wesley Clair. 1927. *Business Cycles: The Problem and Its Setting.* National Bureau of Economic Research.

Montes-Galdón, Carlos, and Eva Ortega. 2022. *Skewed SVARs: tracking the structural sources of macroeconomic tail risks.* Working Papers 2208. Banco de España, March.

Mortensen, Dale T. 1982. "Property Rights and Efficiency in Mating, Racing, and Related Games." *American Economic Review* 72 (5): 968–979.

Neftçi, Salih N. 1984. "Are Economic Time Series Asymmetric over the Business Cycle?" *Journal of Political Economy* 92:307–328.

Newey, Whitney K., and Kenneth D. West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55 (3): 703–708.

Ordoñez, Guillermo. 2013. "The Asymmetric Effects of Financial Frictions." *Journal of Political Economy* 121 (5): 844–895.

Pissarides, Christopher A. 1985. "Short-Run Equilibrium Dynamics of Unemployment, Vacancies, and Real Wages." *American Economic Review* 75 (4): 676–690.

Plagborg-Møller, Mikkel, Lucrezia Reichlin, Giovanni Ricco, and Thomas Hasenzagl. 2020. "When is growth at risk?" *Brookings Papers on Economic Activity* 2020 (1): 167–229.

Primiceri, Giorgio E. 2005. "Time varying structural vector autoregressions and monetary policy." *The Review of Economic Studies* 72 (3): 821–852.

Renzetti, Andrea. 2023. *Modelling and Forecasting Macroeconomic Risk with Time Varying Skewness Stochastic Volatility Models.*

Wolf, Elias. 2021. "Estimating growth at risk with skewed stochastic volatility models." *Available at SSRN 4030094.*

A Appendix

A.1 Bai et al. (2005) test for skewness

Under the null hypothesis of no asymmetry Bai et al. (2005) test statistic is:

$$\hat{\pi}_3 = \frac{\sqrt{T}\hat{\mu}_3}{s(\hat{\mu}_3)} \xrightarrow{d} \mathcal{N}(0,1) \tag{8}$$

where $\hat{\mu}_3$ is the sample estimate of the third central moment of the distribution and $s(\hat{\mu}_3) = (\hat{\boldsymbol{\alpha}}_{\mathbf{2}}\hat{\boldsymbol{\Gamma}}_{\mathbf{22}}\hat{\boldsymbol{\alpha}}_{\mathbf{2}}')^{\frac{1}{2}}$ $\hat{\boldsymbol{\alpha}}_{\mathbf{2}} = [1, -3\hat{\sigma}^2]$. $\hat{\sigma}^2$ is a consistent estimate of the variance σ^2 and $\hat{\boldsymbol{\Gamma}}_{\mathbf{22}}$ is a consistent estimate of the 2×2 sub-matrix of $\boldsymbol{\Gamma} = \lim_{T\to\infty} T\,\mathbb{E}[\bar{Z}\bar{Z}]$ where $\bar{Z}$ is the sample mean of $\boldsymbol{Z}_t$, defined as the deviation of the empirical centered first three moments from the Gaussian's one, namely:

$$\boldsymbol{Z_t} = \begin{bmatrix} (X_t - \mu)^3 - \mu_3 \\ (X_t - \mu) \\ (X_t - \mu)^2 - \sigma^2 \end{bmatrix}$$

The long run variance is estimated following Newey et al. (1987).

A.2 TVSSV-VAR with Skew normal shocks

The TVSSV VAR(p) model with *Skew Normal* shocks is given by:

$$\boldsymbol{y_t} = \boldsymbol{\Pi_0} + \boldsymbol{\Pi_1 y_{t-1}} + \ldots + \boldsymbol{\Pi_p y_{t-p}} + \boldsymbol{A}^{-1}\boldsymbol{H}_t^{0.5}\boldsymbol{\varepsilon_t} \tag{9}$$

$$\varepsilon_{it} \sim \textit{Skew normal}(\zeta_{it}, \omega_{it}, \lambda_{it})$$

$$log(h_{i,t}) = log(h_{i,t-1}) + \eta_{i,t} \qquad \eta_{i,t} \sim \mathcal{N}(0, \sigma^2_{i,\eta}) \tag{10}$$

with $i = \{PMI, \Delta U, CISS\}$ for the euro area and $i = \{CFNAI, \Delta U, NFCI\}$ for the US.

$$\lambda_{\Delta U,t} = \phi_1 \lambda_{\Delta U,t-1} + \boldsymbol{\phi_2 x_{t-1}} + \xi_{\Delta U,t} \qquad \xi_{\Delta U,t} \sim \mathcal{N}(0, \sigma^2_{\xi,\Delta U}) \tag{11}$$

$$\lambda_{i,t} = \phi_1 \lambda_{i,t-1} + \xi_{i,t} \qquad \xi_{it} \sim \mathcal{N}(0, \sigma^2_{\xi,i}) \tag{12}$$

with $i = \{PMI, CISS\}$ for the euro area and $i = \{CFNAI, NFCI\}$ for the US. The *Skew normal* (Azzalini 1986) distribution is:

$$p(\varepsilon_t|\zeta,\omega^2,\lambda) = \frac{2}{\omega}\phi\left(\frac{\varepsilon_t-\zeta}{\omega}\right)\Phi\left(\lambda\left(\frac{\varepsilon_t-\zeta}{\omega}\right)\right)$$

where $\phi(.)$ and $\Phi(.)$ are respectively the probability density function and the cumulative density function of the Standard Normal. The mean and the variance of ε_t are given by $\mathbb{E}[\varepsilon_t] = \zeta+\omega\delta\sqrt{\frac{2}{\pi}}$ and $var(\varepsilon_t) = \omega^2\left(1-\frac{2\delta^2}{\pi}\right)$. Assuming $\mathbb{E}[\varepsilon_t] = 0$ and $var(\varepsilon_t) = 1$ implies the following constraints on the location and scale parameters $\zeta = -\omega\delta\sqrt{\frac{2}{\pi}}$ and $\omega^2 = \left(1-\frac{2\delta^2}{\pi}\right)^{-1}$.

To estimate the TVSSV model we exploit the fact that $\varepsilon_t \sim Skew-Normal(\zeta,\omega^2,\lambda)$ has the following stochastic representation:

$$\varepsilon_t = \zeta + \delta\omega v_t + \sqrt{(1-\delta^2)}\omega z_t \tag{13}$$

where $v_t \overset{i.i.d}{\sim}$ Truncated normal$_{[0,\infty)}(0,1)$ $z_t \overset{i.i.d}{\sim} \mathcal{N}(0,1)$ and $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$, with $-1<\delta<1$.

Using the stochastic representation in equation (13) for the shocks, we can write the VAR system as:

$$\boldsymbol{y_t} = \boldsymbol{\Pi_0} + \boldsymbol{\Pi_1 y_{t-1}} + \ldots + \boldsymbol{\Pi_p y_{t-p}} + \boldsymbol{A}^{-1}\boldsymbol{H_t}^{0.5}(\boldsymbol{\zeta_t} + \boldsymbol{\Omega_t\Delta_t v_t} + \boldsymbol{\Omega_t}(\boldsymbol{I}_n - \boldsymbol{\Delta_t^2})^{0.5}\boldsymbol{z_t}) \tag{14}$$

where:

$\boldsymbol{\zeta_t} = [\zeta_{1,t},\ldots,\zeta_{N,t}]'$

$\boldsymbol{\Omega_t} = diag(\omega_{1t}\ldots\omega_{Nt})$

$\boldsymbol{\Delta_t} = diag(\delta_{1t}\ldots\delta_{Nt})$

$\boldsymbol{v_t} = [v_{1,t},\ldots,v_{N,t}]'$ $\qquad v_{i,t} \sim TruncatedNormal_{(0,\infty)}(0,1)$

$\boldsymbol{z_t} = [z_{1,t},\ldots,z_{N,t}]'$ $\qquad z_{it} \sim N(0,1).$

This representation implies that conditionally on the mixing variables in $\boldsymbol{v}_t$, on the volatilities stored in the matrix $\boldsymbol{H_t}$ and on the elements in the diagonal matrix $\boldsymbol{\Delta_t}$, which are one to one map of the shape parameters (since they are defined as $\delta_{it} = \frac{\lambda_{it}}{\sqrt{1+\lambda_{it}^2}}$), the likelihood is Gaussian. This allows to resuscitate and adapt many of the closed form formulas for the full conditional posterior distributions for the parameters of the model from the standard Gaussian stochastic volatility VAR model (Carriero et al. 2019). As a matter of fact, the diagonal elements in the vector $\boldsymbol{\zeta_t}$ and in the diagonal matrix $\boldsymbol{\Omega_t}$ are neither parameters nor latent states to be estimated. They satisfy the constraints (2) and (3) so that the parameterization of the shocks is

correct. The diagonal elements in $\boldsymbol{\Delta_t}$ satisfy $\delta_{it} = \frac{\lambda_{it}}{\sqrt{1+\lambda_{it}^2}}$. The elements in the diagonal matrix $\boldsymbol{H}_t = diag(h_{1t}, \ldots, h_{1N})$ and the shape parameters λ_{it} are instead latent states satisfying the transition equations (10) (11) and (12). For further details on the estimation of the model see Renzetti (2023).

A.3 Priors of the TVSSV-VAR

Table 6 presents the details on the priors for the parameters of the time varying skewness stochastic volatility model:

Table 6: Priors for the parameters of the TVSSV-VAR model

Parameter	Prior
$vec(\boldsymbol{\Pi})$	$\mathcal{N}(vec(\underline{\boldsymbol{\mu}_{\boldsymbol{\Pi}}}), \underline{\boldsymbol{\omega}_{\boldsymbol{\Pi}}})$
a_{ij}	$\mathcal{N}(0, 100)$
$\phi_{i,1}$	$\mathcal{N}(1, \theta_1)$
$\boldsymbol{\phi_2}$	$\mathcal{N}(\underline{\mu_{\boldsymbol{\phi_2}}}, \underline{\Sigma_{\boldsymbol{\phi_2}}})$
$log(h_{i0})$	$\mathcal{N}(h_{i0}, \overline{100})$
λ_{i0}	$\mathcal{N}(0, 10)$
$\sigma^2_{i,\xi}$	$InverseGamma(5, 0.16)$
$\sigma^2_{i,\eta}$	$InverseGamma(5, 0.16)$

Notes: The table presents the prior distribution of the parameters of the TVSSV-VAR model.

where $\underline{\boldsymbol{\omega}_{\boldsymbol{\Pi}}}$ has the Minnesota type (Litterman 1986) prior:

$$v_{ij,l} = \begin{cases} \frac{\theta_1}{l^{\theta_4}} & \text{if} \quad i = j \\ \frac{\sigma_i^2 \theta_1 \theta_2}{\sigma_j^2 l^{\theta_4}} & \text{if} \quad i \neq j \end{cases} \tag{15}$$

The elements of $vec(\underline{\boldsymbol{\mu}_{\boldsymbol{\Pi}}})$ are equal zero for the coefficients on the cross-equation lags and equal to one for the coefficients of the own lags. As for the hyper-parameters, we set $\theta_1 = 0.04 \quad \theta_2 = 0.25 \quad \theta_3 = 100 \quad \theta_4 = 2$. We estimate σ_i^2 from univariate AR(12) regressions. For the initial state of the volatility, the prior mean h_{i0} where $h_{i,0}$ is the estimated variance from an AR(4) model to each series using as sample the first 40 observations. For the Normal prior for the coefficients on the risk factors in the state equation for the shape parameters of the shocks to changes in the unemployment rate $\boldsymbol{\phi_2}$ we assume a mean $\underline{\mu_{\boldsymbol{\phi_2}}} = [0, \ldots, 0]'$ and variance covariance matrix $\underline{\Sigma_{\boldsymbol{\phi_2}}} = diag\left(\frac{\theta_1 \theta_5}{\sigma_1^2 l^{\theta_4}}, \ldots \frac{\theta_1 \theta_5}{\sigma_J^2 l^{\theta_4}}\right)$ with $\theta_5 = 0.1$.

A.4 Competing models in the forecasting exercise:

In Section 4.2 we compare the forecasts from the time varying skewness stochastic volatility VAR model to the forecasts from:

- A Bayesian VAR with *Independent Normal Inverse-Wishart* prior.
- A Bayesian VAR with stochastic volatility.
- The two step approach based on quantile regression and *Skew-t* interpolation by (Adrian et al. 2019).

A.4.1 Bayesian VAR with Independent Normal Inverse-Wishart prior

The Bayesian VAR with Independent Normal Inverse-Wishart prior is given by:

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\Pi} + \boldsymbol{U} \qquad \boldsymbol{U} \sim MVN(\boldsymbol{0}, \boldsymbol{\Sigma}, \boldsymbol{I_T}) \tag{16}$$

where $\boldsymbol{Y}$ is $T \times N$, $\boldsymbol{X}$ is $T \times k$ with $k = Np+1$, $\boldsymbol{\Pi}$ is $k \times n$, $\boldsymbol{U}$ is $T \times N$ and MVN stands for the matricvariate normal. The prior for the autoregressive coefficients and the variance covariance matrix is:

$$vec(\boldsymbol{\Pi}) \sim \mathcal{N}(vec(\underline{\boldsymbol{\mu}_{\boldsymbol{\Pi}}}), \underline{\boldsymbol{\Omega}_{\boldsymbol{\Pi}}}) \tag{17}$$

$$\boldsymbol{\Sigma} \sim \mathcal{IW}(\boldsymbol{S_0}, v_0) \tag{18}$$

$vec(\boldsymbol{\Pi})$ and $\underline{\boldsymbol{\Omega}_{\boldsymbol{\Pi}}}$ have the same structure in A.3, and we assume $\boldsymbol{S_0} = (T-2)diag(\sigma_1^2 \ldots, \sigma_N^2)$ where we estimate σ_i^2 from univariate AR(12) regressions $v_0 = N + 2$.

A.4.2 Bayesian VAR with stochastic volatility

The BVAR with SV is given by:

$$\boldsymbol{y_t} = \boldsymbol{\Pi_0} + \boldsymbol{\Pi_1 y_{t-1}} + \ldots + \boldsymbol{\Pi_p y_{t-p}} + \boldsymbol{A}^{-1}\boldsymbol{H_t}^{0.5}\boldsymbol{\varepsilon_t} \tag{19}$$

$$\varepsilon_{it} \sim \mathcal{N}(0,1)$$

where $\boldsymbol{A^{-1}}$ is a lower triangular matrix with ones on the main diagonal, $\boldsymbol{H_t} = diag(h_{1,t}, \ldots, h_{i,t})$ is the diagonal matrix collecting the volatilities of the shocks and $\boldsymbol{\varepsilon_t}$ is a column vector collecting

the *Normal* shocks. The log-volatilities evolve according to:

$$log(h_{i,t}) = log(h_{i,t-1}) + \eta_{i,t} \qquad \eta_{i,t} \sim \mathcal{N}(0, \sigma^2_{i,\eta}) \tag{20}$$

with $i = \{PMI, \Delta U, CISS\}$ for the euro area and $i = \{CFNAI, \Delta U, NFCI\}$. As for the prior for the free elements in $\boldsymbol{A}^{-1}$ and $vec(\boldsymbol{\Pi})$ and $\underline{\boldsymbol{\Omega}_{\boldsymbol{\Pi}}}$, have the same structure in A.3 as well we assume the same Inverse Gamma prior for $\sigma^2_{i,\eta}$.

A.4.3 Quantile regression and Skew-t interpolation

Following Adrian et al. (2019) we adopt a two-step procedure to estimate the entire predictive distribution of changes in the unemployment rate as a function of real and financial risk factors. In the first step, we use predictive quantile regression to estimate the quantiles of the conditional distribution, namely:

$$\hat{Q}_{\Delta U_{t+h}|\mathcal{I}_t}(\tau) = \hat{\boldsymbol{\beta}}^{\tau}_{\mathbf{1}} \Delta U_t + \hat{\boldsymbol{\beta}}^{\tau}_{\mathbf{2}} realrisk_t + \hat{\boldsymbol{\beta}}^{\tau}_{\mathbf{3}} financialrisk_t$$

for $\tau = 0.05, \ldots, 0.95$ $realrisk_t = PMI$ and $financialrisk_t = CISS$ for the euro area and $realrisk_t = CFNAI$ and $financialrisk_t = NFCI$ for the US. Then, in the second step, the estimated quantiles we interpolate using a flexible *Skew-t* distribution, so as to obtain a complete predictive density for the dependent variable.

A.5 Forecasts metrics

Defining y the realization of the series to predict, $f(.)$ the density forecast and $F(.)$ corresponding the cumulative distribution, CRPS are defined as:

$$CRPS(f, y) = \int_{-\infty}^{\infty} PS(F(z), \mathbb{1}\{y \leq z\}) dz = \int_0^1 QS_{\alpha}(F^{-1}(\alpha), y) d\alpha \tag{21}$$

where

$$PS(F(z), \mathbb{1}\{y \leq z\}) = (F(z) - \mathbb{1}\{y \leq z\})^2 \tag{22}$$

is the Brier probability score and

$$QS_{\alpha}(F^{-1}(\alpha), y) = 2(\mathbb{1}\{y \leq F^{-1}(\alpha)\} - \alpha)(F^{-1}(\alpha) - y) \tag{23}$$

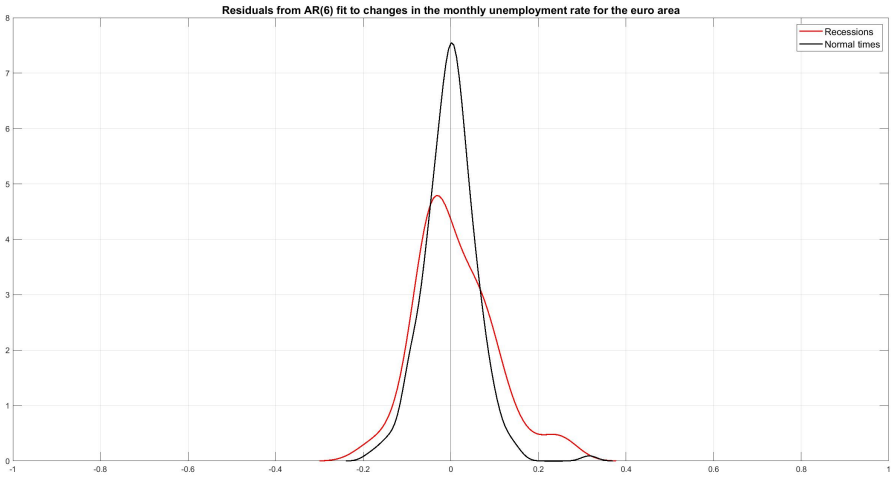
is the Quantile Score. The Quantile Weighted CRPS are computed as:

$$twCRPS = \int_{-\infty}^{\infty} PS(F(z), \mathbb{1}\{y \leq z\})^2 w(z) dz = \int_0^1 QS_\alpha(F^{-1}(\alpha), y) v(\alpha) d\alpha \tag{24}$$

where $v(\alpha) = (1-\alpha)^2$ assigns higher weights to the lower quantiles of the distribution function.

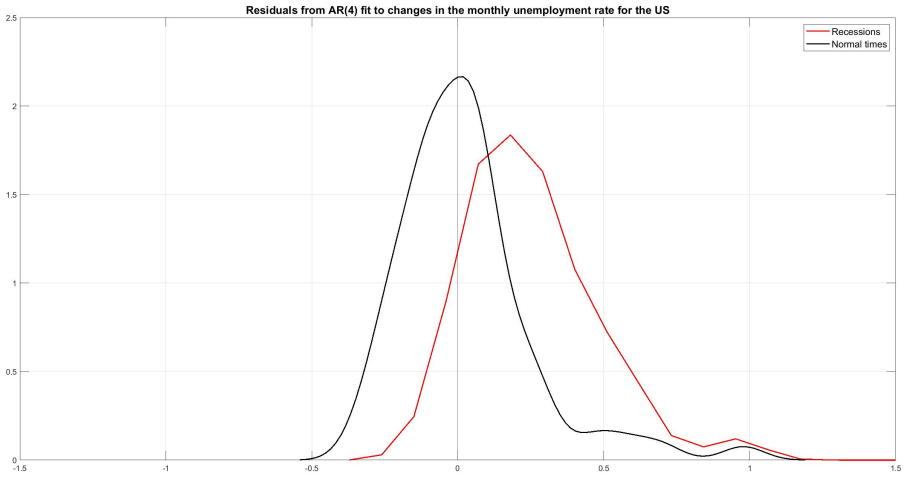
A.6 Additional figures

Figure 23: Residuals from autoregressive fit to changes in the unemployment rate in the euro area distinguishing recessions form normal times



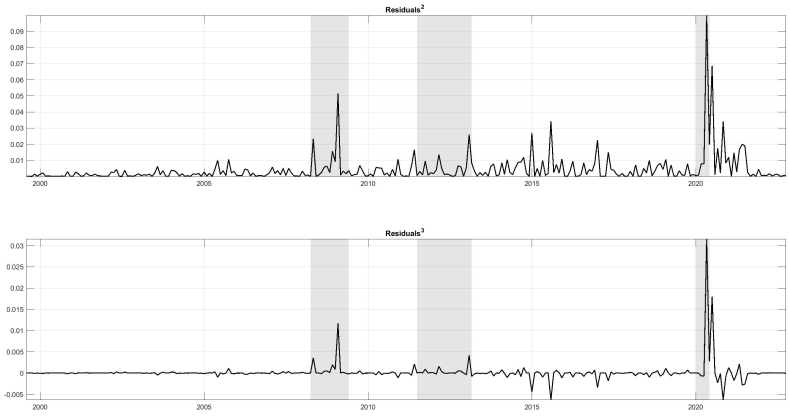
Notes: The figure shows the estimated distribution of the residuals of an autoregressive model for the changes in the unemployment rate in the euro area for recession periods, identified by the EABCN and for normal times.

Figure 24: Residuals from autoregressive fit to changes in the unemployment rate in the US distinguishing recessions form normal times



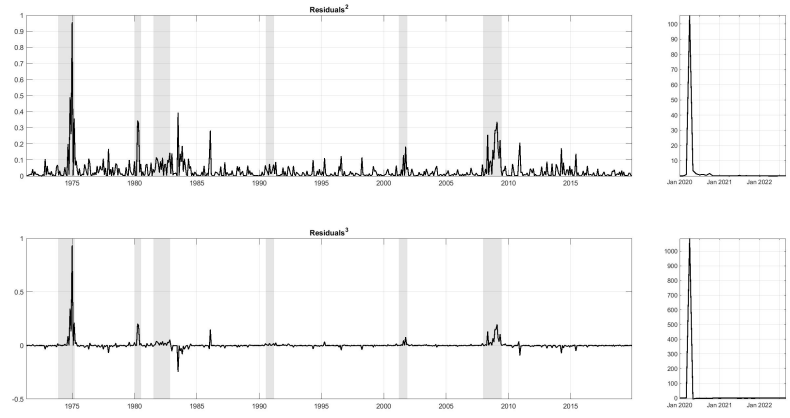
Notes: The figure shows the estimated distribution of the residuals of an autoregressive model for the changes in the unemployment rate in the US for recession periods, identified by the NBER and for normal times.

Figure 25: Squared and cubed residuals from autoregressive fit to changes in the unemployment rate in the euro area



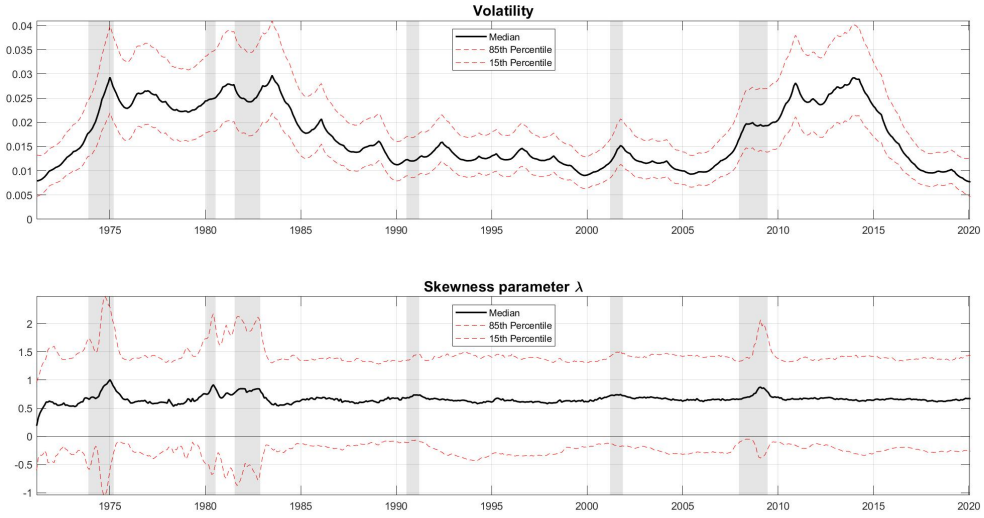
Notes: The figure shows the time series of the squared and cubed residuals from autoregressive model to changes in the unemployment rate in the euro area. The shadow bands indicate the EACN recessions periods.

Figure 26: Residuals from autoregressive fit to changes in the unemployment rate in the US distinguishing recessions form normal times



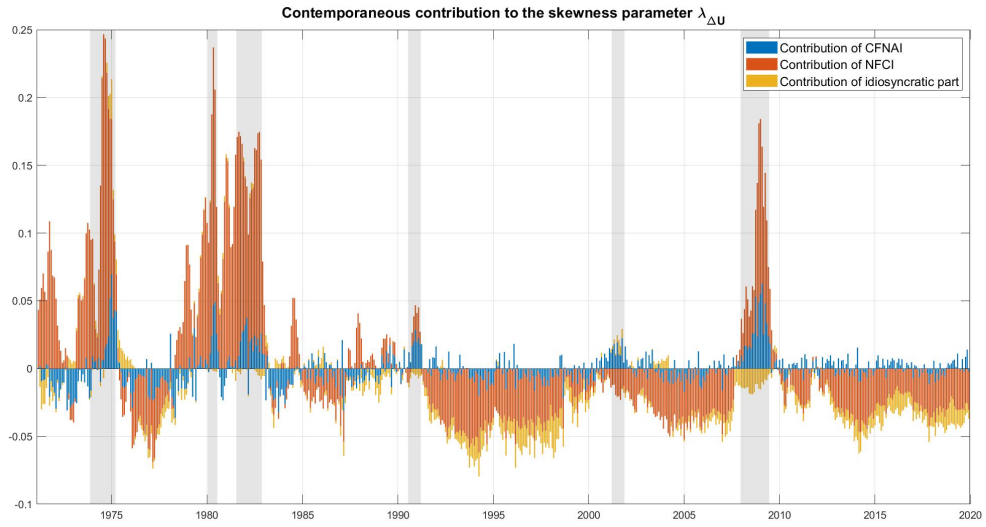
Notes: The figure shows the time series of the squared and cubed residuals from autoregressive model to changes in the unemployment rate in the US. The shadow bands indicate the NBER recessions periods.

Figure 27: Estimated volatility and shape parameter of shocks to changes in the unemployment rate for the US pre-covid



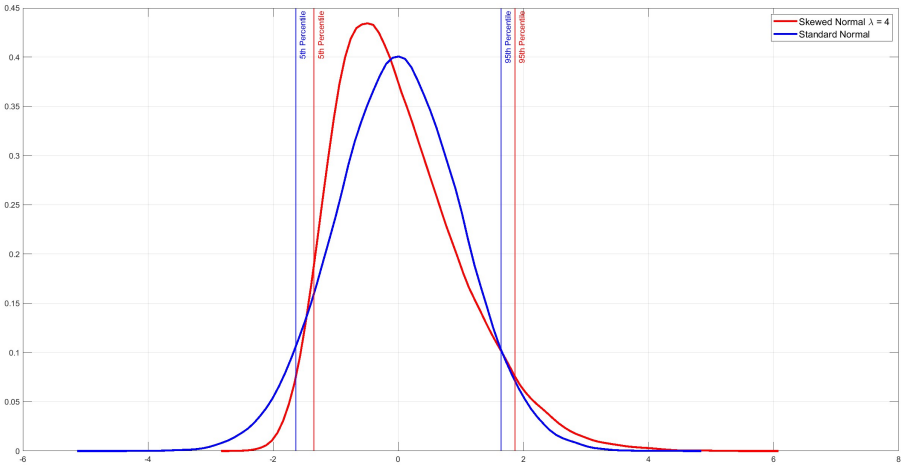
Notes: The figure shows the time series of the estimated volatilities and shape parameters of the shocks to changes in thew unemployment rate in the trivariate TVSSV VAR model for the US. The shadow bands are for the NBER recessions periods.

Figure 28: Contemporaneous contributions to the skewness parameter of the labour market shocks over time in the US pre-covid



Notes: The bar chart presents the contribution to the skewness parameter of the real activity (CFNAI) and the financial (NFCI) risk factors over time in the US.

Figure 29: *Standard normal* and *Skew normal* distributions



Notes: The figure shows the Standard normal distribution together with the Skew normal distribution with shape parameter $\lambda = 4$ re-parameterized to have zero mean and unit variance.

A.7 Additional tables

Table 7: Percentiles of $U_{t+h} - U_t$

Percentiles	**Euro Area**			**United States**		
	$h = 1$	$h = 3$	$h = 12$	$h = 1$	$h = 3$	$h = 12$
10%	-0.11	-0.28	-0.94	-0.23	-0.43	-1.12
15%	-0.09	-0.25	-0.89	-0.18	-0.33	-0.91
20%	-0.08	-0.23	-0.84	-0.15	-0.28	-0.81
25%	-0.07	-0.20	-0.80	-0.12	-0.24	-0.69
50%	-0.03	-0.09	-0.36	-0.01	-0.06	-0.29
75%	0.03	0.06	0.34	0.10	0.13	0.38
80%	0.05	0.13	0.51	0.13	0.20	0.83
85%	0.07	0.18	0.70	0.17	0.33	1.27

Notes: The table shows the percentiles of the distribution of the month on month, quarter on quarter and year on year changes in the unemployment rate for the euro area and the United States.