CIRRELT-2025-27

COVID-19, US macroeconomic tail risk, and inflation forecasts*

Mohamed Mnasri
Georges Dionne
Mohamed Jabir
Akouété-Tognikin Fenou

September 2025

Bureau de Montréal
Université de Montréal
C.P. 6128, succ. Centre-Ville
Montréal (Québec) H3C 3J7
Tél : 1-514-343-7575
Télécopie : 1-514-343-7121

Bureau de Québec
Université Laval,
2325, rue de la Terrasse
Pavillon Palasis-Prince, local 2415
Québec (Québec) G1V 0A6
Tél : 1-418-656-2073
Télécopie : 1-418-656-2624

COVID-19, US macroeconomic tail risk, and inflation forecasts*

Mohamed Mnasri[2], Georges Dionne[1]*, Mohamed Jabir[1,2], Akoueté-Tognikin Fenou[2]

1. Interuniversity Research Centre on Enterprise Networks Logistics and Transportation (CIRRELT)
2. Canada research chair in risk management, HEC Montréal

Abstract. The COVID-19 pandemic undermined well-established macroeconomic dynamics and created outliers in the distribution of main macroeconomic variables. The inclusion of extreme observations in the data substantially distorts the estimated parameters and out-of-sample forecasting from standard Bayesian Vector Autoregressions (BVARs). To capture these tail events, we use a newly developed BVAR model, with a generalized hyperbolic skew Student's *t*-distribution and stochastic volatility for the innovations, which considers both skewness and heavy tails. Our empirical study, based on 12 key US macroeconomic variables ending in 2023:Q4, shows that the pandemic created macroeconomic tail risk due to a simultaneous shift in the tail fatness of macrovariables. The COVID-19 shock generated a long-lasting increase in stochastic volatility for the real GDP, inflation, and the Fed rate. By contrast, the increase in volatility for the labor market was transient. We also find that inflation responded differently to shocks in the real output, monetary policy, and tightness in the labor market during the pandemic as compared to the pre-COVID period. Our analysis reveals a changing behavior of the Phillips curve, which steepened at the beginning of the pandemic and flattened again by its end. Finally, we compare the inflation forecasts by the proposed BVAR model with those generated by a battery of competing models. We find evidence of added forecastability in point and density forecasts of the inflation rate during the pandemic period and for multiple subsequent periods.

Keywords:

BVAR, *t*-distribution, inflation forecast, macroeconomic tail risk, stochastic volatility, Phillips curve, COVID-19

Acknowledgements. We thank Claire Boisvert and Josée Lafrenière for their important contribution to the preparation of this manuscript. The research was financed by the SCOR Foundation for Science and SSHRC Canada.

---

* Corresponding author: georges.dionne@hec.ca

Dépôt légal – Bibliothèque et Archives nationales du Québec
Bibliothèque et Archives Canada, 2025

1. Introduction

Introduced by the pioneering work of Sims (1980), vector autoregressions (VARs) are linear multivariate time-series models that capture the linear interdependencies between macroeconomic variables over time. Although they are primarily based on simple formulations, i.e., constant parameter, constant variance, normally distributed error terms, VARs gained widespread recognition and broad application in macroeconomic analysis by policymakers and forecasters. However, estimating VARs requires a lot of parameters, depending on the number of endogenous variables and the lag structure. VARs are now analyzed using Bayesian inference to overcome this practical issue related to parameter estimation.

Different from frequentist statistics, which are based on the data-generating process and treat the model's parameters as nonrandom, Bayesian methods treat the VAR parameters as unobserved random variables and provide a framework to fully characterize our knowledge about the probability distributions of these random parameters, conditional on the observed data, which is assumed to be nonrandom. Thanks to their flexibility, fast computation, and improved forecast performance, Bayesian VARs are considered benchmark models for economic analysis and forecasting, as emphasized by, among others, Doan, Litterman, and Sims (1984), Litterman (1986), and Sims and Uhlig (1991).

Since then, VARs and Bayesian VARs have been continually extended and improved to consider important features of macroeconomic data, such as time-varying parameters. Sims (1993) stipulated that stationary VAR models appear to be inadequate to model and forecast macroeconomics variables and developed a Bayesian VAR that allows for time-varying variances as well as time-varying autoregressive coefficients and found that the model produces drastically better inflation forecasts. Canova (1993) employed a multivariate Bayesian time-varying coefficients approach to improve the modeling and forecasting of exchange rates. Cogley and Sargent (2001) developed a Bayesian VAR model with random coefficients to measure parameter drift in the dynamics of US inflation, unemployment, and interest rate.

Subsequently, many researchers tried to account simultaneously for the time variation in both the coefficients and the volatilities of residuals. For example, Primiceri (2005), Cogley, and Sergent (2005), and Del Negro and Primiceri (2015) proposed a Bayesian VAR model in which regression coefficients and variance parameters are allowed to evolve according to random walks, to reflect both the time variation of the simultaneous relations among the model's variables and the heteroscedasticity of the innovations in studying the impacts of the monetary policy shocks in inflation and the unemployment rate. Time variation of the variance-covariance matrix of the innovations is modeled using a multivariate stochastic volatility framework, which is meant to capture the heteroscedasticity of the shocks and the cyclical variations in macroeconomic uncertainty. Many studies using multivariate stochastic volatility supposed that all the innovations in the model are jointly normally distributed.

A large body of macroeconomic research has documented the presence of heavy tails and asymmetries in the empirical distributions of macroeconomic variables, especially during economic downturns. Fagiolo et al. (2008) investigated the statistical properties of GDP and industrial-production growth-rate distributions in OECD countries and found that such distributions appear to be well approximated by an exponential power density, with tails much fatter than the Gaussian ones. Acemoglu et al. (2017) asserted that the normal distribution does not provide a good approximation of the distribution of aggregate macroeconomic variables at the tails and developed a theoretical model, which is defined as systematic departures in the frequency of large economic downturns from what is predicted by the normal distribution. Adrian et al. (2019) empirically modeled the full distribution of future real GDP growth as a function of current financial and economic conditions using quantile regressions and found that the estimated distribution is left-skewed during recessions and closer to being symmetric during expansions. Jensen et al. (2020) documented that the US and other G7 economies have been characterized by an increasingly negative business-cycle asymmetry over the last three decades due to an increase in the financial leverage of households and firms. Many models were used to capture the asymmetries in the distribution of economic times series, such as quantile regression methods, Markov switching, and more advanced models based on copulas. However, most of these models suppose a Gaussian distribution for residuals and constant volatility.

Moreover, most empirical studies using VAR applications are based only on a small number of variables, particularly studies relating to the impacts of monetary policy shocks on real GDP and inflation or unemployment. In their seminal paper, Banbura et al. (2010) pointed out that small VAR applications could create an omitted-variable bias, with adverse consequences for both structural analysis and forecasting. Hence, Banbura et al. (2010) emphasized the need for estimating large VAR models to capture complex relationships between macroeconomic variables, which entails a risk of overparametrization, and they proposed a novel Bayesian VAR framework based on shrinkage to overcome this issue and to improve structural analysis and forecast performance.

In sum, researchers face a twofold computational challenge in modeling complex interactions between macroeconomic time series, particularly during downturn episodes with higher tail risk and volatility: i) the need to include a large number of time series, thereby creating a risk of parameter proliferation (Banbura et al., 2010); and ii) the need to allow for time-varying coefficients and volatilities (Primiceri, 2005; Cogley and Sargent, 2005). Both challenges are efficiently accommodated with Bayesian VARs (BVARs), which rely on the Markov chain Monte Carlo framework to get a posterior distribution that is the product of the likelihood and prior distribution. Prior probability distribution comes from the researcher's subjective beliefs about the value of the parameters. Bayesian shrinkage is helpful to handle large VARs. The Bayesian framework also accommodates non-Gaussian distributions for error terms and stochastic volatility. All these successive developments to BVARs remarkably enhanced structural analyses based on impulse- response functions and improved accuracy in predicting the real-time density forecasts of macroeconomic variables, particularly during periods with enormous data movements.

Chiu et al. (2017) constructed a Bayesian VAR model where the orthogonalized shocks feature a Student's *t*-distribution as well as time-varying variance. Using US data on industrial production growth, inflation, interest rates, and stock returns, they showed that their VAR model outperforms alternatives that assume Gaussianity in terms of in-sample fit and out-of-sample forecasting, in particular during 2008. Chan (2020) introduced a class of large Bayesian VAR that allows for non-Gaussian, heteroscedastic, and serially dependent innovations. He applied his model to 20 macroeconomic variables and found that this model outperforms the standard variant with independent, homoscedastic Gaussian innovations and better accommodates large shocks such as in the 2008–2009 financial crisis.

More recently, the COVID-19 pandemic created intractable and longer-lasting economic turbulence: a surge in the inflation rate, distortions in the labor market with spikes in the unemployment rate, shortages due to constraints on supply, and an increase in commodity prices. The financial market was also deeply and adversely affected. COVID-19 created outliers and higher volatility for many key macroeconomic variables and triggered forces that might alter macroeconomic interactions going forward. Although, the COVID-19 posed challenges for estimation and forecasting with Bayesian VARs due to outliers, it was an excellent laboratory to enhance the theory and empirically validate new developments in BVARs; and so, it prompted many empirical studies (Lenza and Primiceri, 2022; Hartwig, 2024; Bobeica and Hartwig, 2023; Carriero et al., 2024a, 2024b).

Our paper belongs to the rapidly expanding literature on using large flexible BVARs, with heavy tails and stochastic volatility for the innovations (MST-SV hereafter), for macroeconomic forecasting and structural analysis (Chan, 2024). From an empirical application perspective, our contribution is related to the literature that focuses on the effects of large economic downturns on macroeconomic dynamics and transmission channels. More particularly, our model is most closely connected to and builds on the recent emerging literature that analyzes and addresses extreme observations (outliers) of the main US macroeconomic variables that were created by the COVID-19 pandemic.

Our analysis contributes to the relevant literature in several ways. First, we apply a novel and flexible BVAR framework, with multivariate skewed *t*-distribution and time-varying volatility for the innovations, to a large and recent dataset of US macroeconomic variables ending in 2023:Q4, hence, including a quite long period after the COVID-19 pandemic. The MST-SV tackles the problem of unstable estimated parameters when adding the COVID-19 observations. Second, we provide further evidence on the presence of skewness and fat tails in the distributions of key US macroeconomic variables and investigate how these non-Gaussian features evolved during the pandemic period. Third, we analyze the macroeconomic shock transmission channels during the pandemic era. In particular, we study the responses of inflation to shocks in real activity, monetary policy, and labor market conditions at different points in time during the pandemic era. Fourth, we gauge the performance of the MST-SV model in real-time forecasts of the price level and inflation rate during the pandemic, using various performance variables and various models.

Our results suggest that the COVID-19 pandemic created outliers for major US macroeconomic variables, as shown by the simultaneous decrease in the degree of freedom of the *t*-distribution of residuals, leading to a macroeconomic tail risk. Our findings reveal that the COVID-19 shock induces a long-lived increase in stochastic volatility for the real GDP, CPI, and Fed rate, but generates transient outliers for the labor market. More importantly, the stochastic volatility of the residuals for major US macroeconomic variables responded differently to the pandemic shock. We also find that the COVID-19 pandemic greatly altered how the price level, as measured by the CPI, responded to shocks in the real GDP, Fed rates, unemployment rate, and cost of employment, as compared to the pre-pandemic period. We find evidence that macroeconomic transmission channels and established interconnexions between macrovariables were hugely undermined by the pandemic. Finally, we obtain added forecastability in point and density forecasts of the price level and the inflation rate during the pandemic by considering tail fatness and time-varying volatility for innovations, as compared to standard Gaussian BVAR and Professionals' forecasts.

Our paper is organized as follows. Section 2 reviews the related recent empirical literature. Section 3 describes the real-time data used. Section 4 presents the BVAR with stochastic volatility and heavy tails. Section 5 details our empirical, in-sample and out-of-sample, analysis; and Section 6 concludes.

2 Related recent empirical literature

Bobeica and Hartwig (2023) estimated a small-scale BVAR using eurozone macroeconomic data over the sample 1980:Q1–2021:Q2 to isolate the drivers of euro area inflation. They showed that the COVID-19 observations strongly affect the parameter estimates of Gaussian BVAR models and the inflation forecasts. They then proposed using a fat-tailed distribution for the error terms. The multivariate Student's *t*-distribution allows the residuals to absorb the unusually large shocks, thus stabilizing the parameter estimates. The authors also suggested including off-model information for forecasts, such as information from the ECB Survey of Professional Forecasters, to produce more plausible forecasts during the pandemic. For the US context, Lenza and Primiceri (2022) asserted that the COVID-19 pandemic produced an unprecedented variation in many key macroeconomic variables, creating a challenge for the estimation of standard time-series models. Using US monthly data on employment, unemployment, consumption, and prices, up to June 2020, they showed that the estimation of standard VAR models including COVID-19 observations leads to misleading impulse responses with explosive error bands. They proposed downweighing the impact of abnormal observations related to the epidemic by explicitly modeling the change in residuals volatility during the COVID-19 era. They also assumed that the residual variance after May 2020 decays at a constant monthly rate to be estimated. They exploited the fact that the time of the volatility change is known in the case of COVID-19, which simplifies the model estimation. Although tractable, their procedure requires identifying exactly when the extreme observations start and estimating the decay rate of the residuals' volatility.

Lenza and Primiceri (2022) argued that dropping the extreme observations from the epidemic era may be inappropriate for forecasting because it underestimates uncertainty. Hartwig (2024) used quarterly observations on six core US macroeconomic variables, over an expanding window from 1988:Q4 to 2019:Q4 and to 2022:Q1, to investigate the ability of several generalized BVARs to cope with the extreme COVID-19 observations. Hartwig (2024) also discussed their impact on prior calibration for inference and forecasting purposes. He addressed two main issues: i) whether a multivariate heavy-tailed or common time-varying volatility error structure is better suited to account for the extreme observations related to the epidemic period, and what works better for forecasting; and ii) the extent to which the Minnesota prior is sensitive during the pandemic and how it should be recalibrated to improve forecasting ability. He found that a flexible BVAR with a combined error structure (heavy tails and time-varying volatility) is preferred to account for pandemic observations. More specifically, he ascertained that the COVID-19 shock is interpreted as a rare event rather than a persistent increase in macroeconomic volatility. Besides the error structure, he proposed an outlier-robust calibration of the standard Minnesota prior, based on the scaled median absolute deviation and median AR(*p*) residual, to alleviate the sensitivity of the parameter estimates and to improve predictability.

Carriero et al., (2024a) examined the ability of BVARs with stochastic volatility to capture tail risks in macroeconomic forecast distributions and outcomes, using US real-time data vintages from 1985:Q1 to 2022:Q1, on GDP growth, inflation, and the unemployment rate. The findings of Carriero et al., (2024a) reveal that both conventional BVAR specifications featuring time-varying volatility for residuals and BVAR models extended to feature a common volatility factor are equally able to capture tail-risk forecasts as with quantile regression methodology. These extensions can allow a contemporaneous correlation between shocks to the levels and volatilities of macroeconomic variables. Carriero et al. (2024b) viewed the extreme observations, due to the economic turbulence created by the COVID-19 pandemic, as possible outliers that are characterized by transient and infrequent increases in volatility, which influences parameter estimates and forecasts from conventional constant-parameter BVARs. Afterward, they developed BVAR models with outlier-augmented stochastic volatility that combine transitory and persistent changes in volatility to alleviate the effects of these extreme realizations. More specifically, the

BVAR models developed with stochastic volatility combine two features: i) large but infrequent volatility outliers; and ii) fat-tailed errors. The empirical results, with monthly observations for 16 US macroeconomic and financial variables for the period from March 1959 to March 2021, show the efficacy of BVAR models with outlier-augmented stochastic volatility and fat-tailed errors for mitigating the influence of outliers induced by the pandemic on parameter estimates and forecast distributions.

Even though our paper has much in common with the empirical literature mentioned above and shares its emphasis on COVID-19 outliers, it also features major differences from this literature. Complementing these empirical studies, our analyses cover a longer post-pandemic period and a larger dataset, consisting of 12 key US macroeconomic variables ending in 2023:Q4. We then use timelier and richer information on macroeconomic dynamics. Our paper is most closely related to the work of Bobeica and Hartwig (2023). However, we differ from them in several important ways. First, we use quarterly US data. Second, our methodology to account for COVID-19–related outliers is also different from their framework, which is based on a multivariate *t*-distribution and constant second-order moments for errors. Our estimated BVAR model is based on a generalized hyperbolic skew Student's *t*-distribution with stochastic volatility for the innovations, to consider both skewness and heavy tails. Third, in Bobeica and Hartwig (2023), the innovations' volatility and the heavy tails parameters are common across all variables. By contrast, our framework explicitly allows the skewness, heavy tails, and stochastic volatility to differ between macrovariables, which appears to be better suited for the pandemic. Fourth, we apply a Minnesota type prior to the autoregressive parameter matrices that shrink the regression coefficients towards univariate random walks.

3 What is the data telling us?

We use a quarterly dataset consisting of 12 US macroeconomic variables during the period 1980:Q1–2023:Q4, sourced from the FRED database of the Federal Reserve Bank of St. Louis. When choosing variables, we started with the most relevant ones, according to the ranking proposed in Jarociński and Maćkowiak (2017) and applied log transformation for some variables. We also followed recent recommendations by Bernanke and Blanchard (2025) and included the employment cost index to capture the effect of the tightness of the labor market on inflation. Our relatively large dataset might mitigate the impact of variables with abnormal dynamics in the context of unstable environments, as suggested by Rossi (2021).

Table 1 gives the different macrovariables included in the empirical analysis,

Table 1: Variables

Variable	Transformation
Price index	
Consumer Price Index (CPI)	log
Real activity variables	
Real GDP	log
Total employment	log
Unemployment rate	none
Employment cost index (ECI)	log
Economic policy uncertainty index (EPU)	log
Survey of Professional Forecasters (SPF)	none
Crude oil price	log
Financial market variables	
Federal fund rate	none
Real exchange rate	log
VIX index	log
Ten-year government bond spread	none

Figure 1 shows the data over the sample period, along with the NBER recession indicators (dashed area). Except for the real GDP, the CPI, and the ECI indices, which have been in a steady increase over time, we observe that most of the others macrovariables are evolving with a motion like a random walk, particularly the 10-year T-Bond spread, the crude oil price, the FX rate, the Fed rate, and the VIX and EPU indices. More importantly, a visual inspection of Figure 1 shows drastic changes in the behavior of the time series of the different variables during recession periods, which could create outlier observations.

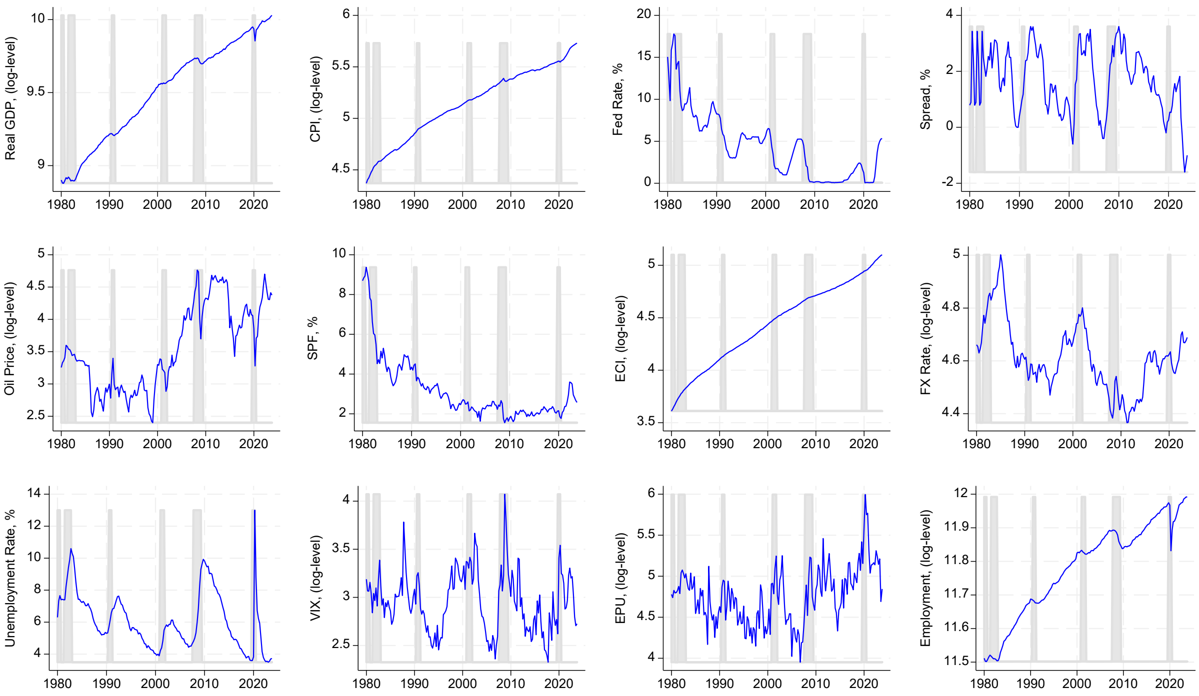

Figure 1: Dataset for the estimation of the BVAR

Note: Real GDP, CPI, price of oil, ECI index, FX rate, VIX, EPU index, and employment are in log-levels; short-term rate, 10-year government bond spread, SPF, and unemployment rate are in percentage points.

Panel A of Table 2 shows summary statistics for the macroeconomic variables in our study. We find significant skewness and excess kurtosis for most of the variables. We also run a normality test based on skewness and kurtosis that rejects the normality hypothesis for all variables, at conventional confidence levels. However, this does not necessarily imply that the error terms have skewed or heavy-tailed distributions. For deeper insight, Panel B of Table 2 reports summary statistics for the residuals from an OLS fit of a homoscedastic VAR (4) model. The skewness and kurtosis values indicate, interestingly, that residuals are non-Gaussian, which is confirmed by the normality test. Despite many of our variable being log-transformed, these statistics show unconditional skewedness as well as fat-tailed distribution in the dataset and therefore a need for flexible modeling tools that can capture these features of the data.

Table 2: Summary statistics

Variables	Obs	Mean	STD	Skew	Kurt	Min	Max	Normality Test: *p-value*
Panel A: Summary statistics of macroeconomic variables								
Real GDP (log-level)	176	9.507	0.338	-0.319	1.867	8.879	10.029	0.000
CPI (log-level)	176	5.145	0.343	-0.355	2.099	4.370	5.730	0.000
Fed rate (%)	176	4.408	3.981	1.050	4.035	0.060	17.78	0.000
Spread (%)	176	1.651	1.165	-0.293	2.378	-1.600	3.600	0.028
Oil price (log-level)	176	3.561	0.651	0.195	1.816	2.401	4.762	0.000
SPF (%)	176	3.176	1.579	2.054	7.505	1.555	9.373	0.000
ECI (log-level)	176	4.445	0.389	-0.284	2.004	3.611	5.100	0.000
FX rate (log-level)	176	4.613	0.130	0.592	3.312	4.366	5.002	0.008
Unemployment rate (%)	176	6.102	1.798	0.813	3.430	3.500	13.000	0.000
VIX (log-level)	176	2.957	0.315	0.398	3.122	2.327	4.072	0.079
EPU (log-level)	176	4.764	0.346	0.390	3.447	3.953	5.996	0.048
Employment (log-level)	176	11.782	0.139	-0.534	2.199	11.502	11.992	0.000
Panel B: Summary statistics for the residuals from the OLS fit of a homoscedastic VAR (4)								
Real GDP (log-level)	172	0.000	0,007	-2,771	24,696	-0,056	0,016	0,000
CPI (log-level)	172	0.000	0,004	-1,134	9,711	-0,022	0,011	0,000
Fed rate (%)	172	0.000	0,374	-0,213	3,772	-1,357	1,080	0,082
Spread (%)	172	0.000	0,417	0,660	5,691	-1,156	1,917	0,000
Oil price (log-level)	172	0.000	0,119	-0,378	4,839	-0,428	0,416	0,002
SPF (%)	172	0.000	0,154	-0,169	4,144	-0,608	0,465	0,039
ECI (log-level)	172	0.000	0,002	-0,457	3,687	-0,006	0,004	0,017
FX rate (log-level)	172	0.000	0,022	-0,138	2,748	-0,060	0,054	0,641
Unemployment rate (%)	172	0.000	0,580	5,716	59,121	-1,229	5,793	0,000
VIX (log-level)	172	0.000	0,165	1,359	6,684	-0,393	0,653	0,000
EPU (log-level)	172	0.000	0,180	0,452	4,465	-0,513	0,683	0,003
Employment (log-level)	172	0.000	0,009	-5,748	59,377	-0,086	0,017	0,000

For further analysis of our main variables, we apply the Hodrick–Prescott filter, which breaks down a time series into a short-run (cyclical) component and a smooth long-run (trend) component. It then produces a detrended time series. We do this detrending for our main variable, namely, the CPI (in log). Figure 2 presents the trend and cyclical components alongside their quantile-quantile (Q-Q) plots against the standard normal distribution. The Q-Q plots indicate deviations of the quantiles of both long-term trend and short-term fluctuations in inflation from the standard normal. They reveal that normal distribution significantly underestimates the frequency of large movements in inflation. Importantly, Figure 2 also reveals that huge cyclical fluctuations in inflation occur during or subsequently to recession periods (dashed area), as for the second oil-price shock in 1979 and the subsequent Volcker monetary policy of 1980–1982, the financial crisis of 2008–2009, and, more recently, the COVID-19 pandemic.

The presence of skewness and fat tails in the distributions of key macroeconomic variables calls for the use of VARs with non-Gaussian error terms, to better capture the changing macroeconomic transmission channels during the economics shocks. In fact, Lenza and Primiceri (2022) and Schorfheide and Song (2024) have reported that forecasts generated since March 2020 from homoskedastic BVARs are often distorted due to COVID-19 outliers.

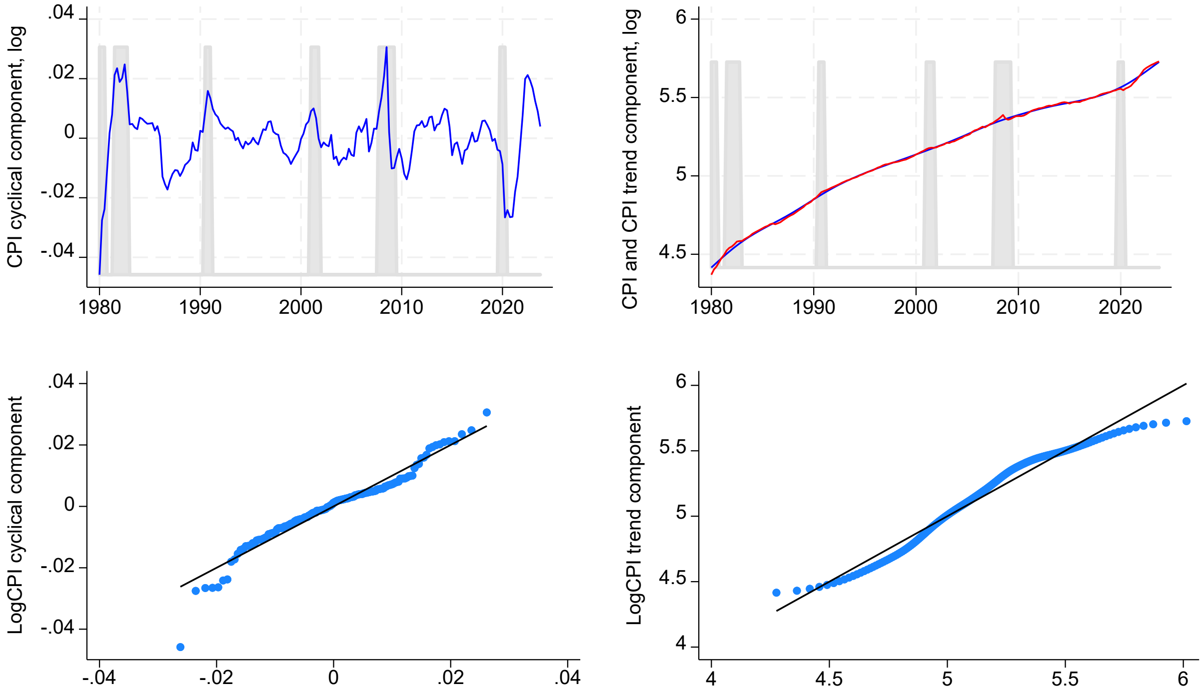

Figure 2: CPI trend and cyclical components (in log) and their Q-Q plots

Note: For the Q-Q plots, the vertical and horizontal axes correspond, respectively, to the quantiles of the CPI trend (cyclical) component against the normal probability distribution.

4 VAR model with skewness and heavy tails

In the spirit of Primiceri (2005) and Del Negro and Primiceri (2015), we use a vector autoregressive (VAR) model allowing for a time-varying variance-covariance matrix of the innovations. In fact, the standard deviations of error terms are assumed to evolve as geometric random walks, that is, stochastic volatility.

Let y_t be an $n \times 1$ vector of endogenous variables that is observed over the periods $t = 1, \ldots\ldots, T$.

Consider the following generic VAR (p) model with Gaussian stochastic volatility (Gaussian-SV) as

$$\mathrm{y}_t = c + \mathrm{B}_1 y_{t-1} + \cdots + \mathrm{B}_p y_{t-p} + u_t \tag{1}$$

where c is a k-dimensional vector of intercepts; B_j is a $k \times k$ matrix of regression coefficients with $t = 1, \ldots, p.$ u_t is a k-dimensional vector of heteroskedastic unobservable shocks associated with the VAR equations, with

$$u_t = A^{-1} H_t^{1/2} \varepsilon_t \tag{2}$$

where A^{-1} is a $k \times k$ lower triangular matrix with ones on the diagonal that describe the contemporaneous interaction of the endogenous variables,

$$A = \begin{bmatrix} 1 & 0 & \ldots & 0 \\ \alpha_{21} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ \alpha_{n1} & \ldots & \alpha_{nn} & 1 \end{bmatrix},$$

and H_t is a $k \times k$ diagonal matrix that captures the heteroskedastic volatility,

$$H_t = \begin{bmatrix} h_{1,t} & 0 & \ldots & 0 \\ 0 & h_{2,t} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & h_{k,t} \end{bmatrix}.$$

The law of motion for each element of $H_t = \mathrm{diag}(h_{1,t}, \ldots., h_{k,t})$ is specified as an independent autoregressive process (random walk) with

$$log(h_{i,t}) = log(h_{i,t-1}) + \sigma_i \eta_{it} \tag{3}$$

and

$$i = 1, \ldots, k, \qquad \eta_{it} \sim N(0, \mathrm{I}).$$

Following Karlsson et al., (2023), the reduced-form shocks u_t are modeled directly as correlated vectors of univariate skew-*t* distributions, which allow for skewness and heavy tails u_t. This gives a VAR model with multi-skew-*t* innovations.

$$u_t = (W_t - \bar{W})\gamma + W_t^{1/2} A^{-1} H_t^{1/2} \varepsilon_t, \tag{4}$$

with $\gamma = (\gamma_1, \dots, \gamma_k)'$ as a k-dimensional vector of skewness parameters, and ε_t is a k-dimensional vector of innovations that follows a multivariate Gaussian distribution with $\varepsilon_t \sim N_k(0, \mathrm{I})$. The mixing matrix $W_t = \mathrm{diag}\,(\xi_t) = \mathrm{diag}\,(\xi_{1t}, \dots, \xi_{kt})$ is a $k \times k$ diagonal matrix with $\xi_{it} \sim \mathcal{IG}\left(\frac{\nu_i}{2}, \frac{\nu_i}{2}\right)$. Degree of freedom parameters are collected in the vector $\nu = (\nu_1, \dots, \nu_k)'$. $\bar{W} = \mathrm{E}(W_t) = \frac{\nu}{\nu - 2}$ identifies whether the innovations have a zero mean.

5 Empirical illustration[1]

We now estimate[2] an empirical version of the Bayesian VAR model, with stochastic volatility and fat tails as introduced above, using quarterly US data. We employ a generous lag structure consisting of four quarterly lags to accommodate flexible dynamics between macroeconomic variables, as suggested in the relevant literature (Bernanke and Blanchard, 2025; Carriero et al., 2024a). We also estimate the model on the pre- and post-COVID periods to discover the impact of the pandemic on the behaviors of macrovariables and their interconnexions. The analysis is conducted both in- and out-of-sample.

---

[1] We use the R package developed in Karlsson et al. (2023), available at: https://github.com/hoanguc3m/fatBVARS.

[2] For each estimation at each quarter, the estimation is based on a simulation with 11,000 iterations with a burin of 1000 iterations.

5.1 In-sample analysis

The in-sample analysis allows us to delve deeply into the features of the time series of our macroeconomic dataset. Figure 3 plots the posterior means of the estimated time-varying stochastic volatility of the innovations, which presents many interesting features. Figure 3 reveals that the standard deviations of residuals vary over time, with broadly low-magnitude movements, and rise around crisis times. The stochastic volatility for the real GDP (in log) decreases during the 80s and 90s and trends upward afterward. The volatility of both CPI and EPU indices (in log) are fluctuating around an upward trend over time. The volatility of the Fed rate drops drastically during the beginning of the 80s, on the eve of the great moderation by Volcker and his unprecedented monetary policy, and it becomes low and substantially constant afterward, before rising again at the onset of the 2007–2009 financial crisis. The volatility of the 10-year T-Bond spread evolves in the same way as the Fed rate. The stochastic volatility of residuals in the VIX index equation peaks during the 2007–2009 financial crisis and during the COVID-19. The stochastic volatility of the employment rate attains its historical record during the pandemic. These findings support the use of stochastic volatility as a given feature of the residuals for US macroeconomic variables.

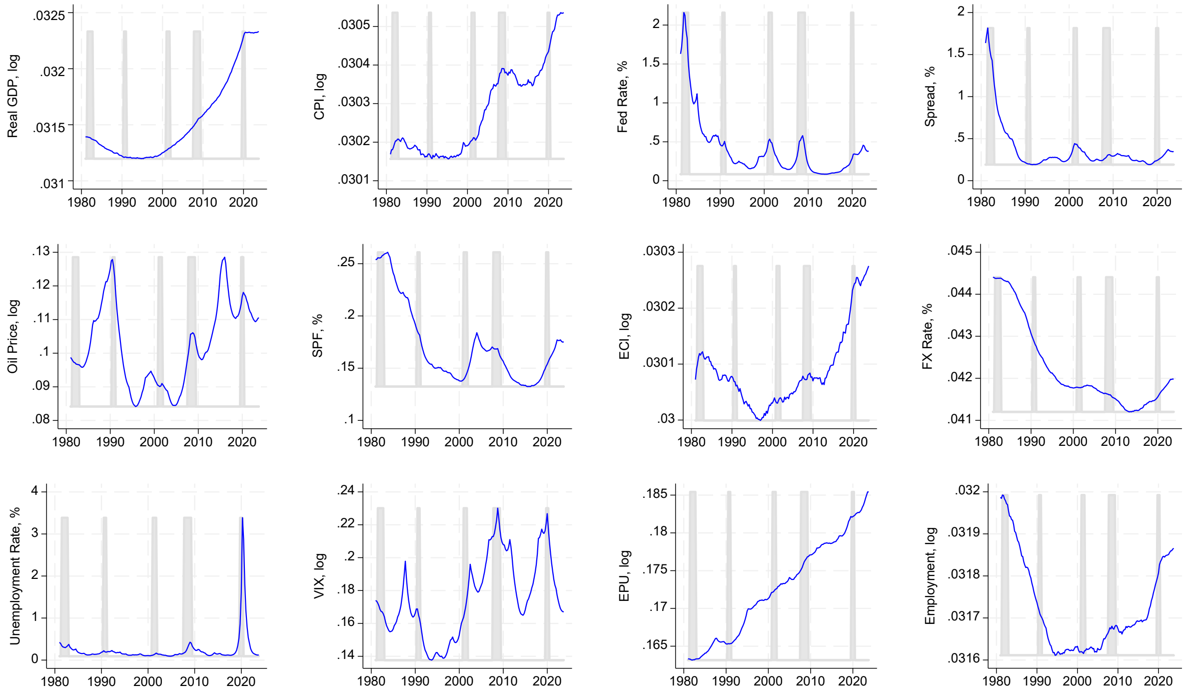

Figure 3: Posterior means of the estimated stochastic volatility

Note: Stochastic volatility $\left(e^{h_t}\right)^{1/2}$ of the residuals in the VAR model with multi-*t* skew distributed innovations and stochastic volatility (MST-SV).

Next, we analyze the skewness parameters of the posterior distributions of the residuals in the VAR models with MST with (without) stochastic volatility, namely, MST-SV (MST-noSV). Apart from the Fed rate and the 10-year T-Bond spread, Figure 4 exhibits large shifts in the marginal distributions of the skewness parameter, coming from both models. Surprisingly, skewness parameters coming from the MST-SV are centered around zero, although with some spread. Conversely, residuals from the MST model appear to have nonzero skewness parameters, especially for the real GDP, oil price, SPF, employment cost index, and unemployment rate. In sum, these findings indicate that accounting for time-varying volatility may preserve the symmetry of the presumed *t*-distribution of residuals. In fact, macroeconomic shocks translate to a higher stochastic volatility without altering the distribution symmetry of residuals. We will come back to this observation.

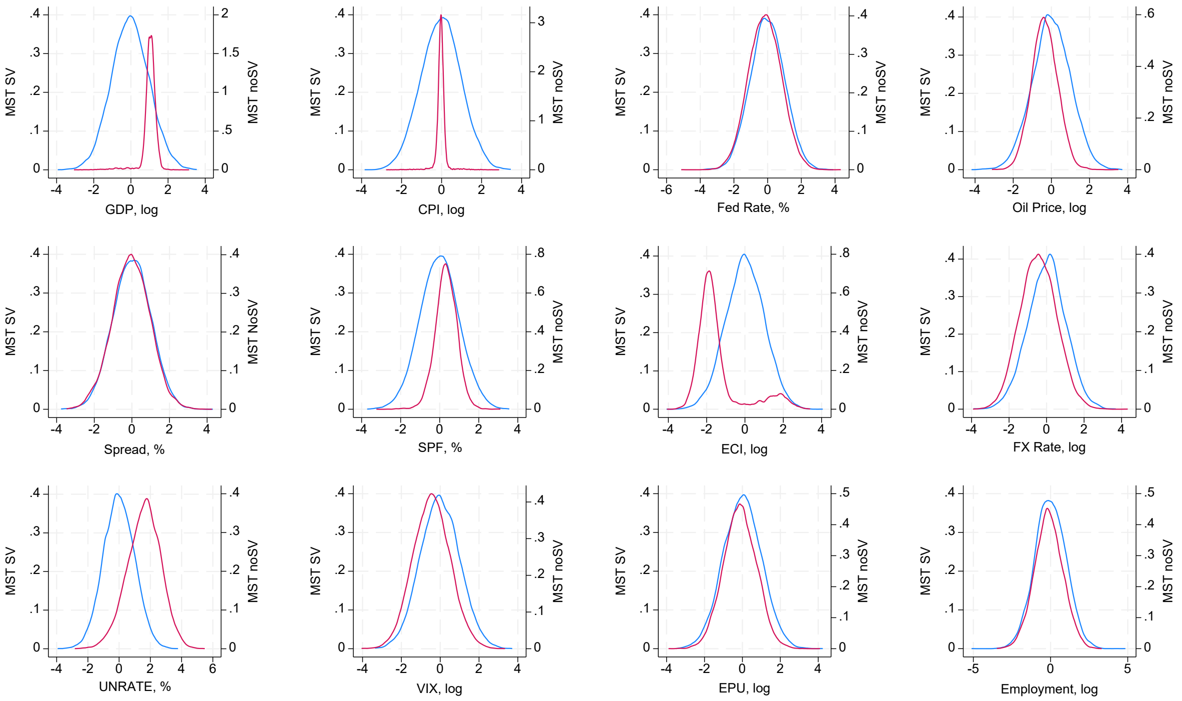

Figure 4: Posterior densities of the skewness parameters of the VAR models

Note: MST without stochastic volatility (in red) and MST with stochastic volatility (in blue).

5.2 Macroeconomic tail risk and volatility during the pandemic era

In this section, we investigate the extent to which major US macroeconomic variables were distorted by the COVID-19 pandemic. We then re-estimate the MST-SV model recursively over an expanding estimation window, beginning from 1980:Q1 and until 2019:Q4, 2020:Q2, 2020:Q4, 2021:Q2, 2021:Q4, and 2022:Q2, respectively. We uncover how the pandemic observations affect the stochastic volatility of residuals and the tail fatness and asymmetry of their distributions. For brevity, we limit the analysis in this section to the four major US macrovariables: real GDP (log-level), CPI (log-level), Fed rate, and unemployment rate, as in Chan (2020).

Figure 5 depicts the medians of the time-varying stochastic volatility of residuals, estimated at different dates at the beginning and during COVID-19. Interestingly, Figure 5 exhibits different patterns for the stochastic volatility of different variable/time intersections. At the onset of the COVID-19 pandemic, the stochastic volatility increases during 2020, compared to its level at 2019:Q4, for the real GDP, Fed rate, and unemployment rate (UNRATE), indicating that the pandemic shock translates into an increase in macroeconomic volatility during 2020. Surprisingly, the stochastic volatility for the CPI residuals shows a completely different pattern by decreasing during 2020, the first year of the pandemic. This may reveal a delay in the inflation response to COVID-19’s impacts, in terms of commodity and goods prices surging as a result of pervasive shortages and lockdowns. During 2021 and the first half of 2022, the stochastic volatility for the real GDP and the Fed rate continues to be higher than the pre-pandemic level. The stochastic volatility for the CPI reverses, and increases during 2021, to attain its pre-pandemic level and then exceed it during the first half of 2022.

These findings show, curiously, that the MST-SV model treats the pandemic effects on the real GDP, the Fed rate, and the CPI as a persistent increase in uncertainty, and not as a rare and transient shock. The stochastic volatility for the unemployment rate decreases during 2021 and the first half of 2022, and then returns to its pre-pandemic level, indicating that the pandemic shock in the labor market is assumed to be transitory. This latter finding corroborates the results by Carreiro et al., (2024b), who found a similar pattern for the stochastic volatility of the payroll growth rate. Remarkably, our findings reveal that the pandemic shock translates to a persistent and long-lasting increase in the time-varying volatility for the real GDP, CPI, and Fed rate, but short-lived outliers for the labor market. It is worth mentioning that our results provide evidence of an important fact: that the stochastic volatility of the residuals for major US macroeconomic variables responded differently to the pandemic shock.

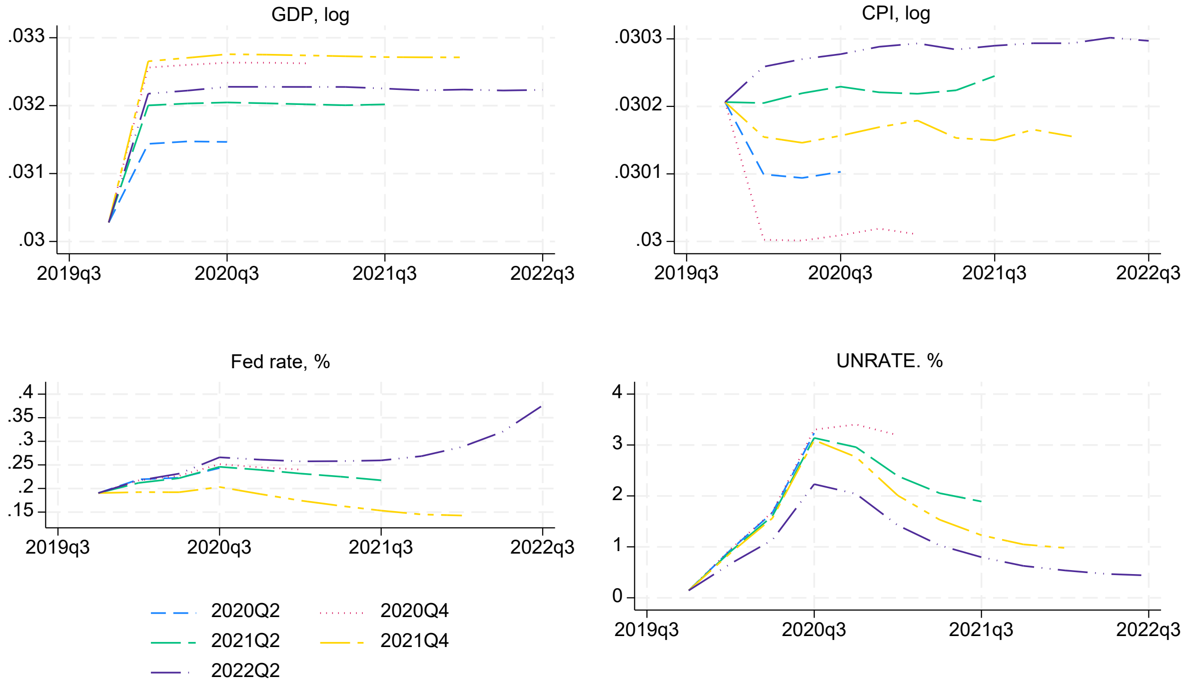

Figure 5: Posterior medians of the estimated stochastic volatility

Note: Stochastic volatility $(e^{h_t})^{1/2}$ of the residuals in the model with multi-*t* skew distributed innovations and stochastic volatility. Medians were obtained from different data samples ending at different dates as indicated in the figure. The origin for each panel corresponds to the estimated volatility in 2019:Q4 for comparison.

We turn to the macroeconomic tail risk and investigate how the posterior distributions of the degrees of freedom (DoF) evolved during the pandemic. Recall that the lower the degree of freedom parameter ν, the stronger the *t*-distribution departs from Gaussianity and the fatter are its tails. Figure 6 presents de posterior densities of the DoF obtained over an expanding estimation window from 1980:Q1 to 2019:Q4 (left-hand panel) and to 2022:Q2 (right-hand panel) for the four key macroeconomic variables. Noticeably, there is no evidence of tail fatness in the distribution of residuals before the pandemic, as indicated by the high DoF[3]. In 2022:Q2, the densities move to the left and the DoF decreases significantly, particularly for the unemployment rate. It appears that the pandemic created macroeconomic tail risk, comparatively to the pre-pandemic era, but with a delay. We observe that tail risk is more pronounced for the labor market, which corroborates the huge spike in the unemployment rate during the pandemic, as depicted in Figure 1. It is worth mentioning that tail fatness in the distributions of shocks concretized simultaneously during 2022 for the four key variables, thereby generating tail co-movements. This is in line with the macroeconomic tail risk discussed by Acemoglu et al. (2017).

In unreported results, we did the same estimations with the expanding samples at different time points over the pandemic period, and we investigated the posterior distributions of the skewness parameters for the residuals. Overall, we find qualitatively similar shapes in the posterior densities of the skewness parameters, as depicted in Figure 4, with more concentration around zero and less spread. In conclusion, results substantiate that the macroeconomic shock due to the COVID-19 pandemic materialized first in a higher stochastic volatility of the model's residuals, which later created tail fatness in their distributions; however, the symmetry parameters were unaffected on average.

---

[3] The densities estimated with samples ending in 2020 and in 2021 give qualitatively the same shape as for 2019:Q4.

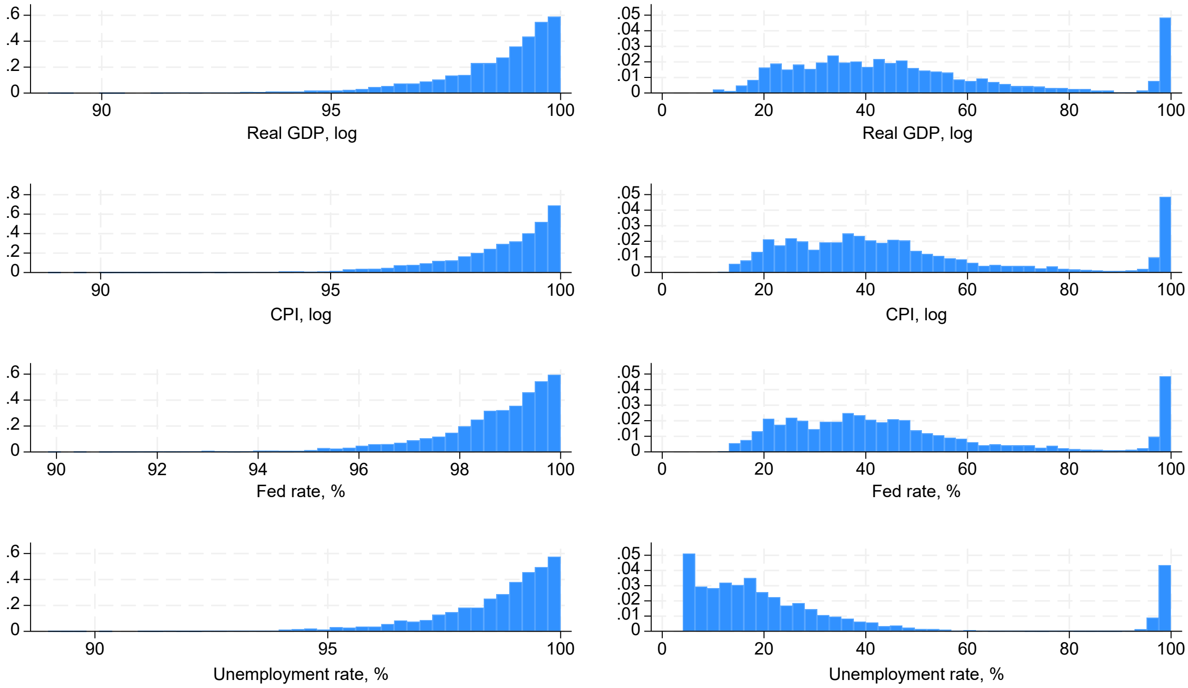

Figure 6: Posterior densities of the degrees of freedom (DoF) of the residuals

Note: Residuals from the VAR model with multi-*t* skew distributed innovations and stochastic volatility (MST-SV). Left-hand panel exhibits the DoF estimated from a sample ending in 2019:Q4. The right-hand panel is for DoF for a sample ending in 2022:Q2.

5.3 COVID-19 and the transmission of macroeconomic shocks

Following COVID-19, an exceptional contraction in real activity led to the complete breakdown of established interactions and statistical relationships between key economic variables. Macroeconomic transmission channels were also altered. In this section, we investigate the extent to which these interactions and channels were modified, by examining the response of the price level, measured by the CPI (in log-level), to shocks in the real output, to monetary policy shocks, and to labor market shocks, as suggested by Bernanke and Blanchard (2025). To do so, we use the impulse response functions (IRFs) estimated with our BVAR model with multi-*t* skew distributed innovations and stochastic volatility. The IRFs are estimated recursively over an expanding estimation window, from 1980:Q1 to 2019:Q4, 2020:Q2, 2021:Q2, and 2022:Q2, respectively, to gain greater insight into the responsiveness of the price level to macroeconomic shocks before and during the pandemic. Figure 7 depicts the IRFs of the CPI (in log-level) to a one-standard-deviation shock in the real GDP, Fed rate, employment cost index, and unemployment rate, estimated over the different expanding windows. Figure 7 shows that the inclusion of the COVID-19 observations noticeably affects the IRF estimates, as compared to pre–COVID-19. The transmission of the real GDP shock to the price level spikes with the inclusion of the 2020:Q2 observations, when the US real GDP slumped deeply due to global lockdowns. Afterward, the reaction of the CPI to real GDP shocks gradually weakens and stabilizes at lower levels than pre-pandemic and flattens in mid-2022. Bobeica and Hartwig (2023) found qualitatively similar reactions of the price level to real GDP shocks for the European context, using a Bayesian VAR with multivariate *t*-distributed errors without stochastic volatility.

The impulse responses of the CPI to a contractionary monetary shock are significantly positive over the first 10–12 quarters, before becoming strongly negative afterward. It appears that the effects of the monetary policy on price levels are not immediate but materialize with some delay. These findings corroborate the results in the relevant literature using the extended narrative measure of Romer and Romer (2004) to identify monetary policy shocks (see also Miranda-Agrippino and Giovanni, 2021). The inclusion of the COVID-19 observations gives different shapes for the IRF of the price level to monetary shocks. In the early stage of the pandemic, the positive effect is weaker and shorter, and the negative effect is stronger as compared to pre–COVID-19. In the middle stage of the pandemic, the IRF of the CPI flattens and becomes positive for all the horizons. In mid-2022, the CPI responsiveness to the monetary policy shocks become stronger than during the pre-pandemic era, either positive during the first periods or negative afterward. Recall that, during 2022, the Fed fund rate increased gradually after historical low levels, and the CPI spiked. Very importantly, the transmission of monetary policy shocks depends on the state of the business cycle, that is, if the economy is in a recession or expansion cycle, as was suggested by Tenreyro and Thwaites (2016). In fact, we find evidence that contractionary policy shocks are more powerful than expansionary shocks (compare the IRF in 2020:Q2 versus the IRF in 2022:Q2).

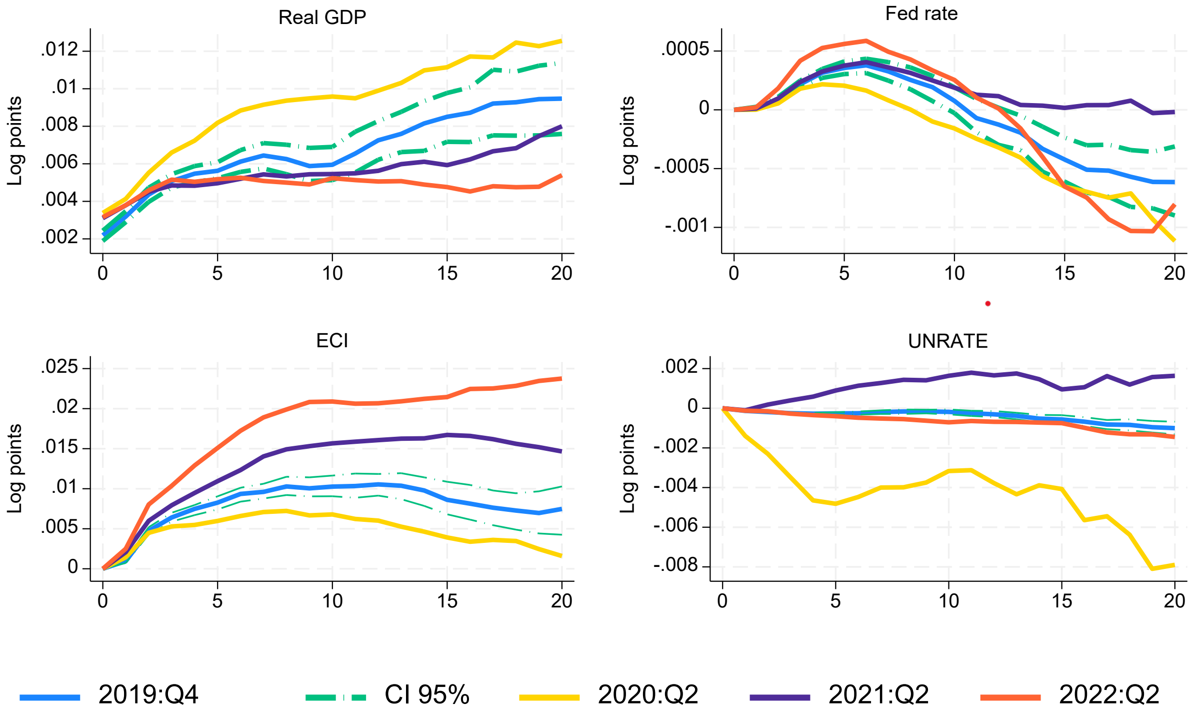

Figure 7: Impulse response functions in a BVAR with multi-*t* skew distributed innovations and stochastic volatility

Note: Medians of the IRF of the CPI (in log-level) to a one-standard-deviation shock in real GDP, Fed rate, employment cost index (ECI), and unemployment rate (UNRATE) are plotted. The 95% credible interval is for the estimation window until 2019:Q4. IRFs are estimated over 20 quarters ahead.

Next, we investigate the impacts of the tightness of the labor market on the price levels and inspect how the Phillips curve, linking inflation and the unemployment rate, evolved before and during the pandemic. As in Bernanke and Blanchard (2025), we find that the price level reacts positively to labor tightness, measured by a one-standard deviation shock in the employment cost index (ECI). An increase in nominal wages has twofold impacts on the price level: i) higher demand, and ii) higher cost of goods and services. Figure 7 shows that the reaction of the CPI to a one-standard-deviation shock in the ECI significantly decreases in the early stage of the pandemic compared to its level pre–COVID-19, due to the unprecedented contraction in activity at the beginning of 2020. Afterward, the IRFs gradually strengthened and exceeded their pre-pandemic levels, as the economy recovered and the labor market became overheated due to monetary and fiscal policies.

5.4 COVID-19 and the labor market

Turning now to the Phillips curve, which links unemployment and inflation and illustrates the trade-off between maintaining price stability and achieving full capacity utilization. It appears that the Phillips curve was flat during the pre-pandemic era. In fact, IRFs of the CPI to the unemployment shocks are flat around zero. The insensitivity of inflation to changes in unemployment during the last two decades has led many economists to suggest that the Phillips curve had disappeared—or was “hibernating,” that is, its slope approached zero (Blanchard, 2016; Stock and Watson, 2020; Hazell et al., 2022).

COVID-19 reawakened the Phillips curve. The reaction of the price level is strongly negative at the beginning of the pandemic, indicating a steepening of the Phillips curve, that is, higher unemployment reduces inflation. Recall that at that time of 2020:Q2, the unemployment rate reached historic records. The Phillips curve then surprisingly reverses in mid-2021, when the unemployment rate remained high and inflation was surging. Rossi et al (2024) found qualitatively similar evidence that the Phillips curve is becoming alive again, using estimation data to the end of 2021. In the late stage of the pandemic, the Phillips curve hibernates again, in mid-2022, when the IRF was hovering close to zero. The changing behavior of the Phillips curve was shown early on by Stock and Watson (1999), who argued that the coefficients of the Phillips curve are time-varying.

In sum, our findings indicate how transmission channels of macroeconomic shocks were altered by the COVID-19 pandemic. Moreover, these transmissions channels were not constant during the pandemic era but changed from one period to another, reflecting the moving macroeconomic dynamics and labor market conditions of the period.

5.5 Real-time forecast evaluation

To assess the out-of-sample predictive accuracy of the BVAR with multi-*t* skew, distributed innovations, and stochastic volatility (MST-SV), we conducted a recursive out-of-sample forecasting exercise using expanding time windows. Starting from 1989:Q4 to 2023:Q4, the MST-SV model was re-estimated quarterly using all available data from 1980:Q1. The goal was to forecast the CPI (in log-level) for 12 quarters ahead (t+1 to t+12), where t represents the current quarter. To evaluate the MST-SV model's performance, we conducted a parallel analysis using a rival standard Gaussian Bayesian VAR as a benchmark model. This simpler model assumes a Gaussian error distribution and excludes stochastic volatility. The metric used to evaluate the point forecasts is the loss functions of the mean squared and absolute forecast errors, denoted by MSFE and MAFE, respectively, and calculated as follows for a variable i at h steps ahead, for $h = 1, \ldots, 12$:

$$MSFE_{i,h} = \frac{1}{T_1 - T_0 - h + 1} \sum_{t=T_0}^{T_1 - h} \left(\bar{y}_{i,t+h|t} - y_{i,t+h} \right)^2 \tag{5}$$

and

$$MAFE_{i,h} = \frac{1}{T_1 - T_0 - h + 1} \sum_{t=T_0}^{T_1 - h} \left| \bar{y}_{i,t+h|t} - y_{i,t+h} \right| \tag{6}$$

where T_0 is the last observation in the first estimation sample, namely, 1989:Q4; T_1 is the last observation on variable i, namely, 2024:Q4; and $\bar{y}_{i,t+h|t}$ is the median of the posterior predictive sample for step h using the expanding windows up to time t. $y_{i,t+h}$ is the actual realization of variable *i* at *h* steps ahead. The loss function is then calculated as the ratio of the MSFE (MAFE) for the MST-SV model divided by the MSFE (MAFE) of the Gaussian BVAR, that is, the benchmark model. Hence, if the ratio is less than 1, this means that the loss of the MST-SV forecasts is smaller than that of the benchmark forecasts, and vice versa if the ratio is greater than 1.

Table 3 reports the point forecast results for the three horizons 4, 8, and 12 quarters ahead, or alternatively, 1 year, 2 years, and 3 years ahead. Table 3 shows evidence of added forecastability by the MST-SV model over the standard Gaussian BVAR in forecasting the price levels at different forecasting horizons. The longer the forecasting horizon, the larger is the improvement in point forecasts by the MST-SV, relative to the Gaussian BVAR. Il appears then that considering heavy tails and skewness in the distribution of residuals and allowing for time-varying variances substantially improves the point forecasts of inflation.

Table 3: Relative improvements in MSFE and MAFE relative to the standard Gaussian BVAR model

	Relative MSFE			Relative MAFE		
	$h = 4$	$h = 8$	$h = 12$	$h = 4$	$h = 8$	$h = 12$
Loss function	0.471	0.147	0.039	0.594	0.373	0.258

Note: Results for the forecasting of the CPI (in log-level). Forecasting horizons are 4. 8, and 12 quarters ahead.

It is worth mentioning that the MSFE and MAFE loss functions evaluate forecast performance averaged over the sample period. However, it is well-known that macroeconomic time series are prone to instabilities, particularly during turmoil periods when several macroeconomic relationships are changing drastically and transmission channels are deeply altered, as during the 2008–2009 recession or COVID-19 pandemic. In such a context, variations in unstable environments are observed where structural breaks are often found in macroeconomic variables over a relatively long-time span. For this context Giacomini and Rossi (2010) suggested that selecting a model with the best average forecasting performance could lead to incorrect selection decisions. In fact, traditional tests of forecast evaluation are not reliable in the presence of instabilities because they assume stationarity, which is violated in the presence of instabilities. Rossi (2021) raised an important question: “How can one assess models’ ability to accurately predict the target variable when the predictive ability changes over time?”

We use the fluctuation test developed by Giacomini and Rossi (2010) to compare the relative forecasting performance of our competing models over the entire time path in the presence of possible instabilities. The fluctuation test provides useful information that is lost when one is using averaged performance measures. The null hypothesis of the fluctuation test is an equal predictive ability at each point in time for the competing models, and it is rejected when the test statistic is higher than a critical value. The alternative is inequal predictive ability. Our competing models are the MST-SV against a benchmark Gaussian BVAR. We run the fluctuation test to compare the forecasting performance of the MST-SV and the Gaussian BVAR models for the three forecasting horizons, namely 4, 8, and 12 quarters ahead.

As an illustration, Figure 8 plots the fluctuation test statistics of Giacomini and Rossi (2010) for the forecasts of CPI (in log-level) for the three horizons,[4] respectively. The test statistics are 6.70, 9.54, and 7.30, respectively, for the horizons 4, 8, and 12 quarters ahead. Overall, these test statistics are higher than the critical value of 3.179. We then reject the null hypothesis of equal forecasting ability and consider the evidence of the alternative, i.e., the dominance of the Bayesian VAR with skew-*t*-distributed residuals and stochastic volatility, over a rival Gaussian BVAR with constant volatility. Moreover, a visual inspection of Figure 8 reveals that the fluctuation test shows that the MST-SV has a varying forecasting efficiency over the time path. In fact, the MST-SV model could perform better at certain time points and might have a similar forecasting accuracy as the Gaussian BVAR in other subperiods. More importantly, two remarks arise from Figure 8. The forecasting performance of the MST-SV appears to be strongly better i) around turmoil periods, such as the 2007–2009 financial crisis and COVID-19 pandemic; and ii) for longer forecasting horizons. As expected, this is essentially due to the ability of BVAR models, with the multi-skew *t*-distributed innovations and stochastic volatility, to better capture outliers and the higher volatility in macroeconomic variables created by turmoil episodes.

---

[4] We use a two-sided test with a confidence level of 95% and a rolling window of 20 quarters (5 years). Using others rolling windows gives qualitatively the same results. Giacomini and Rossi's (2010) test rejects the null hypothesis of equal predictive ability when the test statistic is outside the band lines (2-sided alternative). When the test statistic is below the lowest band line, the first model forecasts significantly better. Our first model is the MST-SV.

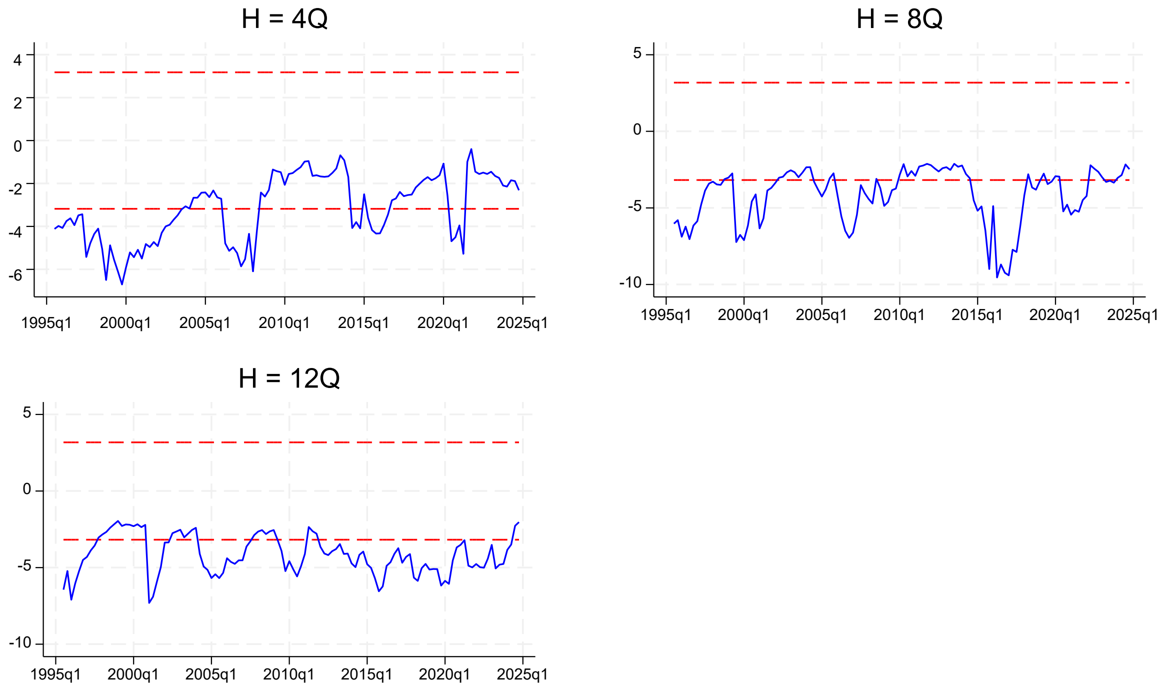

Figure 8: Fluctuation tests of Giacomini and Rossi (2010)

Note: For forecasts of quarterly CPI (in log-level) from the MST-SV model using a recursive window (our first model), the forecasts start from 1989:Q4 to 2023:Q4 using all available datasets from 1980:Q1. The rival model is a Gaussian BVAR (our second model). Dashed lines are critical values, at the 95% confidence level. If a test statistic (solid lines) is inside the band lines at a time point, then the null hypothesis of equal predictive ability holds. When a test statistic (solid lines) is below the lowest band line, the MST-SV forecasts are significantly better than the standard Gaussian BVAR forecasts.

5.6 How does COVID-19 impact inflation forecasting?

The fluctuation test of Giacomini and Rossi (2010) reveals that the MST-SV model has time-varying forecasting efficiency and outperforms the standard Gaussian BVAR, particularly at the onset of the financial crisis and the COVID-19 pandemic, for forecasting the CPI (in log-level). To gain deeper insight into the magnitude of COVID-19's impact on the forecasting ability of our two competing models, we rerun the recursive forecasting exercises of the log CPI at six time points before, during, and at the end of the pandemic, namely, 2019:Q4, 2020:Q2, 2020:Q4, 2021:Q2, 2021:Q4, and 2022:Q2. We predict the log CPI for 4 and 8 quarters ahead. Because we have only six estimation points, we rely more on forecast variables based on density forecast measures, namely: i) the average log predictive likelihood (ALPL), ii) the average continuous rank probability score (ACRPS), and iii) the average continuous rank probability score with quantile weighting (qwACRPS). CRPS is obtained as the quadratic difference between the predictive cumulative distribution function and the empirical distribution of the variable. CRPS is less sensitive to outliers than the log predictive likelihood (Clark and Ravazzolo, 2015). The ALPL can be written as follows:

$$ALPL_{i,h} = \frac{1}{T_1-T_0-h+1}\sum_{t=T_0}^{T_1-h} \log p\left(y_{i,t+h}^{o}|y_{1:t}\right), \tag{7}$$

where $p\left(y_{i,t+h}^{o}|y_{1:t}\right)$ is the h-step ahead posterior predictive density function evaluated at the realization of variable i. Higher values indicate that actual observations are more likely under the predictive density and, hence, a better density-forecasting performance by the model. The ACRPS is equal to

$$ACRSP_{i,h} = \frac{1}{T_1-T_0-h+1}\sum_{t=T_0}^{T_1-h}\left[-\mathrm{E}_f\left|y_{i,t+h|t}-y_{i,t+h}^{o}\right|+0.5\mathrm{E}_f\left|y_{i,t+h|t}-y_{i,t+h}^{\prime}\right|\right] \tag{8}$$

where f is the predictive density of the variable $y_{i,t+h|t}$. $y_{i,t+h|t}$ and $y_{i,t+h}^{\prime}$ are independent random draws from f. We randomly generated 10,000 draws from f using Monte Carlo simulations. Lower values indicate that actual observations are more likely under the predictive distribution and, hence, a better density-forecasting performance by the model. It is worth mentioning that these scoring rules could give different answers regarding the best-performing model in the event of extreme observations.

Gneiting and Ranjan (2011) developed the qwCRPS as a proper scoring function of the entire predictive density to measure the density forecast accuracy. qwCRPS is derived from a weighted sum of quantile scores for a range of quantiles, and it allows higher weighting for specified tails to emphasize selected portions of the density. The qwCRPS is given by

$$qwCRSP_{i,h} = \frac{1}{T_1-T_0-h+1}\sum_{t=T_0}^{T_1-h}\left[\frac{2}{J-1}\sum_{j=1}^{J-1}\upsilon(\tau_j)QS_{\tau_j,i,t+h}\right] \tag{9}$$

with $J = 20$ and $\tau_j = \frac{j}{J} = 0.5,0.10,0.15,\ldots,0.90,0.95$. We considered both a left-tail and a right-tail weighting. The weighting function for the left-tail weighted version (qwCRPS-left) is set to $\upsilon(\tau_j) = \left(1-\tau_j\right)^2$ and set to $\upsilon(\tau_j) = \tau_j^2$ for the right-tail weighted version (qwCRPS-right).

Tables 4 and 5 report performance valuation for the density forecasts of the log CPI by the MST-SV model relative to a benchmark standard Gaussian BVAR for four and eight-quarters ahead, respectively. For the different variables selected, the relative improvement in the density forecast is computed as the difference between the MST-V and the benchmark model. Entries greater than 0 indicate that the MST-SV model performs better at forecasting price level (log CPI)[5]. Overall, the results indicate the outperformance of the MST-SV model over the standard Gaussian BVAR in forecasting inflation during the pandemic for both horizons, namely, four and eight quarters ahead. However, the relative performance of the MST-SV model is stronger in predicting inflation eight quarters ahead. It is then evident that forecasters are better off using models permitting flexible modeling of innovations in terms of both fat tails and, potentially, asymmetry. It may be worth noting that the relative outperformance of the MST-SV model becomes slightly lower when COVID-19 observations are added, until 2021:Q2. Afterward, the MST-SV model gains in relative performance when the macroeconomic shocks due to COVID-19 are well-established and

---

[5] For the CRPS, qwCRPS-right and qwCRPS-left, the relative performance could be computed as the ratio of the metric of the MST-SV model over the benchmark. Entries less than 1 indicate that the MST-SV is better. We did that and found ratios largely below 1.

crystallized in the parameters' estimation. This corroborates our findings from the previous sections relating to the delay tail fatness and asymmetry parameters' response to the COVID-19 shock.

Table 4: Relative improvements in density forecasts of CPI for four quarters ahead (in log-level)

	2019:Q4	2020:Q2	2020:Q4	2021:Q2	2021:Q4	2022:Q2	Average
				Log Predictive Likelihood (LPL)			
Gauss BVAR	-0.752	-0.663	-0.620	-0.616	-0.625	-0.606	-0.647
MST-SV	1.264	1.116	1.156	1.097	1.120	1.934	1.281
				CRSP			
Gauss BVAR	-0.212	-0.191	-0.182	-0.185	-0.181	-0.180	-0.189
MST-SV	0.174	0.120	0.118	0.116	0.117	0.154	0.129
				qwCRSP-right			
Gauss BVAR	-0.074	-0.062	-0.060	-0.060	-0.061	-0.062	-0.063
MST-SV	0.050	0.043	0.041	0.041	0.041	0.052	0.045
				qwCRSP-left			
Gauss BVAR	-0.072	-0.074	-0.069	-0.068	-0.068	-0.062	-0.069
MST-SV	0.050	0.044	0.042	0.039	0.041	0.053	0.045

Note: For four quarters ahead by the MST-SV model relative to the benchmark model, that is, a standard BVAR model with Gaussian innovations and without stochastic volatility. Forecasts start at the mentioned dates and use all available data from 1980:Q1. The density forecast variables are log predictive likelihood (LPL), CRPS, qwCRSP-right, and qwCRSP-left. The first line in each panel reports the density forecast metric of the benchmark model. The figures for the MST-SV model are the relative improvements calculated as the differences of the different variables of the MST-SV model minus the ones of the benchmark models. Entries greater than 0 indicate that the MST-SV model is better.

Table 5: Relative improvements in density forecasts of CPI for eight quarters ahead (in log-level)

	2019:Q4	2020:Q2	2020:Q4	2021:Q2	2021:Q4	Average
Log Predictive Likelihood (LPL)						
Gauss BVAR	-1.568	-1.547	-1.412	-1.371	-1.404	-1.460
MST-SV	1.336	1.236	1.207	1.268	1.263	1.261
CRSP						
Gauss BVAR	-0.619	-0.574	-0.511	-0.491	-0.482	-0.535
MST-SV	0.476	0.410	0.380	0.353	0.351	0.394
qwCRSP-right						
Gauss BVAR	-0.210	-0.184	-0.163	-0.157	-0.166	-0.176
MST-SV	0.164	0.141	0.123	0.117	0.123	0.136
qwCRSP-left						
Gauss BVAR	-0.217	-0.218	-0.185	-0.176	-0.176	-0.194
MST-SV	0.162	0.153	0.128	0.124	0.123	0.138

Note: For eight quarters ahead by the MST-SV model relative to the benchmark model, that is, a standard BVAR model with Gaussian innovations and without stochastic volatility. Forecasts start at the mentioned dates and use all available data from 1980:Q1. The density forecast variables are log predictive likelihood (LPL), CRPS, qwCRSP-right, and qwCRSP-left. The first line in each panel reports the density forecast metric of the benchmark model (Gauss BVAR). The figures for the MST-SV model are the relative improvements calculated as the differences of the different variables of the MST-SV model minus the ones of the benchmark models. Entries greater than 0 indicate that the MST-SV model is better. Remark: the forecasting exercise for eight quarters ahead ended at 2021:Q4 because our data ended at 2023:Q4.

5.7 Using other benchmarks for inflation expectations

In the previous section, we gauged the forecast performance of the MST-SV model against a competing standard Gaussian BVAR. We now set MST-SV model against more elaborate forecasts by practitioners, namely, the Survey of Professional Forecasters (SPF) and the Federal Reserve Bank of Cleveland model. Professional forecasters do not solely rely on models; they use their judgement extensively when forming their beliefs about the future. Hence, they are more able to readjust and update their predictions during challenging times with quick and drastic changes. For example, inflation had puzzling behavior during the pandemic. It declined at the beginning of the pandemic and afterwards surged quickly and drastically, driven by the global lockdowns and disrupted supply chains. Personal judgements and beliefs allow professional forecasters to have much more flexibility in their prediction scheme by distinguishing between the behavior of inflation in the short run and the long run. This is very challenging to do even by more advanced quantitative models based on rationality and backward-looking interconnexions between macroeconomic variables to extrapolate future movements. For instance, the Federal Reserve Bank of Cleveland estimates the expected rate of inflation over the next 30 years along with the inflation risk premium, the real risk premium, and the real interest rate. Their estimates are model-based and use Treasury yields, inflation data, inflation swaps, and survey-based measures of inflation expectations.[6]

We extrapolated the predicted inflation time series from the already forecasted price levels (in log), namely, the log CPI. In keeping with Federal Reserve practices, we constructed time series of inflation predictions at different horizons— one year, two years, and three years ahead—based on the median of the predicted log CPI. The one-year-ahead inflation forecasts are calculated as the average of predicted log CPI for t+1 to t+4, minus the average of realized log CPI over the previous four quarters. Two-years-ahead inflation forecasts are constructed as the difference between the average log CPI for t+5 to t+8 and t+1 to t+4. Lastly, three-years-ahead inflation forecasts are derived as the difference between the average log CPI for t+9 to t+12 and t+5 to t+8.

---

[6] The model is based on the work by Haubrich et al. (2012).

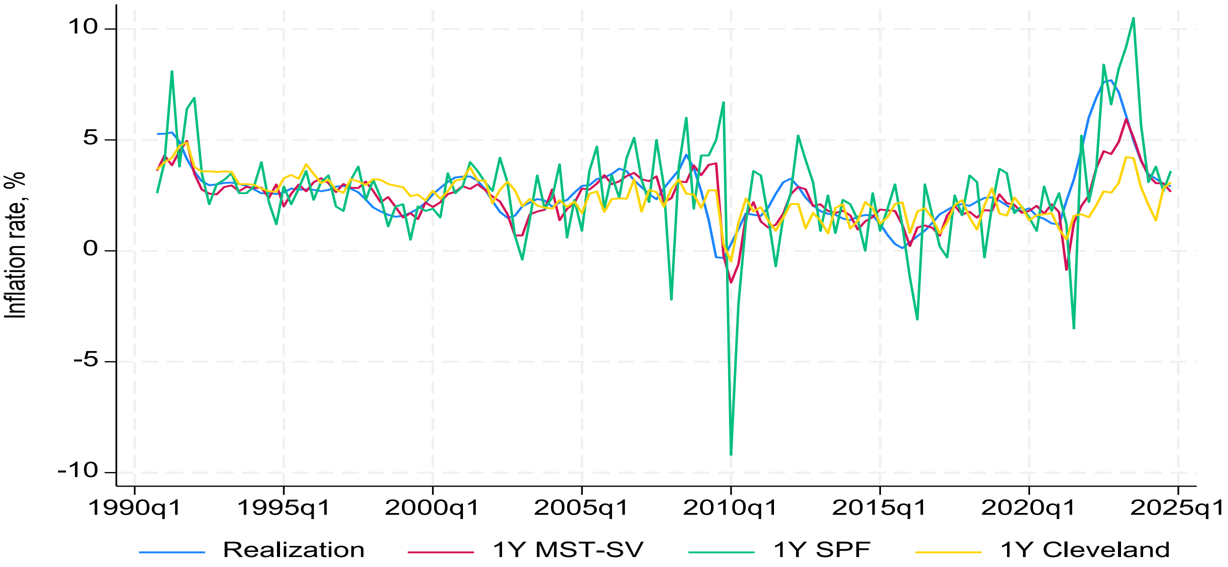

Figure 9*:* One-year-ahead forecasts of the inflation rate

Note: The figure plots one-year-ahead forecasts of the inflation rate made by the Survey of Professional Forecasters (the green line labeled “1Y SPF”), by the MST-SV model (red line labeled “1Y MST-SV”), and by the Fed of Cleveland (orange line labeled “1Y Cleveland”), together with the actual realization of the annual inflation rate (blue line labeled “Realization”). The period is from 1990:Q4 to 2024:Q4.

Figure 9 plots the realized annual inflation rate over the period 1990:Q4 to 2024:Q4, alongside the matched one-year-ahead inflation forecasts coming from the Survey of Professional Forecasters (SPF)[7], the MST-SV model, and the Fed of Cleveland. A visual inspection of Figure 9 reveals that the realized inflation is time varying and that these variations might be sudden and drastic, particularly during downturn episodes (the global financial crisis and COVID-19 pandemic). The forecasted inflation by SPF appears to be very volatile in time. Particularly, SPF forecasts have dramatic shifts at the onset of the financial crisis and COVID-19 pandemic, indicating that professional forecasters could quickly and radically adjust their beliefs about future patterns of inflation, based on new information and circumstances. The predictions by the MST-SV model and by the Fed of Cleveland seem less volatile and are adjusted smoothly. In addition, Figure 9

---

[7] We use the median responses coming from individual forecasts.

highlights interesting time-varying patterns in the forecast performance of the three alternatives. The three prediction models underpredicted the disinflation during the Great Recession of 2008–2009 and greatly underpredicted inflation during the pandemic era.

Figure 10 plots the realized annual inflation rate over the period 1992:Q4 to 2024:Q4, alongside the matched three-year-ahead inflation forecasts coming from our three competing alternatives. Surprisingly, the forecasts from our MST-SV model appear to be very volatile compared to the other two models, which lead to more stable inflation forecasting for three years ahead. It appears that the mean-reversion behavior of inflation at the medium horizons, due to monetary and fiscal policies, is better forecasted by the SPF and the Fed of Cleveland models than by the MST-SV model.

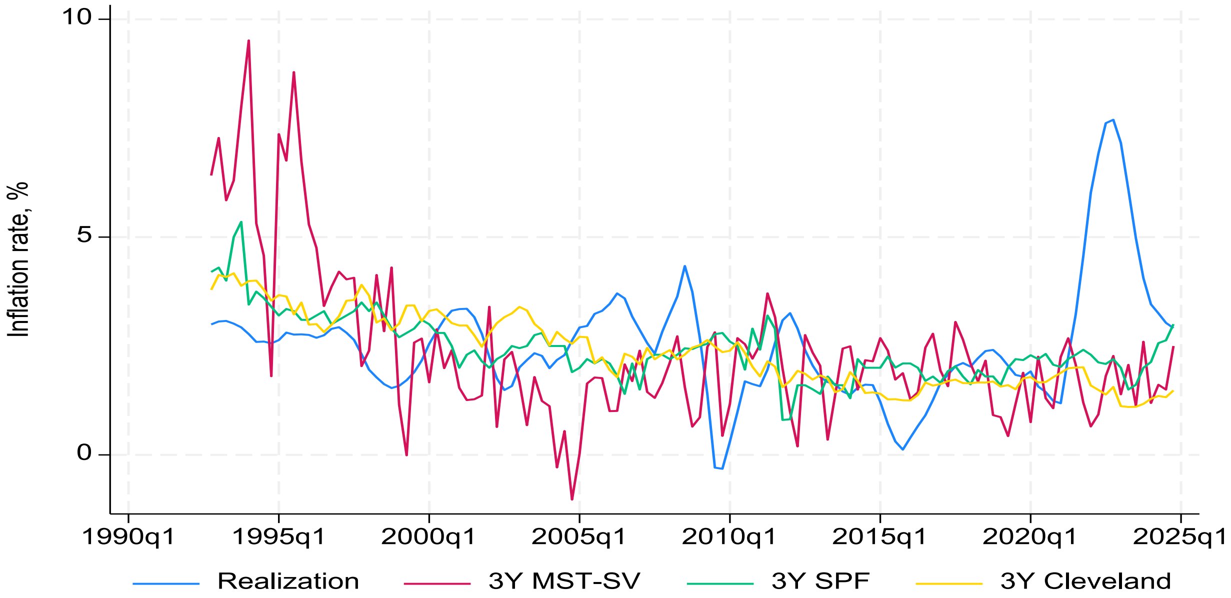

Figure 10: Three years ahead of the inflation rate

Note: The figure plots 12-quarters-ahead forecasts (3 years) of the inflation rate, made by the Survey of Professional Forecasters (green line labeled "3Y SPF"), by the MST-SV model (red line labeled "3Y MST-SV"), and by the Fed of Cleveland (orange line labeled "3Y Cleveland"), together with the actual realization of the annual inflation rate (blue line labeled "Realization"). The period is from 1992:Q4 to 2024:Q4.

To gain a deeper understanding of the forecasting performance of the three rival models, we use the fluctuation test developed by Giacomini and Rossi (2010). Surprisingly, the fluctuation test indicates that our MST-SV model outperforms the SPF in forecasting one-year-ahead inflation for a long period of time, ending at the onset of the financial crisis. Afterwards, the two models, SPF and MST-SV, have similar forecasting ability for the one-year-ahead inflation. In addition, the test indicates that the MST-SV and the Fed of Cleveland models have similar forecasting ability for the one-year-ahead inflation during all periods. For the three-years-ahead forecasts, the MST-SV model has a similar forecasting ability to the SPF, but they part ways as of the 2000s. Compared with the Fed of Cleveland model, our MST-SV was outperformed from 2015 to the end of the COVID-19 pandemic. The time-varying forecasting performance of the competing models suggests that combining these models could improve overall forecasting ability in the context of unstable economic environments.

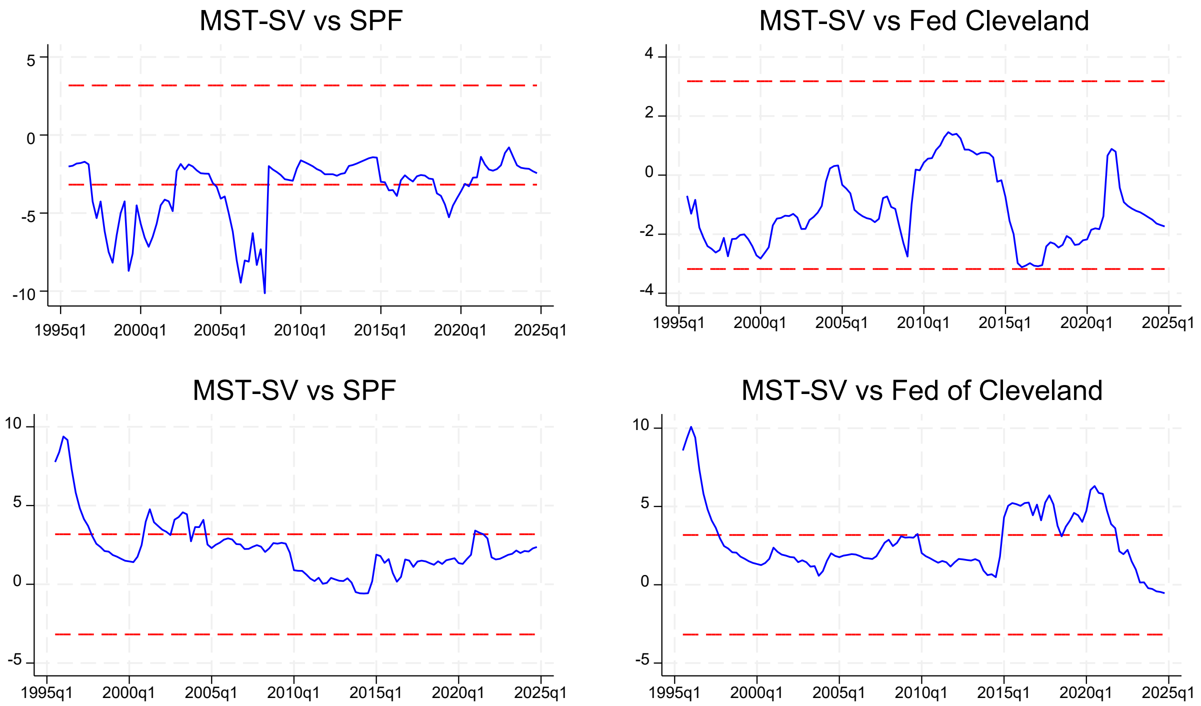

Figure 11: Fluctuation tests of Giacomini and Rossi (2010) for forecasts of quarterly CPI (in log-level) from the MST-SV model

Note: Test made using a recursive window. The forecasts start from 1989:Q4 to 2023:Q4 using all available datasets from 1980:Q1. The top panel is for one-year-ahead forecasts, and the lower panel is for the three-years-ahead forecasts. The rival models are SPF forecasting and the Fed of Cleveland (our competing models). Dashed lines are critical values, at the 95% confidence level. If a test statistic (solid lines) is inside the band lines at a time point, then the null hypothesis of equal predictive ability holds. When a test statistic (solid lines) is below the lowest band line, the MST-SV forecasts are significantly better than the competing forecasts.

6 Conclusion

The COVID-19 pandemic was an unprecedented shock to the global economy in recent history. It triggered enormous data movements and created outliers in the distribution of the main macroeconomic variables. These outliers could alter well-established macroeconomic dynamics where interconnexions were deeply undermined. The pandemic posed several challenges for the estimation and analysis of multivariate macroeconomic time-series. More specifically, estimated models may become unstable and generate unlikely forecasts once the extreme pandemic observations are included.

We tackle these methodological challenges using a flexible Bayesian VAR framework allowing for heavy tails and stochastic volatility for innovations, to appropriately handle these extreme observations, as has been suggested, among others, by Lenza and Primiceri (2022) and Schorfheide and Song (2024). We apply this novel model for a large US dataset starting from 1980:Q1 to 2023:Q4. We hence include in our estimations the full pandemic period, to gain a better understanding of its macroeconomic effects. We find empirical evidence that the COVID-19 pandemic created outliers for the main US macroeconomic variables, as the degrees of freedom of the Student's *t*-distributions of residuals became lower with the inclusion of post-pandemic observations.

The simultaneous shifts in the tail fatness of macrovariables could create macroeconomic tail risk, as suggested by Acemoglu et al (2017). The COVID-19 shock generated a long-lasting increase in stochastic volatility for the real GDP, price level (measured by the log CPI), and Fed rate. However, the increase in volatility for the labor market was transitory. Notably, the magnitude of the effects of COVID-19 on macroeconomic volatility varies from one period to another, and they vary cross-sectionally. Specifically for inflation, the impulse response functions reveal that the CPI (in log-level) responded differently to shocks in terms of real output, monetary policy, and tightness in the labor market during the pandemic era, as compared to pre-COVID periods. Hence, our analysis shows that macroeconomic transmission channels were highly altered by the pandemic shock. More importantly, forecasts accuracy metrics and the fluctuation tests of Giacomini and Rossi (2010) confirm the added forecastability in point and density forecasts of the price level and the inflation rate during the pandemic period and for multiple periods ahead by considering the tail fatness and time-varying volatility in macroeconomic disturbances, compared to standard Gaussian BVARs and other more elaborate competing forecasts.

References

Acemoglu D, Ozdaglar A, Tahbaz-Salehi A, 2017. Microeconomic origins of macroeconomic tail risks. *American Economic Review* 107, 1, 54-108.

Adrian T, Boyarchenko N, Giannone D, 2019. Vulnerable growth. *American Economic Review* 109, 4, 1263-1289.

Banbura M, Giannone D, Reichlin L, 2010. Large Bayesian vector auto regressions. *Journal of Applied Econometrics* 25, 1, 71–92.

Blanchard OJ, 2016. The Phillips curve: Back to the '60s? *American Economic Review* 106, 5, 31-34.

Bernanke B, Blanchard O, 2025. What Caused the US Pandemic-Era Inflation? *American Economic Journal: Macroeconomics* 17, 3, 1–35.

Bobeica E, Hartwig B, 2023. The COVID-19 shock and challenges for inflation modelling. *International Journal of Forecasting* 39, 1, 519-539.

Carriero A, Clark TE, Marcellino M, 2024a. Capturing macro-economic tail risks with Bayesian vector autoregressions. *Journal of Money, Credit and Banking* 56, 1099-1127. https://doi.org/10.1111/jmcb.13121.

Carriero A, Clark TE, Marcellino M, 2024b. Addressing COVID-19 outliers in BVARs with stochastic volatility. *Review of Economics and Statistics* 106, 5, 1403-1417.

Chan JCC, 2024. BVARs and Stochastic Volatility. In: M. Clements and A. Galvã (Eds.), *Handbook of Research Methods and Applications in Macroeconomic Forecasting*, 43-67, Edward Elgar Publishing.

Chan JCC, 2020. Large Bayesian VARs: A flexible Kronecker Error Covariance Structure. *Journal of Business and Economic Statistics* 38, 1, 68-79.

Chiu CW, Mumtaz H, Pinter G, 2017. Forecasting with VAR models: Fat tails and stochastic volatility. *International Journal of Forecasting* 33, 4, 1124-1143.

Cogley T, Sargent TJ, 2001. Evolving post-World War II US Inflation dynamics. *NBER Macroeconomics Annual* 16, 331-388.

Cogley T, Sargent TJ, 2005. Drift and volatilities: Monetary policy and output in Post WWII US. *Review of Economic Dynamics* 8, 275-308.

Del Negro M, Primiceri GE, 2015. Time-varying structural vector autoregressions and monetary policy: a corrigendum. *Review of Economic Studies* 82, 4, 1342-1345.

Doan T, Litterman R, Sims C, 1984. Forecasting and conditional projection using realistic prior distributions. Econometric reviews 3, 1, 1-100.

Fagiolo G, Napoletano M, Roventini A, 2008. Are output growth-rate distributions fat-tailed? Some evidence from OECD countries. *Journal of Applied Econometrics* 23, 639-669. https://doi.org/10.1002/jae.1003.

Giacomini R, Rossi B, 2010. Forecast comparisons in unstable environments. *Journal of Applied Economics* 25, 595-620. https://doi.org/10.1002/jae.1177.

Gneiting T, Ranjan R, 2011. Comparing density forecasts using threshold-and quantile-weighted scoring rules. *Journal of Business & Economic Statistics* 29, 3, 411-422. http://www.jstor.org /stable/23243806.

Hartwig B, 2024. Bayesian VARs and prior calibration in times of COVID-19. *Studies in Nonlinear Dynamics & Econometrics* 28, 1, 1-24. https://doi.org/10.1515/snde-2021-0108.

Haubrich J, Pennacchi G, Ritchken P, 2012. Inflation expectations, real rates, and risk premia: Evidence from inflation swaps. *The Review of Financial Studies* 25, 5, 1588–1629. https://doi.org/10.1093/rfs/hhs003.

Hazell J, Herreño J, Nakamura E, Steinsson J, 2022. The Slope of the Phillips Curve: Evidence from U.S. States, *The Quarterly Journal of Economics*, 137, I3, 1299–1344. https://doi.org/10.1093/qje/qjac010

Jarociński M, Maćkowiak B, 2017. Granger causal priority and choice of variables in vector autoregressions. *Review of Economics and Statistics* 99, 2, 319-329.

Jensen H, Petrella I, Hove Ravn S, Santoro E, 2020. Leverage and deepening business-cycle skewness. *American Economic Journal: Macroeconomics* 12, 1, 245-281.

Karlsson S, Mazur S, Nguyen H, 2023. Vector autoregression models with skewness and heavy tails. *Journal of Economic Dynamics and Control*, 146, 104580.

Lenza M, Primiceri GE, 2022. How to estimate a vector autoregression after March 2020. *Journal of Applied Econometrics* 37, 4, 688-699. https://doi.org/10.1002/jae.2895.

Litterman R, 1986. Forecasting with Bayesian vector autoregressions | Five years of experience. *Journal of Business and Economic Statistics* 4, 25-38.

Miranda-Agrippino S, Ricco G, 2021. The transmission of monetary policy shocks. *American Economic Journal: Macroeconomics* 13, 3, 74-107.

Primiceri GE, 2005. Time varying structural vector autoregressions and monetary policy. *Review of Economic Studies* 72, 3, 821-852.

Romer CD, Romer DH, 2004. A new measure of monetary shocks: Derivation and implications. *American Economic Review* 94, 4, 1055-1084.

Rossi B, 2021. Forecasting in the presence of instabilities: How we know whether models predict well and how to improve them. *Journal of Economic Literature* 59, 4, 1135-1190.

Rossi B, Inoue A, Wang Y, 2024. Has the Phillips curve flattened? French Stata Users' Group Meetings 2024 22, Stata Users Group.

Schorfheide F, Song D, 2024. Real-time forecasting with a (standard) mixed-frequency VAR during a pandemic. *International Journal of Central Banking* 20, 4, 275-320.

Sims CA, 1980. Macroeconomics and reality. *Econometrica* 48, 1, 1-48.

Sims CA, Uhlig H, 1991. Understanding unit rooters: A helicopter tour. *Econometrica* 59, 6, 1591-1600.

Sims CA, 1993. A nine-variable probabilistic macroeconomic forecasting model. In: Stock JH, Watson W (Eds.), Business cycles, indicators, and forecasting. University of Chicago press, 179-212.

Stock JH, Watson MW, 1999. Forecasting inflation. *Journal of Monetary Economics* 44, 2, 293-335.

Stock JH, Watson MW, 2020. Slack and cyclically sensitive inflation. *Journal of Money, Credit and Banking* 52, S2, 393-428.

Tenreyro S, Thwaites G, 2016. Pushing on a string: US Monetary policy is less powerful in recessions. *American Economic Journal: Macroeconomics* 8, 4, 43-74.