

Discussion Paper

Deutsche Bundesbank
No 52/2022

**Bayesian VARs and prior calibration
in times of COVID-19**

Benny Hartwig

Editorial Board:

Daniel Foos
Stephan Jank
Thomas Kick
Martin Kliem
Malte Knüppel
Christoph Memmel
Hannah Paule-Paludkiewicz

Deutsche Bundesbank, Wilhelm-Epstein-Straße 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-0

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

Internet <http://www.bundesbank.de>

Reproduction permitted only if source is stated.

ISBN 978-3-95729-931-4

ISSN 2749-2958

Non-technical summary

Research Question

The COVID-19 pandemic poses major challenges for many macroeconomic time series models for inference and forecasting purposes due to the extreme variation in some economic variables. Models based on Bayesian vector autoregression (BVAR) are also affected. A BVAR derives the relationships between economic variables from their historical properties and from a-priori distributions of the model parameters (priors). However, the extreme variation of some variables can severely distort these relationships and make the results of conventional BVARs appear implausible. This paper examines to which extent modified BVARs can deal with the extreme events caused by the pandemic.

Contribution

This paper shows that changing the error distribution makes the BVAR more robust to the extreme events caused by the pandemic: Instead of the usual normal distribution, a multivariate distribution with heavy tails or common time-varying volatility or a combination of both features can be used for this purpose. Since the common Minnesota prior is also very sensitive to extreme events, a modified calibration for the Minnesota prior is proposed as an alternative.

Results

When estimating BVARs using quarterly U.S. data, it turns out that the data favors the model with a combined error structure (heavy tails and common volatility). This model interprets the huge swings during the pandemic as a rare event rather than a persistent increase in macroeconomic volatility. Further, it becomes evident that the Minnesota prior is very sensitive to the extreme variation of some real variables. For instance, using this prior causes a structural break in the BVAR relation between employment and inflation. If the proposed modifications of the BVAR are taken into account, the accuracy of point forecasts for recent years tends to increase significantly.

Nichttechnische Zusammenfassung

Fragestellung

Die COVID-19 Pandemie stellt viele makroökonomische Zeitreihen-Modelle für Analyse- und Prognosezwecke aufgrund extremer Ausschläge einiger ökonomischer Variablen vor große Herausforderungen. Auch Modelle, die auf einer bayesianischen Vektorautoregression (BVAR) beruhen, sind davon betroffen. Ein BVAR leitet die Beziehungen zwischen ökonomischen Variablen aus deren historischen Verläufen sowie aus A-Priori-Verteilungen der Modellparameter (Priors) ab. Die extremen Ausschläge einiger Variablen können diese Beziehungen allerdings stark verzerren und lassen die Ergebnisse üblicher BVARs oft wenig plausibel erscheinen. Dieses Papier untersucht, inwiefern modifizierte BVARs mit den pandemiebedingten Extremereignissen umgehen können.

Beitrag

Es wird gezeigt, dass eine Änderung der Fehlerverteilung den BVAR robuster gegenüber den pandemiebedingten Extremereignissen werden lässt: Statt der üblichen Normalverteilung kann zu diesem Zweck eine multivariate Verteilung mit schweren Rändern oder gemeinsamer zeit-variierender Volatilität oder deren Kombination verwendet werden. Da zudem der gängige Minnesota-Prior sehr empfindlich gegenüber Extremereignissen ist, wird eine modifizierte Kalibrierung für den Minnesota-Prior als Alternative vorgeschlagen.

Ergebnisse

Bei der Schätzung von BVARs mit vierteljährlichen US-Daten stellt sich heraus, dass das Modell mit einer kombinierten Fehlerstruktur (schwere Ränder und gemeinsame Volatilität) von den Daten bevorzugt wird. Dieses Modell interpretiert die enormen Ausschläge während der Pandemie eher als ein seltenes Ereignis anstatt eines nachhaltigen Anstiegs der makroökonomischen Volatilität. Darüber hinaus zeigt sich, dass der Minnesota-Prior sehr empfindlich gegenüber den extremen Ausschlägen einiger realer Variablen ist. Die Verwendung dieses Priors ruft im BVAR beispielsweise einen Strukturbruch zwischen Beschäftigung und Inflation hervor. Wenn die vorgeschlagenen Modifikationen des BVARs berücksichtigt werden, erhöht sich die Treffgenauigkeit von Punktprognosen für die letzten Jahre oft stark.

Bayesian VARs and Prior Calibration in Times of COVID-19*

Benny Hartwig
Deutsche Bundesbank

Abstract

This paper investigates the ability of several generalized Bayesian vector autoregressions to cope with the extreme COVID-19 observations and discusses their impact on prior calibration for inference and forecasting purposes. It shows that the preferred model interprets the pandemic episode as a rare event rather than a persistent increase in macroeconomic volatility. For forecasting, the choice among outlier-robust error structures is less important, however, when a large cross-section of information is used. Besides the error structure, this paper shows that the standard Minnesota prior calibration is an important source of changing macroeconomic transmission channels during the pandemic, altering the predictability of real and nominal variables. To alleviate this sensitivity, an outlier-robust prior calibration is proposed.

Keywords: forecasting, multivariate t errors, common time-varying volatility, outlier-robust prior calibration.

JEL classification: C11, C51, C53.

*Contact address: Deutsche Bundesbank, DG Economics, Mainzer Landstraße 46, 60325 Frankfurt am Main, Germany. Phone: +49 69 9566-13582. E-mail: benny.hartwig@bundesbank.de. The author thanks an anonymous referee, Michael Binder, Elena Bobeica, Marek Jarocinski, Sören Karau, Michele Lenza as well as conference and seminar participants at the Bundesbank at the Bundesbank, 4th Annual Workshop on Financial Econometrics, 11th RCEA Money-Macro-Finance Conference and the RCC5 Brownbag for useful comments and suggestions. The views expressed in this paper are those of the author and do not necessarily coincide with the views of the Deutsche Bundesbank or the Eurosystem.

1 Introduction

The COVID-19 pandemic wreaked havoc in the global economy and triggered unprecedented movements in many key macroeconomic indicators. This unusual shock poses several challenges for the estimation and analysis of macroeconomic models. Specifically, Bayesian vector autoregressions (BVARs) may become unstable and generate implausible forecasts when recent observations are included in the estimation sample.¹

The pandemic, thus, calls for a careful treatment of the data in an econometric model and requires a view as to how it affects fundamental workings of the economy (Schorfheide and Song, 2021). Though the pandemic may have triggered forces that might alter macroeconomic interactions going forward, it is questionable whether a few extreme observations should dominate the inference procedure. For this reason, the literature proposes several solutions to make the BVAR more resilient to pandemic-related extreme observations. One prominent strategy is to relax the error structure in a BVAR.² This approach ranges from modelling COVID-19 observations explicitly with a common volatility model (Lenza and Primiceri, 2022), or implicitly by using a multivariate t distribution (Bobeica and Hartwig, 2023), or to variable-specific outliers in the volatility process (Carriero et al., 2022).

This paper addresses two important gaps in the literature. First, there is no clear guidance whether a multivariate heavy-tailed or common time-varying volatility error structure is better suited to account for the extreme variation in a BVAR and what works better in forecasting. Modelling the pandemic as a rare event or persistent increase in volatility is an important choice as it may have different implications for density forecasts. Second, the calibration of standard prior distributions such as the Minnesota prior has not received particular attention yet (except for fixing the calibration at pre-pandemic times). This is somewhat surprising as the COVID-19 observations may lead to a substantial re-calibration of the prior and, thus, may potentially alter inference and forecasting properties of a Bayesian VAR during the pandemic.

To fill these gaps, this paper assesses both error extensions individually and jointly by considering multivariate t errors and common stochastic volatility in the generalized BVAR framework of Chan (2020). The forecasts of these models are compared to the Lenza and Primiceri (2022) approach who also employ a common shock treatment. Modelling the COVID-19 observations as a common shock has some merits. First, the influence of extreme observations is downweighed homogeneously instead of heterogeneously across equations, implying that the COVID-19 observations have a more limited effect on all parameters in the VAR model. Second, inference with a common shock treatment is more simple and requires considerably less computation time than estimating the more flexible and realistic model of Carriero et al. (2022).³ Regarding the second gap, this paper investigates the sensitivity of the Minnesota prior during the pandemic. It argues that

¹See Lenza and Primiceri (2022), Carriero, Clark, Marcellino, and Mertens (2022), Schorfheide and Song (2021) on the U.S. and Bobeica and Hartwig (2023) on the euro area.

²Alternative approaches range from introducing dummy variables for the pandemic period (Carriero et al., 2022), including other exogenous variables related to the pandemic (Ng, 2021) to relaxing the whole structure by a non-parametric and non-linear mixed-frequency VAR approach (Huber, Koop, Onorante, Pfarrhofer, and Schreiner, 2023).

³The flexibility of their model comes at the cost of a peculiar variable ordering sensitivity which is especially relevant when volatility clusters idiosyncratically (Hartwig, 2020) and has the potential to alter conditional projections in an economically significant way (Arias, Rubio-Ramírez, and Shin, 2022).

variable-specific scale estimates used for calibrating the prior are not robust to extreme observations and therefore become a source of changing estimated transmission channels. To alleviate this sensitivity, this paper proposes and discusses several outlier-robust calibration strategies for the Minnesota prior. The importance of these considerations is illustrated in an empirical application with U.S. quarterly data featuring both a medium VAR with six key macroeconomic indicators similar to [Lenza and Primiceri \(2022\)](#) and a large VAR with 20 variables.

This paper makes three contributions to the literature. First, model diagnostics indicate that the data prefers to interpret the COVID-19 shock as a rare event rather than a persistent increase in macroeconomic volatility. Moreover, the persistence of macroeconomic volatility in BVARs featuring only time-varying volatility is inversely related to the amount of cross-sectional information as variation in many variables quickly normalized. Consequently, predictive intervals in medium BVARs may become very imprecise during the pandemic, whereas those of large BVARs are not particularly inflated for a long time. This complements the evidence in [Carriero et al. \(2022\)](#) by the fact that a common volatility structure is not necessarily plagued by excessively wide predictive intervals as compared to a variable-specific volatility structure without outlier correction. Further, the forecast exercises suggest that the robust BVARs coupled with a large cross-sectional information set is the group of superior forecasting models. Among the robust BVARs, however, there is no clear superior forecasting approach for all variables, horizons and prior combinations as the predictive accuracy is broadly similar. Thus, the off-the-shelf robust BVARs of [Chan \(2020\)](#) can be readily used during the pandemic period. Nevertheless, the large BVAR of [Lenza and Primiceri \(2022\)](#) produced the most accurate forecasts overall.

The second contribution of this paper is to document that a mechanical update of the Minnesota prior is another important source of parameter instability during the pandemic. The updated prior distribution differs considerably from the pre-pandemic one as it sets a very tight prior for VAR coefficients of real variables. Relative to the pre-pandemic prior, this updated prior is associated with a decisive loss of the in-sample fit, change in transmission channels and trade-off between the predictability of real variables and inflation. Further, the muted internal propagation improves predictive accuracy of real variables in the beginning of the pandemic, however, this goes at the expense of getting inaccurate predictions for inflation. In fact, this strongly revised prior prevents the large VAR model from capturing the mounting inflationary pressure in 2021.

Third, adopting an extreme observation-robust prior calibration strategy improves the model fit substantially and delivers a comparable pre-pandemic prior distribution. Moreover, the robust calibration strategies offer three key advantages over fixing the prior distribution or a manual fine-tuning of the calibration sample: they are easy to implement, do not sacrifice in-sample fit for robustness and allow for an orderly revision of the prior distribution by robustly processing all available information. Based on theoretical considerations and simplicity, the MAD(q Reg) prior calibration based on the scaled median absolute deviation and median AR(p) residual is proposed as the COVID-19 robust alternative.

The rest of this paper proceeds as follows. Section 2 presents Bayesian VAR models to cope with COVID-19 and discusses prior calibration. Section 3 conducts an empirical analysis with these models and priors using U.S. data. Section 4 concludes this paper.

2 Methodology

2.1 Bayesian VARs to cope with the COVID-19 pandemic

This section presents the Bayesian VAR model with a generalized error structure of [Chan \(2020\)](#) and uses it as a framework to assess whether a heavy-tailed or heteroskedastic error structure should be preferred for modelling the COVID-19 observations.

Let y_t be a $n \times 1$ vector of variables that is observed over the periods $t = 1, \dots, T$. Consider the following generic VAR(p) model:

$$y_t = a_0 + A_1 y_{t-1} + \dots + A_p y_{t-p} + \epsilon_t,$$

where a_0 is an $n \times 1$ vector of intercepts and A_1, \dots, A_p are $n \times n$ coefficient matrices. Let $x'_t = (1, y'_{t-1}, \dots, y'_{t-p})$ be a $1 \times k$ vector of an intercept and lags with $k = 1 + np$ and $A = (a_0, A_1, \dots, A_p)'$ is a $k \times n$. Rewrite the VAR in more compact form as $y'_t = x'_t A + \epsilon_t$, and stack the observations over $t = 1, \dots, T$, which yields

$$Y = XA + E, \tag{1}$$

where Y , X , and E are of dimensions $T \times n$, $T \times k$, and $T \times n$, respectively. In a standard VAR, the innovations $\epsilon_1, \dots, \epsilon_T$ are assumed to be independent and identically distributed (i.i.d.) as $N(0, \Sigma)$. More compactly, rewrite $E \sim MN(0, \Sigma \otimes I_T)$, where Σ is an $n \times n$ covariance matrix, I_T is the identity matrix of dimension T , MN denotes the matrix-variate normal distribution, \otimes is the Kronecker product.

The Kronecker structure allows for a more general covariance structure:

$$E \sim MN(0, \Sigma \otimes \Omega), \tag{2}$$

where Ω is a $T \times T$ covariance matrix. Intuitively, the cross-sectional and serial covariance structures of Y are separately modelled by specifying Σ and Ω , respectively. By choosing a suitable serial covariance structure Ω , the model in (1)–(2) includes a variety of flexible error structures such as heavy-tailed and heteroskedastic innovations.

The main difference between these error structures lies in their treatment of large shocks. Under a heavy-tailed distribution, large shocks receive more probability mass in general, while with heteroskedastic errors large shocks are captured by allowing volatility to change over time. Intuitively, these error structures may limit the effect of extremely large innovations on the parameter estimates because they downscale their contribution through the (implied) serial covariance Ω , see [Section 2.3](#) for details. This idea of down-weighting observations to stabilize VAR parameter estimates in the context of the COVID-19 pandemic was first proposed by [Lenza and Primiceri \(2022\)](#).

From the class of heavy-tailed distributions, this paper considers the multivariate t distribution as it can accommodate substantially thicker tails than the normal distribution. Specifically, ϵ_t is assumed to be independently multivariate t -distributed with mean vector 0, scale matrix Σ and degree of freedom parameter ν , or more compactly, $\epsilon_t \sim t(0, \Sigma, \nu)$. As explained in [Chan \(2020\)](#), many non-Gaussian distributions can be used in this framework as they can be written as a scale mixture of Gaussian distributions.⁴

⁴The outlier-adjusted error distribution of [Stock and Watson \(2016\)](#) or, as discussed in [Chan \(2020\)](#), the multivariate Laplace distribution of [Eltoft, Kim, and Lee \(2006\)](#) may also be used in this context.

The t distribution can be modelled by letting $\Omega = \text{diag}(\lambda_1, \dots, \lambda_T)$, where each λ_t follows independently an inverse-gamma distribution $(\lambda_t|\nu) \sim IG(\nu/2, \nu/2)$, see [Geweke \(1993\)](#). In other words, the scale mixture Gaussian representation of the t distribution reveals that the innovations ϵ_t conditional on λ_t exhibit common but independently distributed shifts of the cross-sectional covariance matrix over time. Note [Bobeica and Hartwig \(2023\)](#) also consider multivariate t errors to model the pandemic observations.

As heteroskedastic error structure, the common drift in stochastic volatility model in the spirit of [Carriero, Clark, and Marcellino \(2016\)](#) is considered with $\epsilon_t \sim N(0, \Sigma \cdot \exp(h_t))$ following [Chan \(2020\)](#). The log variance follows a stationary autoregressive process of order one:

$$h_t = \rho h_{t-1} + \epsilon_t^h, \quad (3)$$

with $\epsilon_t^h \sim N(0, \sigma_h^2)$, $|\rho| < 1$ and $\Omega = \text{diag}(\exp(h_1), \exp(h_2), \dots, \exp(h_T))$.

Note this covariance structure shares some similarities with the [Lenza and Primiceri \(2022\)](#) approach. Specifically, they assume that the COVID-19 observations scale the cross-sectional covariance matrix persistently by a common factor that decays over time. For the pre-pandemic sample, they assume a constant covariance matrix. In contrast, the model with common stochastic volatility allows volatility to change over the whole sample, not just in response to the pandemic.

The third error structure considered here is a combination of common drift in volatility and multivariate t innovations as in [Chan \(2020\)](#). Specifically, $\epsilon_t \sim N(0, \Sigma \cdot \lambda_t \cdot \exp(h_t))$, where $\lambda_t \sim IG(\nu/2, \nu/2)$ and h_t follows an AR(1) process. This model allows for richer (implied) volatility dynamics as opposed to their single ingredients. It can capture both persistent but also abnormally large changes in volatility, thereby allowing the data to decide whether the COVID-19 observations have transient or persistent effects on volatility.⁵ [Table 1](#) summarizes the list of competing Bayesian VARs with a generalized error structure.

Table 1: List of competing models

Model	Description
BVAR- \mathcal{N}	BVAR with i.i.d. Gaussian innovations
BVAR- t	BVAR with i.i.d. t innovations
BVAR- \mathcal{N} -CSV	BVAR with Gaussian innovations and common stochastic volatility
BVAR- t -CSV	BVAR with t innovations and common stochastic volatility

Apart from the COVID-19 related macroeconomic disruptions, there is a vast literature documenting a departure from classical Gaussian errors in a wide range of macroeconomic models. In the context of VARs, [Christiano \(2007\)](#) presents evidence that even before the Global Financial Crisis VAR residuals are not normally distributed. Moreover, [Chiu, Mumtaz, and Pinter \(2017\)](#), [Clark and Ravazzolo \(2015\)](#) and [Chan \(2020\)](#) document that the data favors stochastic volatility and t -distributed errors over Gaussian errors in a VAR model and that these features improve forecast accuracy. In addition, [Cúrdia, Del Negro,](#)

⁵The idea to combine these error structures to discriminate between high-frequency and low-frequency volatility was first put forward by [Jacquier, Polson, and Rossi \(2004\)](#).

and Greenwald (2014) and Chib and Ramamurthy (2014) provide evidence that structural shocks in DSGE models exhibit heavy tails.⁶ Antolín-Díaz, Drechsel, and Petrella (2020) show that accounting for both shifts in variances and large shocks improves nowcast accuracy of a dynamic factor model relative to a model with classical Gaussian errors.

Though stochastic volatility is often found to be a more important feature than t -distributed errors for macroeconomic time series, several studies document that inference and forecast accuracy may hinge on allowing for heavy-tailed errors.⁷ For instance, ignoring fat tails may lead to misleading inference about the historical evolution of volatility, especially during the Global Financial Crisis, which is interpreted as a rare event rather than a persistent increase in macroeconomic volatility when allowing for both features.

The Bayesian VAR models that combine stochastic volatility with t -distributed errors of Clark and Ravazzolo (2015) and Chiu et al. (2017) or with outlier-adjusted errors of Carriero et al. (2022) differ from the models discussed above by their assumption that time-varying volatility and heavy-tailedness is variable-specific rather than common across all variables. Though an individual error distribution may be more realistic and less restrictive, the modelling approach pursued by these papers is prone to be sensitive to the variable ordering due to the triangular factorization of the time-varying covariance matrix in the spirit of Cogley and Sargent (2005). Specifically, this variable ordering sensitivity may give rise to ambiguous conclusions and is especially relevant when volatility clusters idiosyncratically, see Hartwig (2020). It can also affect conditional projections in an economically significant way, see Arias et al. (2022).

2.2 Prior calibration

Besides the error structure in a Bayesian VAR, another source of parameter instability may be due to an unusually strong and potentially unintended re-calibration of the prior distribution during the pandemic. Specifically, many prior distributions rely on simple variable-specific volatility estimates based on the full sample and as such are sensitive to extreme observations as well. The Minnesota prior is just one example of such a commonly used prior distribution in the context of Bayesian VARs. This section discusses the calibration sensitivity in detail and proposes outlier-robust calibration strategies for the Minnesota prior, which can also be used for calibrating other scale-dependent prior distributions.⁸

For all considered Bayesian VAR models, this paper assumes a normal-inverse-Wishart prior for the VAR coefficients A and cross-sectional covariance matrix Σ , i.e. $\Sigma \sim IW(S_0, \nu_0)$, and $(A|\Sigma) \sim MN(A_0, \Sigma \otimes V_a)$. As a naïve benchmark prior, a weakly informative but sample-independent calibration is used to specify the normal-inverse-Wishart prior. The hyperparameters of the inverse-Wishart prior are set to $\nu_0 = 3 + n$ and $S_0 = I_n$. The prior mean A_0 is centered at zero and the covariance matrix, V_a^W , is assumed to be

⁶Cúrdia et al. (2014) consider a model with both stochastic volatility and t -distributed errors, while Chib and Ramamurthy (2014) focus on t -distributed errors only.

⁷See Cúrdia et al. (2014), Chiu et al. (2017), Chan (2020) and Antolín-Díaz et al. (2020).

⁸Further examples of scale-dependent priors are the sum of coefficient (co-integration) and sum of initial conditions prior (Sims and Zha, 1998), steady-state prior (Villani, 2009) and long-run relations prior (Giannone, Lenza, and Primiceri, 2019).

diagonal with i -th diagonal element $v_{a,ii}$ set as:

$$v_{A,ii}^W = \begin{cases} 1 & \text{for a coefficient associated to a lag of variable } r \\ \kappa_2^2 & \text{for an intercept} \end{cases}, \quad (4)$$

where $\kappa_2 = 10$ controls the overall strength of shrinkage for the intercepts.

To maintain the Kronecker structure of the prior, this paper adopts a Minnesota prior without cross-variable shrinkage, see [Carriero et al. \(2016\)](#) and [Chan \(2020\)](#).⁹ The Minnesota prior is implemented as a normal-inverse-Wishart prior using the same notation as above. For the inverse-Wishart prior, the same sample-independent calibration is used to isolate the effects of a sample-dependent calibration to the prior covariance matrix V_a^{MN} of the VAR coefficients only.¹⁰ V_a^{MN} is assumed to be diagonal with i -th diagonal element $v_{a,ii}$ set as:

$$v_{A,ii}^{MN} = \begin{cases} \frac{\kappa_{1,m}^2}{l^2 \hat{s}_{r,m}^2} & \text{for a coefficient associated to lag } l \text{ of variable } r \\ \kappa_2^2 & \text{for an intercept} \end{cases}, \quad (5)$$

where $\hat{s}_{r,m}$ is the scale estimate of variable r and calibration strategy m , specified below, and $\kappa_{1,m}$ controls the degree of shrinkage of the VAR coefficients. For comparability, $\kappa_{1,m}$ is calibrated for each method m to yield the maximum marginal likelihood of the BVAR- \mathcal{N} model at pre-pandemic times, see Section 3.2.

In practice, $\hat{s}_{r,m}$ is calibrated to be the variable-specific sample volatility of the residual from an AR(p) model with Gaussian and homoskedastic errors, see [Litterman \(1986\)](#). Particularly, the root mean squared deviation (RMSD) is used as the sample-specific volatility estimate of variable r :

$$RMSD(x_r) = \frac{1}{T} \sqrt{\sum_{t=1}^T (x_{r,t} - \text{mean}(x_r))^2}, \quad (6)$$

where x_r is the AR(p) residual of variable r in the standard specification.¹¹

There are two main reasons why the Minnesota prior is based on variable-specific volatility estimates. First, the prior is not scale-invariant, i.e. it matters whether the data is entered in decimals or percentage points. Second, a variable-specific volatility estimate is necessary to express the degree of shrinkage relative to the variability of the underlying stochastic process which is similar in spirit to a ridge regression.

By construction, however, this classical prior calibration strategy is sensitive to extreme observations. Because of this, volatility estimates of some variables may be strongly inflated once the pandemic observations are included in the calibration sample, leading to a tighter prior for the affected variables as compared to the pre-pandemic calibration. This mechanical revision of the prior distribution of the VAR coefficients may be unattractive for several reasons. First, an extremely tight prior for some variables may alter the

⁹For more details on this specific Minnesota prior, see Section 3.2.1 in [Karlsson \(2013\)](#).

¹⁰Alternatively, the scale S_0 may be calibrated using sample-dependent information as $S_0 = (\nu_0 - n - 1) \cdot \text{diag}(\hat{s}_{1,m}^2, \dots, \hat{s}_{n,m}^2)$, see Section 3.2.1. [Karlsson \(2013\)](#).

¹¹For ease of exposition, the degree of freedom correction is omitted here.

dynamic propagation of shocks by muting their transmission and breaking links between variables. Second, it is questionable whether a few extreme observations should receive so much weight in revising the prior distribution of the VAR dynamics. Third, it is not clear if and how quickly macroeconomic transmission channels change in response to the pandemic. Therefore, it might be more attractive to adopt a less sensitive prior calibration to not a-priori mute the information in the data through an excessively tight prior.

The unique character of the COVID-19 crisis allows the researcher to easily pin down such a prior by cutting the calibration sample at pre-pandemic times, see, for instance, [Schorfheide and Song \(2021\)](#) and [Lenza and Primiceri \(2022\)](#). However, a drawback of this approach is that it discards all future and potentially useful information from the new normal once macroeconomic variation stabilizes. It also raises the question as to how new observations should be treated when specifying a prior distribution for empirical applications going forward. Moreover, an ad-hoc fine-tuning of the prior calibration sample might also limit the comparability across empirical studies as the appropriate timing to include new observations may vary across applications, countries and the judgement of researchers.

For this reason, this paper searches for outlier-robust calibration strategies that can process all information and yield a prior distribution that is comparable to the standard method in absence of extreme observations. To meet these requirements, the search is guided by looking for robust estimators that yield asymptotically the same point estimate as a mean regression when the data is normally distributed. There are two key elements that may lead to a revision of the variable-specific volatility estimate: 1) the volatility estimator itself and 2) the input series x_r used for the volatility estimator. To investigate the importance of these sources, this paper considers three robust volatility estimators and combines them with two alternative input series to calibrate the Minnesota prior.

The scaled median absolute deviation (MAD) is considered as a first alternative to the RMSD because it is robust to extreme observations, simple to implement and yields asymptotically the same point estimate of volatility when the data is normally distributed, see [Rousseeuw and Croux \(1993\)](#):

$$MAD(x_r) = b \cdot \text{median}(|(x_r - \text{median}(x_r))|), \quad (7)$$

where $b = 1.4826$ is a scaling parameter.

However, a drawback of the MAD is that it takes a symmetric view point on dispersion and has a rather low Gaussian efficiency, see [Rousseeuw and Croux \(1993\)](#). As an alternative, the authors propose the S_n and Q_n statistic as a more efficient and general alternative to the MAD. Specifically, these volatility estimators remain valid even when the input series x_r follows an asymmetric or a heavy-tailed distribution. Therefore, these statistics are considered as further alternatives to the RMSD.

The S_n statistic measures dispersion as the typical distance between observations as opposed to the typical distance from the central value like the MAD.¹²

$$S_n(x_r) = c \cdot \text{median}\{\text{median}|x_{r,t} - x_{r,s}|\}, \quad s = 1, \dots, T, \quad (8)$$

where $c = 1.1926$ is a scaling parameter.

¹²The lack of location dependence makes the S_n statistic valid for asymmetric distributions.

The Q_n statistic is defined as the k -th order statistic of the $\binom{T}{2}$ interpoint distances:

$$Q_n(x_r) = d \cdot \{|x_{r,t} - x_{r,s}|; t < s\}_{(k)}, \quad (9)$$

where $d = 2.219$ is a scaling parameter and $k = \binom{(T/2)+1}{2}$.¹³

The second ingredient to make the prior calibration robust to extreme observations is to use a less sensitive input series than the residual of a standard Gaussian AR(p) model (o Reg). Here, the first difference of the time series (FD) and the residual of a median AR(p) model (q Reg) are considered as an alternative. The median AR(p) model is estimated using the quantile regression framework of [Koenker and Bassett \(1978\)](#).

The advantage of using a first-differenced time series over an AR(p) residual is that the former remains unchanged, while the latter may be revised due to updated model parameters. This invariance property makes the first-differenced time series an ideal input series to assess the resilience of alternative volatility estimators during the pandemic. However, a drawback is that the associated volatility estimate is generally larger but not proportionally larger across the n equations because of different lag properties of the individual variables. Consequently, the prior distribution based on this input series might be somewhat different and the amount of shrinkage may be suboptimal for a VAR model.¹⁴

Relative to the first-differenced time series, the residual of a median AR(p) and mean AR(p) regression tend to be more similar as they take the lag structure into account. For Gaussian data, both estimation methods produce asymptotically the same coefficient estimate as the (conditional) median and mean of a normally distributed variable is identical. However, the median regression has the advantage of being more robust and efficient than the mean regression when the error distribution is heavy-tailed, see [Koenker and Bassett \(1978\)](#). Thus, the median AR(p) residual is expected to be less sensitive during the pandemic. Table 2 summarizes all considered prior specifications.

Table 2: List of prior specifications

Prior	Description
Weak	Weakly informative prior (naïve benchmark)
RMSD (o Reg)	Minnesota prior (standard AR(p) and scale, RMSD)
MAD (o Reg)	Minnesota prior (standard AR(p) but robust scale, MAD)
S_n (o Reg)	Minnesota prior (standard AR(p) but robust scale, S_n)
Q_n (o Reg)	Minnesota prior (standard AR(p) but robust scale, Q_n)
RMSD (FD)	Minnesota prior (first difference and scale, RMSD)
MAD (FD)	Minnesota prior (first difference and robust scale, MAD)
S_n (FD)	Minnesota prior (first difference and robust scale, S_n)
Q_n (FD)	Minnesota prior (first difference and robust scale, Q_n)
RMSD (q Reg)	Minnesota prior (robust- q AR(p) and scale, RMSD)
MAD (q Reg)	Minnesota prior (robust- q AR(p) and scale, MAD)
S_n (q Reg)	Minnesota prior (robust- q AR(p) and scale, S_n)
Q_n (q Reg)	Minnesota prior (robust- q AR(p) and scale, Q_n)

¹³The Q_n statistic has the highest Gaussian efficiency among the MAD and S_n statistic.

¹⁴Note $\kappa_{1,m}$ is common across the n equations and cannot act as a degree of freedom.

To complete the prior specification, this paper follows [Chan \(2020\)](#) and assumes a uniform prior on $(2, 100)$ for the degree of freedom parameter, i.e. $\nu \sim U(2, 100)$. This prior implies that degrees of freedom may become sufficiently large to approximate Gaussian errors. For stochastic volatility, independent priors for σ_h^2 and ρ : $\sigma_h^2 \sim IG(\nu_{h0}, S_{h0})$ and $\rho \sim N(\rho_0, V_\rho)\mathbf{1}(|\rho| < 1)$ are assumed, where $\nu_{h0} = 5$, $S_{h0} = 0.04$, $\rho_0 = 0.9$, and $V_\rho = 0.2^2$. These values imply that the prior mean of σ_h^2 is 0.1 and ρ is centered at 0.9. Further, an independent prior for the parameter blocks (A, Σ) and Ω is assumed, i.e., $p(A, \Sigma, \Omega) = p(A, \Sigma)p(\Omega)$.

2.3 Bayesian estimation

The Bayesian VAR with covariance structure (2) can be easily estimated under a normal-inverse-Wishart prior for (A, Σ) and an independent prior for the parameter blocks (A, Σ) and Ω , see [Chan \(2020\)](#).

Posterior draws can be obtained by sequentially sampling from 1) $p(A, \Sigma|Y, \Omega)$ and 2) $p(\Omega|Y, A, \Sigma)$. Depending on the covariance structure Ω , additional blocks might be needed to sample some extra hierarchical parameters. In particular, the BVAR with t -distributed errors requires additional sampling of the degree of freedom parameter ν , while (ρ, σ_h^2) need to be sampled for the BVAR with common stochastic volatility. Following [Chan \(2020\)](#), these parameters are fitted using univariate time series models.

To understand how (extreme) observations are treated in this model, it is useful to investigate the conditional posterior $p(A, \Sigma|Y, \Omega)$. As the prior $p(A, \Sigma)$ is natural conjugate, the conditional posterior $p(A, \Sigma|Y, \Omega)$ is still normal-inverse-Wishart distributed:

$$p(A, \Sigma|Y, \Omega) \sim MNIW(\hat{A}, \Sigma \otimes K_A^{-1}, \hat{S}, v_0 + T),$$

where

$$\begin{aligned} K_A &= V_A^{-1} + X'\Omega^{-1}X, & \hat{A} &= K_A^{-1}(V_A^{-1}A_0 + X'\Omega^{-1}Y), \\ \hat{S} &= S_0 + A_0'V_A^{-1}A_0 + Y'\Omega^{-1}Y - \hat{A}'K_A^{-1}\hat{A}. \end{aligned}$$

Hence, $(A, \Sigma|Y, \Omega)$ can be sampled in two steps. First, sample Σ marginally from $(\Sigma|Y, \Omega) \sim IW(\hat{S}, v_0 + T)$. Then, given the Σ drawn, sample A from

$$(A|Y, \Sigma, \Omega) \sim MN(\hat{A}, \Sigma \otimes K_A^{-1}).$$

Note under a noninformative prior for the VAR coefficients, i.e. when the inverse of V_A goes to zero, the conditional posterior mean of A converges to the generalized least squared estimator for some conditionally known Ω (weighting matrix).¹⁵ Thus, the data (Y, X) is not equally informative for the parameters (A, Σ) , but their contribution is weighted by Ω . Under t -distributed errors and common stochastic volatility, Ω is a diagonal matrix with generic elements $\omega_t^2, t = 1, \dots, T$. In period t , the observations (y_t, x_t) are weighted by the (implicit) volatility ω_t^{-1} . Specifically, when ω_t is large, then (y_t, x_t) is less informative for (A, Σ) . Intuitively, the common volatility estimate ω_t becomes large if some of the forecast errors ϵ_t are unusually large, either driven by some variable r or a set of variables.

¹⁵When $\Omega = I_T$, then the posterior mean of A converges to the least squares estimator.

3 Empirical Application

3.1 Data

This section assesses the impact of the COVID-19 observations on these more flexible Bayesian VARs and various prior calibration strategies by considering a medium and large VAR model for inference and forecasting at quarterly frequency, see Table 3.¹⁶ The medium VAR consists of six core U.S. macroeconomic variables similar to [Lenza and Primiceri \(2022\)](#).¹⁷ The large VAR adds 14 additional variables typically used in forecasting exercises such as [Bańbura, Giannone, and Reichlin \(2010\)](#), [Carriero, Clark, and Marcellino \(2015\)](#) and [Chan \(2020\)](#). Both models are estimated over an expanding window from 1988:Q4 to 2019:Q4 until 2022:Q1. As in [Lenza and Primiceri \(2022\)](#), the sample is not extended before 1988:Q4 as inflation has been reacting less to real activity since the 1990s ([Del Negro, Lenza, Primiceri, and Tambalotti, 2020](#)).

Table 3: Quarterly dataset from FRED-QD ([McCracken and Ng, 2020](#))

Variable	Abbreviation	Mnemonic	Transformation	Medium	Large
Employment (nonfarm)	EMP	PAYEMS	log	x	x
Unemployment rate	UR	UNRATE	raw	x	x
Hours worked (nonfarm)	HW	HOANBS	log		x
Hourly earnings (goods)	HE	CES0600000008	log		x
Consumption (PCE, real)	CON	PCECC96	log	x	x
Gross domestic product (real)	GDP	GDPC1	log	x	x
Industrial production	IP	INDPRO	log		x
Capacity util. (manuf.)	CU	CUMFNS	log		x
Housing starts (total)	HOUS	HOUST	log		x
Consumer price index (all items)	CPI	CPIAUCSL	log	x	x
PCE, price index	PCE	PCECTPI	log		x
PCE, core price index	PCEc	PCEPILFE	log	x	x
Real crude oil price	OIL	OILPRICEx	log		x
PPI (finished goods)	PPI	WPSFD49207	log		x
1-year government bond yield	T1y	GS1	raw		x
10-year government bond yield	T10y	GS10	raw		x
BAA-10-year govt. spread	BAA	BAA10YM	raw		x
S&P500 stock price index	SP500	S&P 500	log		x
Real M2 money stock	M2	M2REAL	log		x
Exchange rate (trade weighted)	EER	TWEXMMTH	log		x

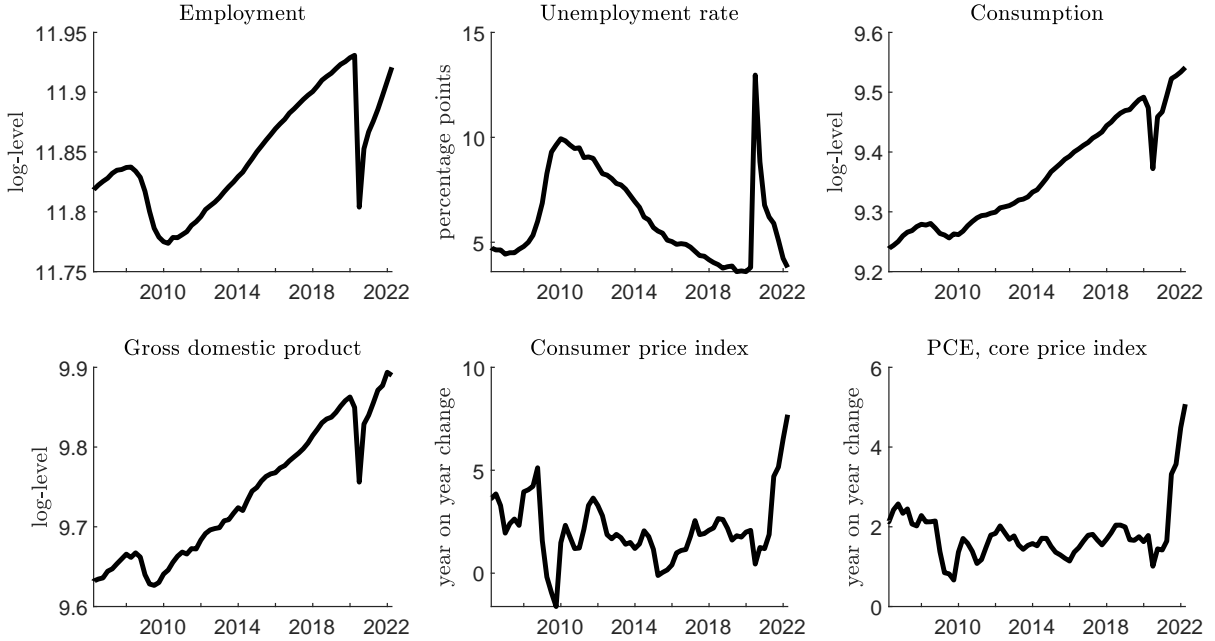
Figure 1 shows an excerpt of the data for the variables in the medium VAR from 2006:Q1 until 2022:Q1. From March 2020 onwards, COVID-19 started to spread through the U.S. and led to a widespread shutdown of the U.S. economy. This triggered an unparalleled decline in labor market indicators like employment but also in consumption and output in the second quarter of 2020. With lockdown measures easing over the summer, the U.S. economy rebounded strongly during the third quarter but lost momentum due to a second wave of infections in the fourth quarter. Surprisingly, CPI and PCE core inflation did not particularly react to the economic turmoil at first. In 2021, the

¹⁶The quarterly frequency is adopted to provide additional evidence on model (in)stability.

¹⁷See June 2020 WP. Industrial production is replaced by GDP as it is a broader measure of activity.

economy recovered steadily but inflation started to accelerate substantially which took many economists by surprise. In fact, inflation turned out to be more persistent than initially expected as supply could not catch up with demand due to the strong re-opening of the economy and sustained supply bottlenecks in the global economy. This raises the question whether standard econometric models are still useful in tracking and forecasting these developments and what specifications work best.

Figure 1: Time-series plots for key U.S. macroeconomic variables



3.2 (Re)-calibration of the Minnesota prior

A careful calibration of the Minnesota prior – especially in turbulent times – is one important aspect for modelling and forecasting macroeconomic time series with a Bayesian VAR. To study what the proposed calibration strategies entail for the Minnesota prior, the variable-specific volatility estimates of various estimation strategies is compared first. Then, the implications for the optimal degree of parameter shrinkage as well as the sensitivity during the COVID-19 pandemic is assessed.

Table 4 (a) reports variable-specific volatility estimates for various calibration strategies of the Minnesota prior at pre-pandemic times (2019:Q4) in absolute and relative terms to $\text{RMSD}(o\text{Reg})$. Three facts stand out. First, Panel (a) shows that the volatility estimates are generally larger but not proportionally larger when the first difference is used as an input series due to different serial correlation patterns, see column $\text{RMSD}(\text{FD})$. For instance, the volatility estimate of the unemployment rate is about 1.4 times higher than the $\text{RMSD}(o\text{Reg})$, while this factor is about 4.1 for the PCE core price index.

Second, volatility estimates are similar when the $\text{AR}(p)$ residual of the mean or median regression is considered, see column $\text{RMSD}(q\text{Reg})$. This is because both regression techniques produce comparable coefficient estimates in the pre-pandemic sample.

Table 4: Calibration of Minnesota priors before the pandemic

	RMSD (oReg)	MAD (oReg)	S_n (oReg)	Q_n (oReg)	RMSD (FD)	MAD (FD)	S_n (FD)	Q_n (FD)	RMSD (qReg)	MAD (qReg)	S_n (qReg)	Q_n (qReg)
EMP	0.19	0.92	0.83	0.90	2.65	1.40	1.48	1.51	1.03	0.73	0.73	0.85
UR	0.19	0.94	0.95	0.97	1.47	0.80	1.07	1.19	1.04	0.87	0.86	0.93
HW	0.45	0.92	0.97	0.98	1.59	0.88	1.04	1.15	1.01	0.94	0.91	0.97
HE	0.23	1.01	1.01	0.99	3.20	1.14	1.07	1.15	1.04	0.96	0.93	1.00
CON	0.39	0.83	0.88	0.91	2.09	1.05	1.10	1.14	1.01	0.85	0.85	0.91
GDP	0.51	0.81	0.88	0.94	1.64	0.87	0.91	0.96	1.02	0.85	0.84	0.90
IP	0.85	0.82	0.92	0.96	1.51	1.05	1.00	1.06	1.04	0.80	0.80	0.90
CU	0.91	0.90	0.98	0.96	1.44	1.01	1.05	1.10	1.05	0.81	0.84	0.89
HOUS	6.68	0.91	0.93	0.98	1.06	0.86	0.90	0.97	1.04	0.80	0.82	0.92
CPI	0.44	0.74	0.76	0.81	1.79	0.68	0.75	0.82	1.02	0.75	0.71	0.78
PCE	0.33	0.81	0.80	0.85	1.93	0.76	0.82	0.91	1.02	0.76	0.78	0.81
PCEc	0.13	0.96	1.01	1.01	4.08	1.27	1.33	1.35	1.03	0.97	0.95	0.99
OIL	14.06	0.86	0.85	0.92	1.05	0.79	0.88	0.91	1.02	0.78	0.81	0.86
PPI	0.98	0.81	0.82	0.86	1.19	0.69	0.79	0.84	1.01	0.83	0.80	0.85
T1y	0.35	0.57	0.64	0.72	1.23	0.88	0.84	0.96	1.01	0.57	0.63	0.72
T10y	0.34	1.05	1.03	1.05	1.08	1.11	1.13	1.11	1.01	1.02	0.99	1.06
BAA	0.31	0.68	0.65	0.71	1.09	0.69	0.70	0.74	1.03	0.64	0.62	0.69
SP500	5.36	0.83	0.83	0.87	1.15	0.80	0.83	0.88	1.02	0.74	0.83	0.85
M2	0.80	0.79	0.84	0.84	1.48	1.10	1.08	1.10	1.02	0.72	0.78	0.82
EER	2.81	1.00	1.02	1.04	1.07	0.93	1.00	1.02	1.01	1.05	1.00	1.02

(a) Variable-specific volatility estimates in absolute and relative terms to RMSD(oReg)

(I) Medium BVAR- \mathcal{N}												
$\kappa_{1,m}^*$	0.35	0.30	0.31	0.32	0.79	0.37	0.39	0.40	0.36	0.29	0.29	0.31
$\hat{\kappa}_{1,m}^*$	0.35	0.30	0.31	0.32	0.80	0.35	0.39	0.41	0.36	0.29	0.29	0.31
SHIFT	1.00	0.87	0.89	0.92	2.29	1.01	1.11	1.16	1.03	0.84	0.82	0.89
STD	0.00	0.08	0.08	0.06	0.88	0.26	0.24	0.23	0.01	0.08	0.08	0.06
(II) Large BVAR- \mathcal{N}												
$\kappa_{1,m}^*$	0.29	0.24	0.25	0.26	0.45	0.27	0.28	0.30	0.30	0.24	0.24	0.25
$\hat{\kappa}_{1,m}^*$	0.29	0.25	0.25	0.26	0.49	0.27	0.29	0.30	0.30	0.24	0.24	0.26
SHIFT	1.00	0.86	0.88	0.91	1.69	0.94	0.99	1.04	1.02	0.82	0.82	0.89
STD	0.00	0.11	0.11	0.09	0.77	0.19	0.19	0.18	0.01	0.12	0.10	0.09

(b) Optimal shrinkage and relation to different volatility estimation strategies

Notes: Panel (a) shows for each calibration strategy the variable-specific volatility estimates in absolute and relative terms to RMSD(oReg) based on the pre-pandemic sample. Panel (b) reports the optimal ($\kappa_{1,m}^*$) and volatility-adjusted ($\hat{\kappa}_{1,m}^*$) shrinkage parameter for both BVAR- \mathcal{N} models. SHIFT and STD are the mean and standard deviation of variable-specific volatility estimates relative to those of RMSD(oReg).

Third, the RMSD generally yields higher volatility estimates than the robust volatility estimators. Overall, the robust estimators produce fairly similar estimates. Moreover, a larger difference between RMSD and robust estimators suggests that the input series for variable r exhibits some non-normal behavior. This difference is generally larger when the first difference is used as the input series. Therefore, the S_n and Q_n statistic should

be preferred over the MAD in these cases as they are more efficient and may be more informative due to their ability to account for skewness in the distribution.

What are the implications of these characteristics on $\kappa_{1,m}^*$, the optimal degree of shrinkage for the VAR coefficients? Table 4 (b) shows that $\kappa_{1,m}^*$ is generally larger for the RMSD estimator and first-differenced input series as compared to analogous robust metrics in both BVARs.^{18,19} This difference can be rationalized by noting from (5) that if the ratio of variable-specific volatility estimates of method j and i is constant for all r variables then the optimal $\kappa_{1,j}$ and $\kappa_{1,i}$ are proportional:

$$\kappa_{1,j} = \kappa_{1,i} \left(\frac{1}{n} \sum_{r=1}^n \left(\frac{\hat{\sigma}_{r,j}}{\hat{\sigma}_{r,i}} \right) \right), \quad (10)$$

where $\hat{\sigma}_{r,\cdot}$ is the variable r specific volatility estimate of method j and i respectively.

The second row in each subtable of Panel (b) shows that the volatility-adjusted hyperparameter $\hat{\kappa}_{1,m}^*$ is fairly close to its optimal value $\kappa_{1,m}^*$. This indicates that the bulk of the variation can be explained by adjusting $\kappa_{1,m}^*$ for an average level shift between volatility estimates (SHIFT). This approximation becomes more accurate the more proportional the variable-specific volatility estimates are across calibration methods. The degree of proportionality may be measured by the standard deviation of relative volatility estimates (STD). The STD statistic shows that residual based strategies yield more proportional volatility estimates than the first-differenced one, and, hence a more comparable prior calibration. Taken together, this is an encouraging empirical fact as established values for $\kappa_{1,m}^*$ in various empirical applications may be easily mapped by (10) into an alternative calibration strategy without necessarily re-optimizing this tuning parameter.

Next, the influence of the pandemic observations on the Minnesota prior is discussed. Table 5 (a) reports the ratio of volatility estimates from 2022:Q1 to 2019:Q4 and (b) presents statistics summarizing the average change of the ratio of volatility (MEAN) and dispersion of the ratio of volatility measured against an unchanged ratio (DEV).

The table shows that the prior based on the standard Minnesota calibration strategy is substantially altered as some variable-specific volatility estimates are strongly inflated, see column RMSD(oReg). Overall, the prior distribution becomes disproportionately tighter for the VAR coefficients associated with indicators of real activity and labor market conditions (except hourly earnings), while it is hardly affected for price and financial variables (except real M2 stock). More specifically, the volatility estimate of employment is over 6 times larger than its pre-pandemic value, for instance. This means that the updated parameters set an extremely tight prior for the own and cross-equation dynamic lags of employment, roughly $42 (\approx 6.46^2)$ times tighter than its pre-pandemic value.

To further illustrate the impact of a prior re-calibration on the VAR dynamics, Figure 2 shows impulse response functions to a one standard deviation shock in employment of the Gaussian BVAR estimated until pre-pandemic times but calibrating the standard Minnesota prior with data until 2019:Q4, 2020:Q4 and 2022:Q1, respectively.²⁰ The figure shows that the revised prior distribution has a decisive impact on the posterior dynamic

¹⁸ $\kappa_{1,m}^*$ is identified as the marginal likelihood maximizing hyperparameter on the grid (0.01 : 0.01 : 10).

¹⁹ $\kappa_{1,m}^*$ is smaller for the larger BVAR model as higher dimensional VAR models require a more parsimonious prior to reduce estimation uncertainty and overfitting, see [Giannone, Lenza, and Primiceri \(2015\)](#).

²⁰The shock has no structural interpretation and is just used to study dynamic propagation.

Table 5: Re-calibration of Minnesota priors during the pandemic

	RMSD (oReg)	MAD (oReg)	S_n (oReg)	Q_n (oReg)	RMSD (FD)	MAD (FD)	S_n (FD)	Q_n (FD)	RMSD (qReg)	MAD (qReg)	S_n (qReg)	Q_n (qReg)
EMP	6.46	1.38	1.80	1.74	2.54	1.05	1.10	1.16	7.71	0.96	1.05	1.02
UR	4.79	1.43	1.70	1.60	3.40	1.17	1.00	1.00	5.39	0.99	1.10	1.08
HW	3.46	1.23	1.22	1.29	2.21	1.08	1.11	1.12	3.90	0.98	1.08	1.05
HE	1.06	1.04	1.07	1.08	1.04	1.04	1.07	1.06	1.04	1.04	1.10	1.05
CON	3.22	1.32	1.35	1.38	1.77	1.00	1.04	1.05	3.50	0.97	1.05	1.11
GDP	2.31	1.08	1.09	1.12	1.59	1.07	1.10	1.10	2.41	1.00	1.09	1.08
IP	2.17	1.13	1.07	1.13	1.50	0.98	1.01	1.04	2.20	1.06	1.08	1.03
CU	2.19	1.06	0.99	1.11	1.57	1.09	1.08	1.07	2.21	1.06	1.00	1.06
HOUS	1.13	0.93	0.96	0.99	1.10	0.99	0.98	1.00	1.12	0.96	1.00	1.01
CPI	1.13	0.96	1.05	1.07	1.07	1.09	1.12	1.09	1.12	1.01	1.14	1.12
PCE	1.10	1.06	1.03	1.06	1.07	1.08	1.13	1.10	1.09	1.12	1.11	1.10
PCEc	1.33	1.20	1.09	1.13	1.07	1.05	1.08	1.09	1.30	1.19	1.14	1.14
OIL	1.08	1.01	1.09	1.05	1.07	1.07	1.01	1.04	1.09	1.07	1.04	1.04
PPI	1.12	1.02	1.10	1.09	1.15	1.19	1.16	1.14	1.11	1.10	1.07	1.11
T1y	1.00	1.02	1.03	1.00	1.00	0.94	0.98	0.98	1.00	1.06	1.07	0.99
T10y	1.00	0.99	1.02	1.00	1.00	1.01	1.04	1.00	1.00	1.00	1.04	0.99
BAA	1.01	0.97	1.00	1.01	1.01	1.03	1.03	1.03	1.01	1.00	1.05	1.01
SP500	1.01	0.97	0.98	1.01	1.01	1.03	1.03	1.03	1.01	1.05	0.99	1.01
M2	1.62	1.12	1.11	1.11	1.42	1.07	1.06	1.06	1.63	1.06	1.03	1.08
EER	0.98	0.95	0.96	0.96	0.98	0.98	0.98	0.96	0.98	0.92	0.97	0.97

(a) Ratio of volatility: 2022:Q1 to 2019:Q4

(I) Medium BVAR- \mathcal{N}												
MEAN	3.20	1.23	1.35	1.34	1.91	1.07	1.07	1.08	3.57	1.02	1.10	1.09
DEV	2.91	0.28	0.46	0.43	1.23	0.09	0.08	0.09	3.48	0.08	0.10	0.10
(II) Large BVAR- \mathcal{N}												
MEAN	1.96	1.09	1.14	1.15	1.43	1.05	1.06	1.06	2.09	1.03	1.06	1.05
DEV	1.73	0.17	0.26	0.25	0.75	0.08	0.08	0.08	2.05	0.07	0.08	0.07

(b) Descriptive statistics

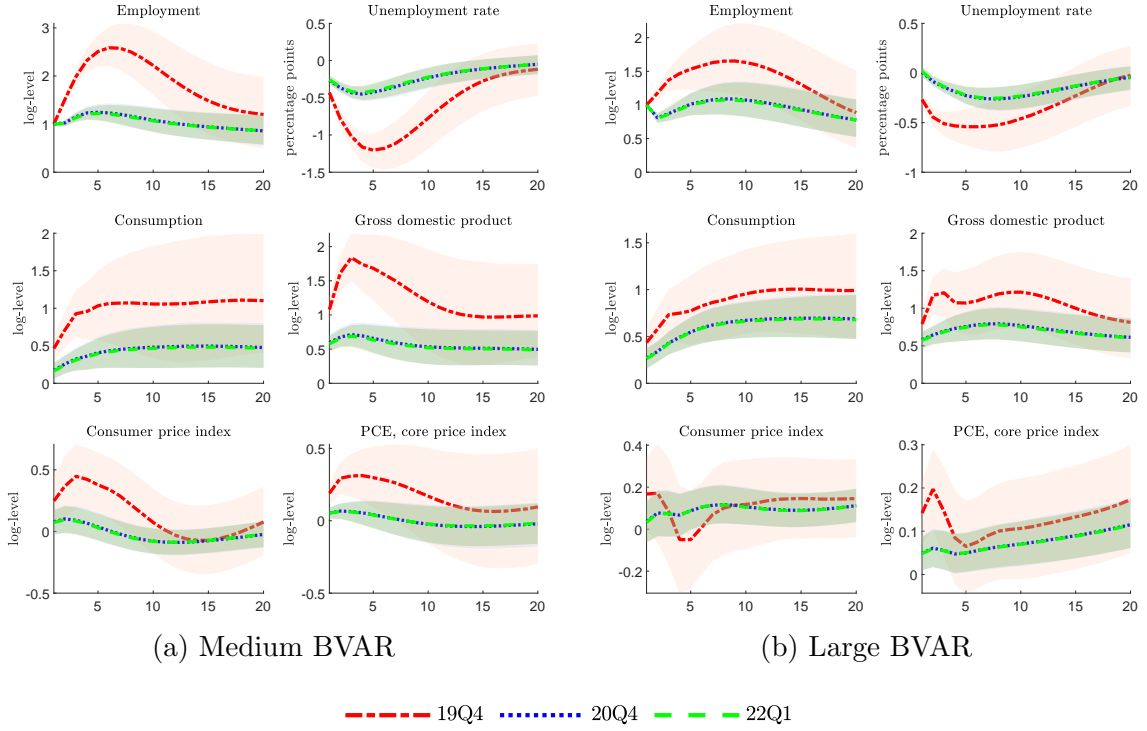
Notes: Panel (a) presents the ratio of variable-specific volatility estimates of 2022:Q1 relative to 2019:Q4. Panel (b) reports descriptive statistics of the ratio of volatilities: MEAN is the average and DEV is the root of mean squared deviation from one.

propagation of the employment shock in both the medium and large VAR.²¹ It a-priori mutes the internal propagation of the shock by shrinking cross-dynamic relations to zero and thereby breaks historical links between variables such as employment and prices. Thus, the COVID-19 shock not only has a decisive impact on the estimation stage of a standard Bayesian VAR but also on the prior calibration stage, which is an important ingredient for modelling macroeconomic transmission channels during the pandemic.

In contrast, volatility estimates based on robust scale estimators (MAD, S_n , Q_n)

²¹Re-optimizing $\kappa_{1,m}^*$ or accounting for an average level shift in volatility estimates is insufficient to recover a comparable pre-pandemic prior distribution as the change in variable-specific volatility estimates is too heterogeneous.

Figure 2: Impact of prior re-calibration on VAR dynamics



Notes: Impulse response functions to a one standard deviation shock in employment of the BVAR- \mathcal{N} estimated until 2019:Q4 with RMSD(*oReg*) Minnesota prior calibrated on samples until 2019:Q4, 2020:Q4 and 2022:Q1, respectively. The thick lines are posterior median estimates and the colored area depicts the 16% to 84% credible interval.

coupled with robust input series (first difference or median $AR(p)$ residual) are hardly affected by the pandemic observations, see Table 5 (a) and (b). Thus, the Minnesota prior based on a two-dimensional robustified calibration strategy remains comparable before and throughout the pandemic and also does not a-priori mute internal propagation of shocks in a VAR when the prior is revised during the pandemic.

3.3 Model and prior comparison

To discriminate between these Bayesian VAR models and prior specifications, the marginal likelihood is used as a formal Bayesian model selection criterion following Chan (2020).²² The marginal likelihood under model $M_{k,m}$ is defined as

$$p(y|M_{k,m}) = \int p(y|\theta_k, M_{k,m})p(\theta_k|M_{k,m})d\theta_k,$$

²²The marginal likelihood does not always favor the most general model as it trades off model fit against complexity. Intuitively, the higher the dimension of the parameter space, the more spread the prior is with respect to the likelihood. This is known as Occam's razor.

where $p(y|\theta_k, M_{k,m})$ is the likelihood function, $p(\theta_k|M_{k,m})$ is the prior distribution, θ_k is a model-specific parameter vector and m is a prior calibration. For the BVAR- \mathcal{N} , the analytical formula is used to compute the marginal likelihood, while for the other BVAR models no closed form exists and Chib’s method (Chib, 1995) is used instead.

Specifically, if the marginal likelihood under model $M_{i,s}$ is larger than under $M_{j,q}$, then the data is more likely under model $M_{i,s}$ as compared to $M_{j,q}$. Given all models are a-priori equally likely, the weight of evidence between two models can be measured by the Bayes factor defined as the ratio of marginal likelihoods, see Chan (2017). Table 6 presents log Bayes factors against the BVAR- \mathcal{N} with weakly informative prior before and during the pandemic for the medium VAR, for the large VAR see Table A.1 in Appendix A.²³

Table 6: Bayes factor in medium BVAR

Prior	BVAR- \mathcal{N}	BVAR- t	BVAR- \mathcal{N} -CSV	BVAR- t -CSV
Weak	•	12.69 (0.06)	41.17 (0.10)	43.38 (0.10)
RMSD (<i>oReg</i>)	75.08 –	85.71 (0.04)	117.68 (0.08)	119.50 (0.10)
MAD (<i>oReg</i>)	72.95 –	84.36 (0.05)	116.39 (0.04)	118.45 (0.07)
S_n (<i>oReg</i>)	74.44 –	85.73 (0.04)	117.46 (0.06)	119.48 (0.08)
Q_n (<i>oReg</i>)	74.51 –	85.69 (0.03)	117.94 (0.05)	119.76 (0.10)
RMSD (<i>FD</i>)	64.36 –	76.04 (0.05)	106.49 (0.06)	108.48 (0.11)
MAD (<i>FD</i>)	66.61 –	78.70 (0.04)	106.99 (0.05)	108.98 (0.12)
S_n (<i>FD</i>)	67.42 –	79.55 (0.04)	109.78 (0.07)	111.79 (0.10)
Q_n (<i>FD</i>)	68.21 –	80.23 (0.05)	111.52 (0.09)	113.77 (0.10)
RMSD (<i>qReg</i>)	74.91 –	85.61 (0.04)	117.48 (0.06)	119.51 (0.14)
MAD (<i>qReg</i>)	75.30 –	86.37 (0.05)	118.61 (0.07)	120.48 (0.10)
S_n (<i>qReg</i>)	74.96 –	86.17 (0.05)	117.60 (0.06)	119.62 (0.08)
Q_n (<i>qReg</i>)	74.75 –	85.92 (0.05)	117.99 (0.07)	119.81 (0.13)

(a) Estimation sample until 2019:Q4

Weak	•	216.54 (0.21)	237.84 (0.24)	244.32 (0.24)
RMSD (<i>oReg</i>)	–20.29 –	160.71 (0.05)	275.63 (0.10)	285.00 (0.10)
MAD (<i>oReg</i>)	65.68 –	287.09 (0.13)	316.52 (0.16)	325.02 (0.16)
S_n (<i>oReg</i>)	64.53 –	283.20 (0.11)	316.23 (0.17)	325.14 (0.17)
Q_n (<i>oReg</i>)	65.58 –	283.89 (0.11)	317.88 (0.10)	326.22 (0.10)
RMSD (<i>FD</i>)	45.17 –	252.61 (0.09)	306.40 (0.20)	315.04 (0.20)
MAD (<i>FD</i>)	64.57 –	289.24 (0.13)	308.96 (0.16)	317.57 (0.16)
S_n (<i>FD</i>)	65.59 –	289.18 (0.10)	310.97 (0.21)	318.93 (0.21)
Q_n (<i>FD</i>)	66.13 –	288.92 (0.17)	311.44 (0.20)	320.45 (0.20)
RMSD (<i>qReg</i>)	–40.31 –	135.77 (0.08)	263.50 (0.10)	272.42 (0.10)
MAD (<i>qReg</i>)	66.07 –	294.17 (0.13)	314.50 (0.19)	322.72 (0.19)
S_n (<i>qReg</i>)	67.85 –	294.78 (0.12)	317.20 (0.16)	325.52 (0.16)
Q_n (<i>qReg</i>)	68.42 –	294.78 (0.16)	317.43 (0.16)	326.27 (0.16)

(b) Estimation sample until 2022:Q1

Notes: Log Bayes factors of various BVARs and prior combinations against BVAR- \mathcal{N} with weakly informative prior. Bold figures indicate maximum Bayes factor for each prior specification. Brackets report the numerical standard error.

²³Complementary to these tables, Figure A.1 in Appendix A shows the log marginal likelihood when the models are estimated over an expanding window. Notice the marginal likelihood cannot be directly compared across different points in time as information sets differ due to alternative sample sizes.

Before the pandemic, all Minnesota priors are decisively favored over the naïve benchmark prior, see Panel (a). For instance, the average log Bayes factor is 72 in the medium BVAR which means the BVARs with Minnesota prior are 1.85×10^{31} more likely than those with a naïve benchmark prior. This overwhelming support for shrinking more distant VAR coefficients to zero is not surprising as more distant lags are relatively less informative for current dynamics. Moreover, the sample fit of the standard $\text{RMSD}(o\text{Reg})$ and robustified calibration strategies is broadly similar. The figures based on $\text{AR}(p)$ residuals are almost identical across volatility estimators, while those based on the first difference are somewhat smaller. Thus, two-dimensional robustified calibration strategies do not sacrifice in-sample fit for robustness.

Across all prior specifications, the BVAR- t -CSV is the best performing model albeit the Bayes factor is just immaterially larger than that of the BVAR- \mathcal{N} -CSV model. Both models have, however, a substantial margin over the BVAR- t model, while the latter clearly outperforms the standard Gaussian BVAR- \mathcal{N} model. Thus, both error extensions are favored over standard Gaussian errors by the data. But on a single ingredient basis, stochastic volatility is a more important feature than a heavy-tailed error distribution. This is line with the findings of [Cúrdia et al. \(2014\)](#), [Chiu et al. \(2017\)](#) and [Chan \(2020\)](#) who also look at the U.S. but consider a different data set and an earlier period.

How does the pandemic affect the model fit across these BVARs? Panel (b) shows that while the overall ranking across BVARs remains unchanged, the relative distance between Bayes factor with respect to the best performing model changes markedly before and during the pandemic. For the naïve benchmark prior, the Bayes factor falls by about 200 and 290 log points in the medium and large BVAR- \mathcal{N} , respectively.²⁴ This indicates that the classical Bayesian VAR is poorly equipped to deal with these extreme observations. Moreover, the Bayes factor increases by 2.91 log points for BVAR- t , while it falls by 4.28 for the BVAR- \mathcal{N} -CSV in the medium VAR. This relative change is even more pronounced in the large VAR – with an increase of 9.9 and a decrease of 13.82 log points, respectively. This suggests that the heavy-tailed error extension might be better equipped to capture the extreme variation during the pandemic than the common stochastic volatility specification.

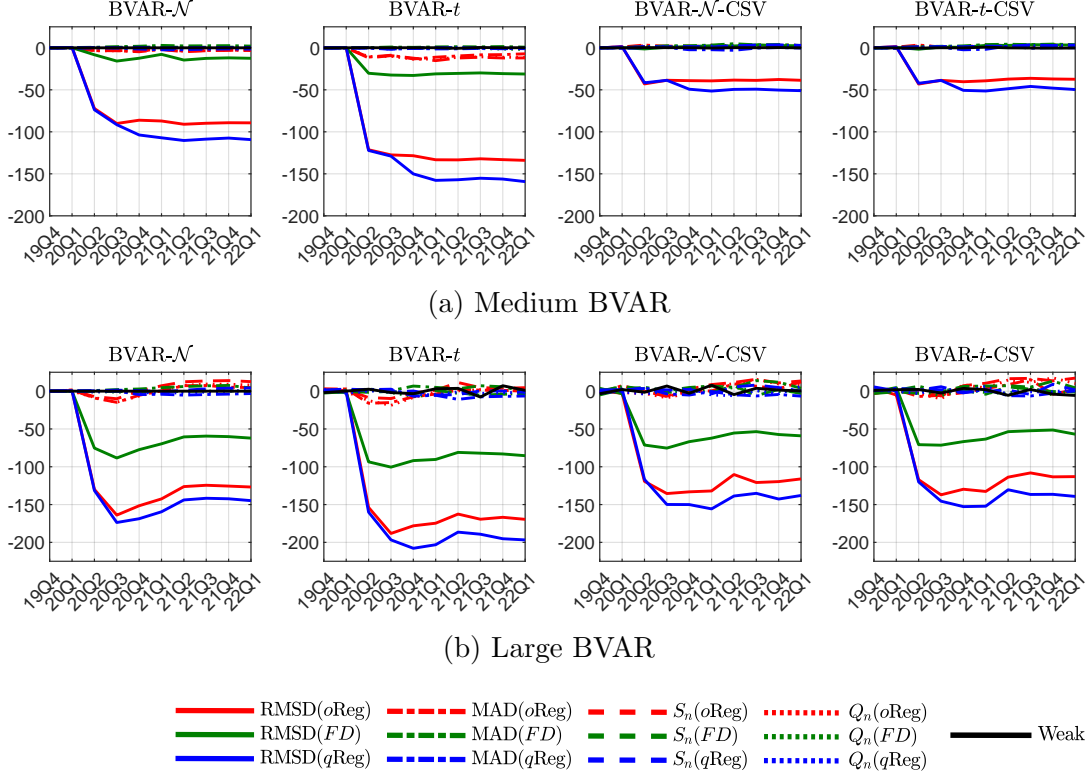
Turning to prior sensitivity, the Minnesota calibrations based on the $\text{AR}(p)$ residual and RMSD volatility estimator yield a substantially lower in-sample fit as compared to the robust calibration strategies. For instance, the Bayes factor between $\text{MAD}(q\text{Reg})$ and $\text{RMSD}(o\text{Reg})$ of the BVAR- t -CSV is 37.72 in the medium and 115.43 in the large model, suggesting that the robust prior calibration is decisively preferred. In addition, the naïve benchmark prior in the medium BVAR with Gaussian and t -distributed errors is now decisively favored over these non-robust Minnesota calibrations. Furthermore, Bayes factors of robust calibration strategies are very similar across specifications. Thus, the more general S_n or Q_n statistic of [Rousseeuw and Croux \(1993\)](#) do not provide a substantial improvement over the scaled MAD in terms of model fit.

To complement this analysis and evaluate the impact of prior revisions on model fit in and over time, Figure 3 shows Bayes factors of the BVARs estimated with re-calibrated prior against a 2019:Q4 fixed prior obtained over an expanding estimation window from 2019:Q4 to 2022:Q1. The figure shows that re-calibrating the Minnesota prior with the

²⁴The Bayes factor is computed as $\Delta BF_{t,t+h,t\text{-CSV}-k,m} = BF_{t,t\text{-CSV}-k,m} - BF_{t+h,t\text{-CSV}-k,m}$ where $BF_{t,t\text{-CSV}-k} = BF_{t,t\text{-CSV}} - BF_{t,k}$, $k = \{\mathcal{N}, t, \mathcal{N}\text{-CSV}\}$, $t = 2019:\text{Q4}$ and $t+h = 2022:\text{Q1}$ and prior m .

standard calibration strategy (red thick colored line) has a decisive negative effect on the overall in-sample fit starting in 2020:Q2. In contrast, a fully robust calibration hardly affects the Bayes factor and thus yields about the same model fit as fixing the prior.²⁵

Figure 3: The costs of re-calibrating the prior distribution



Notes: Log Bayes factor between BVARs with re-calibrated and 2019:Q4 fixed prior distribution estimated over an expanding sample from 2019:Q4 to 2022:Q1.

In the remainder of the paper, empirical results are presented for the MAD(*qReg*) Minnesota prior. This calibration is chosen as the preferred fully robust alternative to the RMSD(*oReg*) for three reasons. First, it is resilient to extreme observations – not only those related to the pandemic. Second, it yields asymptotically the same point estimates as the RMSD(*oReg*) when the data is normally distributed and not contaminated by outliers. Third, it is simpler to implement than similar strategies with the S_n and Q_n statistic.

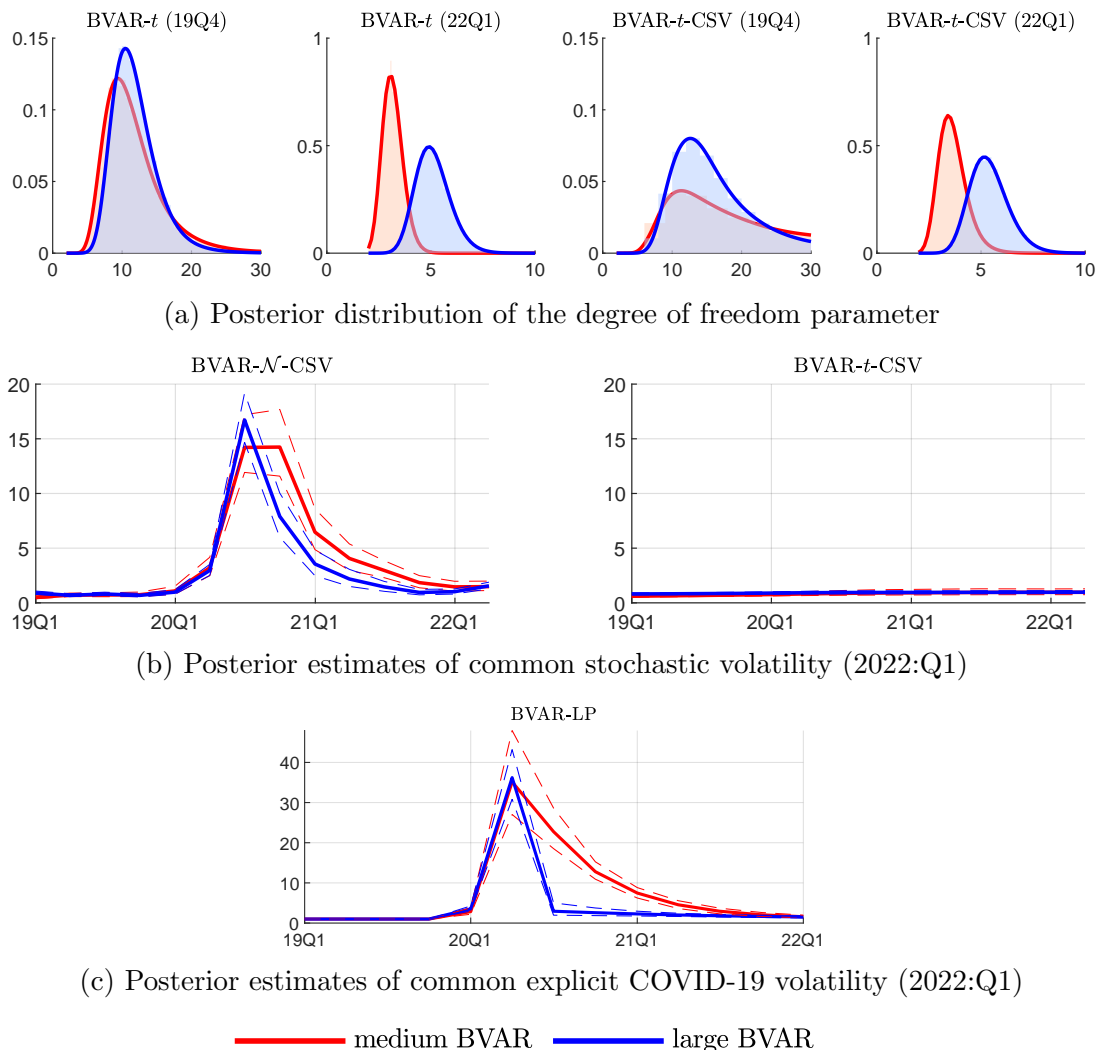
3.4 Macroeconomic tail risk and volatility

Model diagnostics strongly reject Gaussianity, but how do the pandemic observations affect macroeconomic tail risk and volatility? This section discusses the implications and adds the recently proposed explicit common volatility BVAR model of [Lenza and Prim-](#)

²⁵In the large VAR, the Bayes factor of two-dimensional robustified calibration strategies exhibits some fluctuations due to a higher estimation error for the marginal likelihood. The estimation error in the medium VAR is negligible.

iceri (2022) as another benchmark (denoted as BVAR-LP).²⁶

Figure 4: Macroeconomic tail risk and volatility during the pandemic



Notes: BVARs with MAD(q Reg) Minnesota prior estimated until 2019:Q4 and 2022:Q1, respectively. The thick line is the posterior median and the dashed lines are the 16% to 84% credible interval.

Figure 4 shows how macroeconomic tail risk and volatility are affected by the pandemic observations.²⁷ Recall, the lower the degree of freedom parameter ν , the stronger the t distribution departs from Gaussianity and the heavier the tails. Panel (a) shows that the posterior of ν peaks around 10 before the pandemic in all t error models, indicating already considerable tail risk.²⁸ Once the pandemic observations are included, the posterior distribution becomes much tighter and peaks around 3.5 in the medium (red area)

²⁶The BVAR-LP is estimated with the MAD(q Reg) Minnesota prior and $\kappa_{1,m}$ is not re-optimized using the Giannone et al. (2015) method for comparability. For estimation details and results, see Appendix B.

²⁷Volatility estimates prior to the pandemic are shown in Figure A.2 in Appendix A.

²⁸In the BVAR- t -CSV, the posterior distribution of ν is more spread as it can explain large shocks either by changes in stochastic volatility or by allowing for heavier tails.

and 5 in the large model (blue area), respectively. For the BVAR- t , this is not surprising as this is the model’s main channel to accommodate more extreme innovations. However, the much sharper identification in the BVAR- t -CSV means that the tail risk increased strongly besides changes in macroeconomic volatility.

The nature of volatility during the pandemic hinges on the error structure and dimension of the VAR models, see Panel (b) and (c). Both the BVAR- \mathcal{N} -CSV and BVAR-LP interpret the pandemic shock as an enormous increase in macroeconomic volatility in 2020:Q2, about 35 standard deviations in the BVAR-LP and half this size in the BVAR- \mathcal{N} -CSV. After the initial burst, the volatility profiles start to diverge. Volatility remains high for an extended period in the stochastic volatility specification, while in the BVAR-LP the profiles differ across VAR dimensions. In the medium model, volatility decays slowly while in the large model volatility jumps back immediately to a more moderate level.²⁹ Thus, the error structure and cross-sectional information plays an important role for identifying the nature of volatility with these VARs featuring no tail risk.

In contrast, the BVAR- t -CSV – which allows for both tail risk and stochastic volatility – assigns a completely different interpretation to the COVID-19 shock. The common stochastic volatility series is almost entirely flat during the pandemic period. This means that the model interprets the pandemic as a rare event and not as a (persistent) increase in macroeconomic volatility.³⁰ Since fat tails are an important feature of the data before and during the pandemic, ignoring it may lead to misleading inference about the nature of volatility and other objects of interest. This property was also previously pointed out by [Jacquier et al. \(2004\)](#), [Cúrdia et al. \(2014\)](#), [Chiu et al. \(2017\)](#) and [Chan \(2020\)](#) for different data sets and sample periods.

3.5 Model (in)stability and forecasting

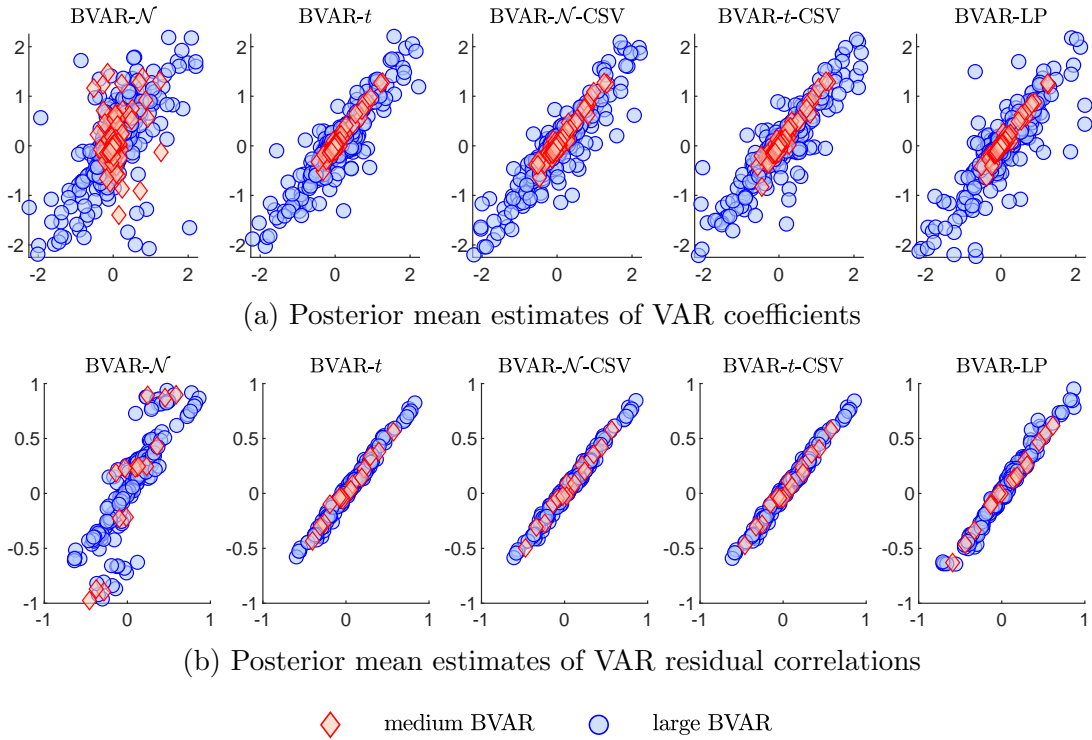
Having discussed the implications of the COVID-19 shock for tail risk and volatility, this section focuses on model (in)stability and forecast properties. Figure 5 shows scatter plots of the posterior mean VAR coefficients and residual correlation matrix obtained in 2019:Q4 against 2022:Q1. Panel (a) shows that the VAR coefficients in the BVAR- \mathcal{N} are substantially revised with intercepts and coefficients measuring (cross)-dynamics of employment, the unemployment rate, consumption and output being particularly affected. For instance, the first-order autoregressive coefficient of employment changes from 1.28 in 2019:Q4 to -0.12 in 2022:Q1. In contrast, the VAR parameters in the more flexible medium BVARs are hardly influenced by the extreme observations. This robustness also extends to the large BVARs which however exhibit some revision in the VAR parameters, but nothing extraordinary as compared to the BVAR- \mathcal{N} .

The pandemic observations induce a drastic and long-lasting change in the residual correlation structure in both BVAR- \mathcal{N} models, see Panel (b). For instance, the average correlation between the residual of the unemployment rate and those of employment, consumption and output changes from -0.37 in 2019:Q4 to -0.91 in 2022:Q1 in the medium BVAR. This abrupt revision is solely driven by the very synchronized co-movement of

²⁹The data informs the large BVAR-LP quickly about the nature of volatility while the estimate of the medium model is sensitive to incoming observations, see Figure B.2 Appendix B.

³⁰Figure A.3 in Appendix A shows the latent volatility of the t error models, which characterize the pandemic as a 28 (BVAR- t) and 21 (BVAR- t -CSV) standard deviation shock, respectively.

Figure 5: Change in common VAR parameters



Notes: Common VAR parameters in 2019:Q4 (x-axis) against 2022:Q1 (y-axis) for various BVARs with MAD(q Reg) Minnesota prior.

these variables in response to the COVID-19 shock. The residual correlations in the BVARs with a more flexible error structure, however, are only mildly influenced by the pandemic observations as they downweigh the information from these extreme realizations. Therefore, inference based on the traditional Gaussian BVAR may be overshadowed by the extreme size of the COVID-19 shock for an extended period of time as its impact on the residual covariance matrix cannot wash out easily at quarterly frequency.

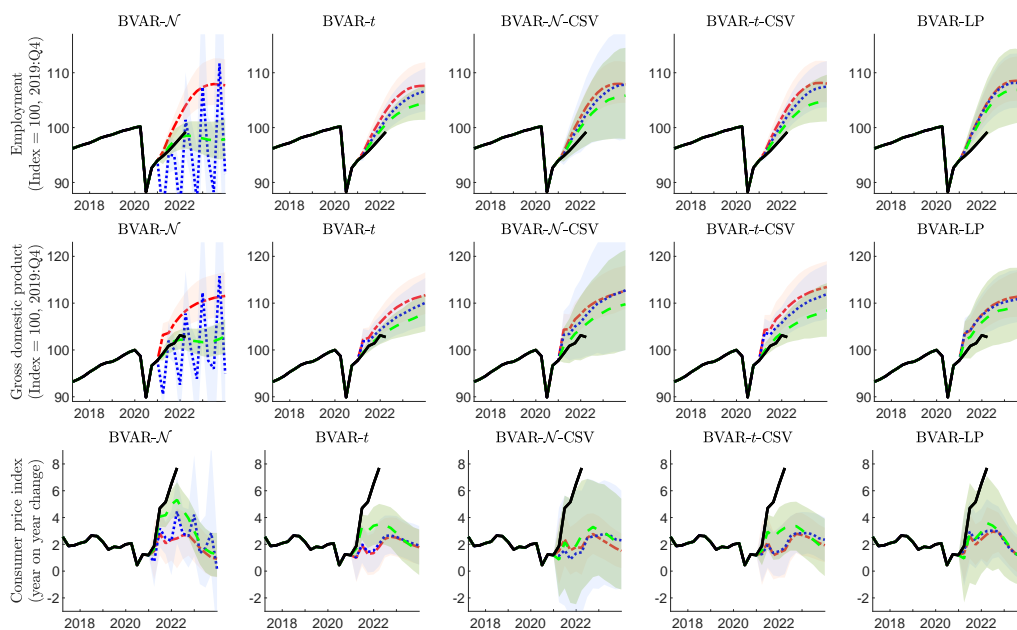
Turning to forecast properties, parameter revisions and prior re-calibrations also have a decisive impact on unconditional projections. Focusing first on the impact of parameter revisions, Figure 6 shows unconditional projections starting in 2021:Q1 for employment, GDP and CPI inflation based on BVARs with MAD(q Reg) Minnesota prior and estimated over different samples, stopping in 2019:Q4, 2020:Q4 and 2022:Q1.³¹ This exercise reveals several important insights into how well these models are able to digest the COVID-19 shock and whether they can put this information to good use in hindsight.

First, parameter instability may cause serious problems for forecasting with Gaussian VARs. Looking at the predictions for employment, for instance, the updated parameters in 2020:Q4 trigger oscillations in the projections of the medium-sized model (blue dotted line).³² Though these oscillating dynamics eventually wash out as more information become available (see predictions with 2022:Q1 parameters, green dashed line), the extreme

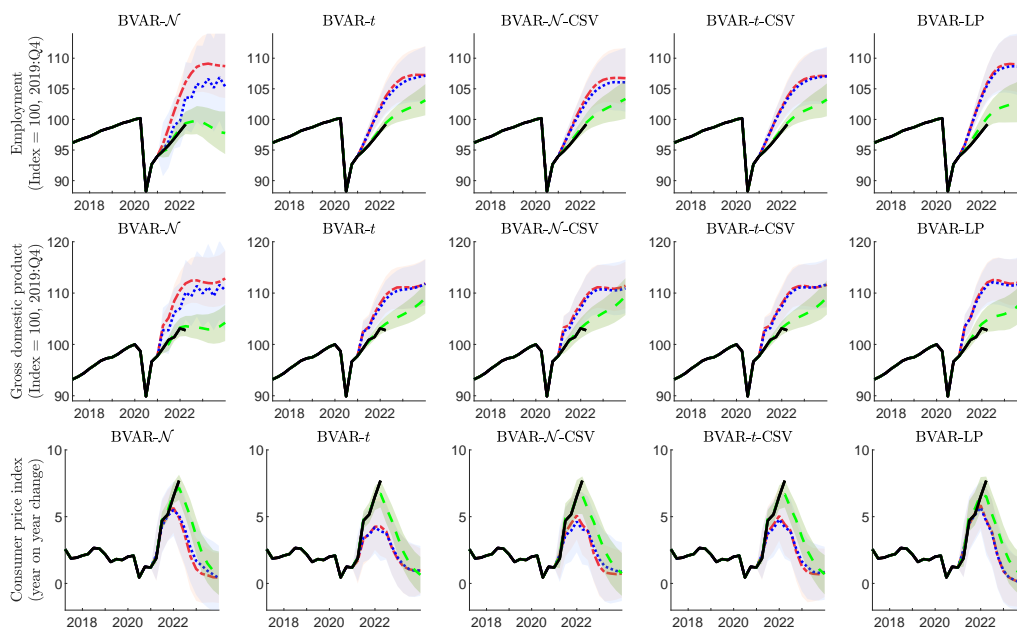
³¹For the unemployment rate, consumption and PCE core inflation, see Figure ?? in Appendix A.

³²The short-run dynamics in the large model are less affected. This might be explained by the fact that it features a larger cross-sectional information set and a tighter optimal shrinkage parameter.

Figure 6: Impact of re-estimated parameters on unconditional forecasts starting 2021:Q1



(a) Medium BVAR



(b) Large BVAR

--- 19Q4 20Q4 --- 22Q1 ——— actual

Notes: BVARs with $MAD(qReg)$ Minnesota prior estimated until 2019:Q4, 2020:Q4 and 2022:Q1, respectively. The thick line is the posterior median and the colored area depicts the 16% to 84% credible interval.

observations also alter long-run forecasts of the variables. For instance, gross domestic product is projected to stagnate in the long run, while all of the more flexible BVARs

suggest a continuation of the pre-pandemic growth pace. This complements the findings in [Lenza and Primiceri \(2022\)](#), [Carriero et al. \(2022\)](#), [Schorfheide and Song \(2021\)](#) and [Bobeica and Hartwig \(2023\)](#) by the fact that the COVID-19 observations may not only distort the short-run dynamics but also have a persistent impact on long-term predictions.

Second, the more flexible BVARs are robust to the COVID-19 shock. They produce stable forecasts in the pandemic period and are able to track the transmission of the shock through the U.S. economy by orderly revising the VAR structure. At first, the projections in both medium and large VARs are hardly affected by the incorporation of four additional quarters, see projections with 2019:Q4 parameters (red line) vs 2020:Q4 parameters (blue dotted line). However, as time passes, the VAR better understands the nature of the COVID-19 shock and unconditional projections are more visibly revised (green dashed line). In fact, the large VAR is able to closely track the actual developments of all six core U.S. macroeconomic variables whereas the medium model misses some key patterns (e.g. underestimation of inflationary pressure). Thus, this analysis suggests that the COVID-19 shock might have altered macroeconomics dynamics and that extreme observation-robust BVARs are well equipped to track these changes if they feature a sufficiently large cross-sectional information set. Furthermore, traditional Gaussian VARs cannot easily cope with such extreme observations and induce non-sensible parameter revisions instead.

Third, the COVID-19 shock may substantially inflate density forecasts intervals when the VAR only allows for common time-varying volatility and lacks a sufficiently large cross-sectional dimension. Specifically, forecasts based on the medium BVAR- \mathcal{N} -CSV and BVAR-LP may exhibit extremely wide forecast intervals whereas the predictive intervals of the large models are comparable to those featuring t errors. This complements the evidence in [Carriero et al. \(2022\)](#) by the fact that though predictive intervals based on variable-specific volatility structure may become excessively wide after the pandemic shock, the same is not necessarily true for a common volatility structure.

The re-calibration of the standard Minnesota prior also affects forecast properties and reveals a trade-off for the prediction of real and nominal variables.³³ To illustrate this, [Figure 7](#) shows unconditional projections starting in 2021:Q1 for employment, GDP and CPI inflation based on BVARs with the fully robust MAD(q Reg) and standard RMSD(o Reg) Minnesota prior calibration estimated until 2020:Q4.³⁴ As discussed in [Section 3.2](#), the RMSD(o Reg) calibration sets a very tight prior on the VAR coefficients for most real variables. For forecasting purposes, this re-calibration may be beneficial when predicting real variables at the beginning of the pandemic as the dampened internal propagation of shocks results in less overshooting predictions. However, the same channel prevents the large VAR from tracking the mounting inflationary pressure in 2021, making it a less attractive choice for understanding and predicting current inflation dynamics.

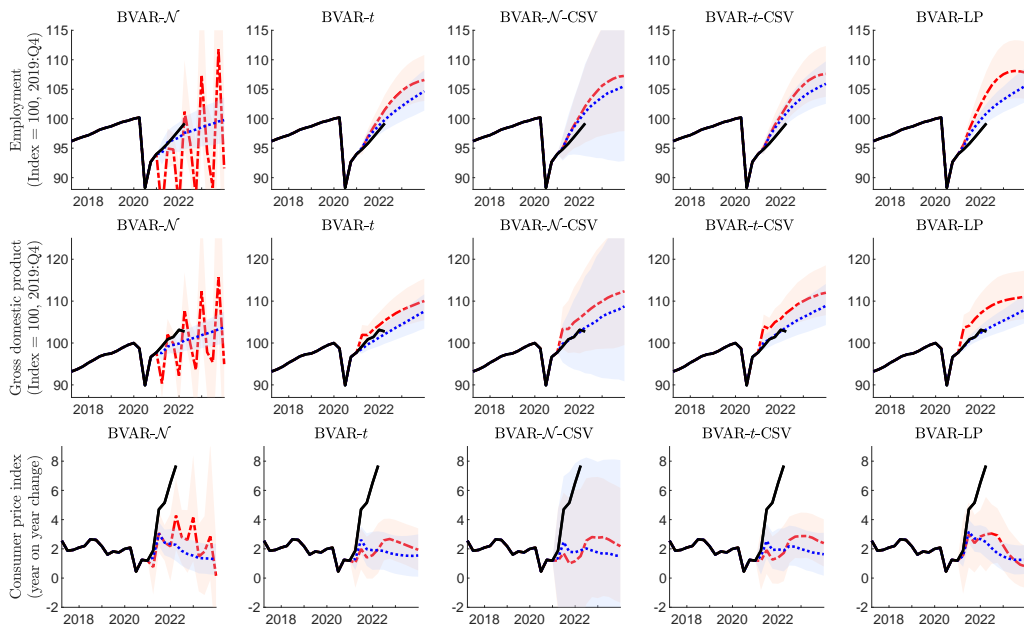
The previously discussed forecast properties are supported by a very short forecast evaluation from 2020:Q1 until 2022:Q1, which should be viewed as suggestive evidence. [Table 7](#) shows root mean squared forecast errors (RMSFE) of a pseudo-out-of-sample forecast evaluation of unconditional forecasts for the one-quarter-ahead and one-year-ahead horizon.³⁵ This exercise underlines the importance of using robust BVARs and carefully

³³The VAR higher-order and cross-dynamic lags are strongly shrunk to zero for many coefficients due to the tighter prior for most real variables, see [Figure A.5](#) in [Appendix A](#).

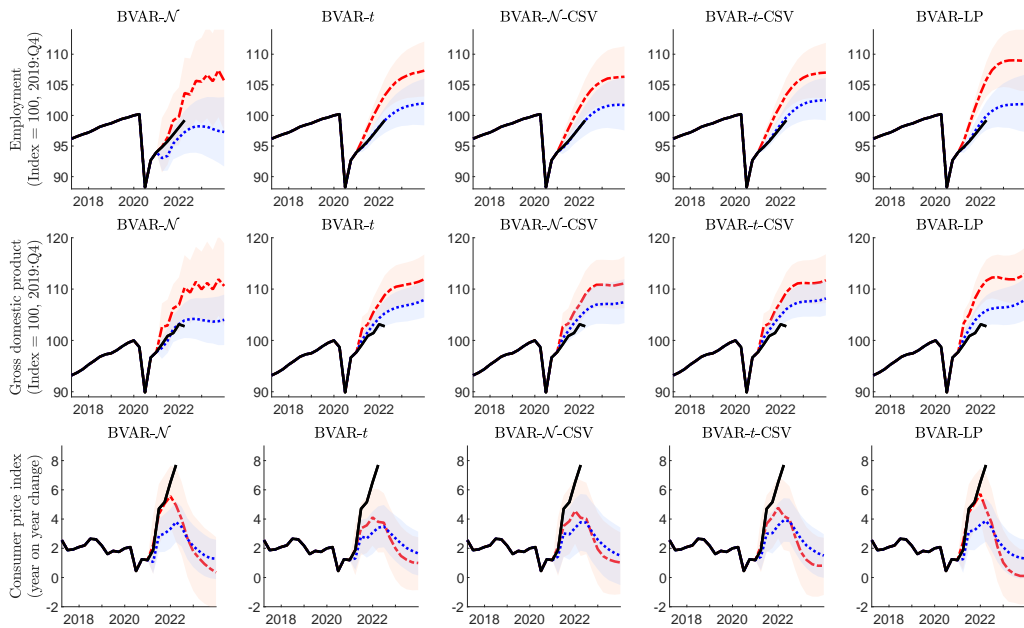
³⁴For unemployment rate, consumption and PCE core inflation, see [Figure A.6](#) in [Appendix A](#).

³⁵This exercise abstracts from real-time revisions and evaluates the FRED-QD vintage in June 2022.

Figure 7: Impact of re-calibrated priors on unconditional forecasts starting 2021:Q1



(a) Medium BVAR



(b) Large BVAR

--- MAD(q Reg) -.-.- RMSD(o Reg) — actual

Notes: BVARs with MAD(q Reg) and RMSD(o Reg) prior, respectively, estimated until 2020:Q4. The thick line is the posterior median and the colored area depicts the 16% to 84% credible interval.

calibrating the prior distribution during the pandemic. It shows that the robust BVARs coupled with a large cross-sectional information set is the group of superior forecast-

Table 7: Pseudo-out-of-sample forecast evaluation

	BVAR- \mathcal{N}		BVAR- t		BVAR- \mathcal{N} -CSV		BVAR- t -CSV		BVAR-LP		RW
	M	L	M	L	M	L	M	L	M	L	
(I) BVARs with MAD(q Reg) Minnesota prior											
Employment	8.78	6.36	5.18	4.25	5.21	4.24	5.28	4.27	5.32	4.16	4.34
Unemployment rate	7.30	4.65	3.96	3.12	3.88	3.10	4.01	3.14	4.14	3.09	3.45
Consumption	8.52	6.00	4.74	4.14	4.51	4.12	4.71	4.17	5.16	3.75	4.41
GDP	9.36	6.15	5.67	4.34	5.68	4.44	5.97	4.51	6.43	4.07	3.87
CPI inflation	1.18	1.21	1.17	0.94	1.24	1.05	1.20	0.95	1.19	1.25	1.30
PCE core inflation	0.76	0.90	0.83	0.74	0.84	0.79	0.95	0.79	0.76	0.70	0.74
(II) BVARs with RMSD(o Reg) Minnesota prior											
Employment	6.05	5.95	4.41	4.00	4.82	3.99	4.88	4.00	4.39	4.00	4.34
Unemployment rate	4.96	4.04	3.43	3.12	3.64	3.10	3.76	3.08	3.42	3.17	3.45
Consumption	6.18	5.98	4.35	4.01	4.37	3.99	4.57	4.02	4.31	3.79	4.41
GDP	6.37	5.67	4.46	3.76	4.77	3.74	4.98	3.85	4.57	3.45	3.87
CPI inflation	1.01	1.18	0.94	0.95	1.07	0.96	1.06	0.92	0.90	1.11	1.30
PCE core inflation	0.55	0.60	0.53	0.51	0.65	0.53	0.68	0.54	0.52	0.49	0.74

(a) One-quarter-ahead horizon

(I) BVARs with MAD(q Reg) Minnesota prior											
Employment	30.23	6.06	9.86	4.96	9.77	4.83	10.31	5.03	11.12	5.73	5.40
Unemployment rate	35.28	2.64	6.58	2.76	6.17	2.58	6.86	2.77	7.80	3.73	3.91
Consumption	28.80	2.96	8.45	3.34	7.74	3.50	8.60	3.58	9.63	3.61	7.46
GDP	30.45	4.68	8.97	4.68	8.69	4.59	9.37	4.88	10.87	6.78	5.64
CPI inflation	6.12	1.96	3.74	2.12	3.91	1.75	3.81	1.77	2.98	1.56	4.00
PCE core inflation	2.77	1.06	2.06	1.48	2.26	1.44	2.08	1.39	1.91	0.94	2.28
(II) BVARs with RMSD(o Reg) Minnesota prior											
Employment	18.77	15.43	5.78	4.27	7.72	4.15	8.37	4.29	5.84	4.08	5.40
Unemployment rate	17.27	7.19	3.72	2.53	4.90	2.47	5.56	2.58	3.75	2.95	3.91
Consumption	18.58	12.59	6.58	3.15	6.92	3.02	7.54	3.16	6.70	2.67	7.46
GDP	19.26	11.92	6.09	3.48	6.96	3.46	7.67	3.81	6.13	3.44	5.64
CPI inflation	3.92	3.25	3.41	2.88	3.45	2.54	3.47	2.50	2.92	2.35	4.00
PCE core inflation	2.09	1.68	2.01	1.58	2.00	1.55	1.90	1.48	1.89	1.41	2.28

(b) One-year-ahead horizon

Notes: RMSFE for unconditional forecasts based on medium (M) and large (L) BVARs with MAD(q Reg) and RMSD(o Reg) Minnesota prior estimated recursively from 2019:Q4 until 2021:Q4. RW denotes prediction from a random walk. Bold figures indicate the minimum RMSFE for each variable.

ing models. However, there is no clear superior forecasting approach among the robust BVARs as the predictive accuracy is broadly similar. Rather, the optimal BVAR depends on the variable and forecast horizon of interest as well as the chosen prior calibration. Nevertheless, the large BVAR-LP yields the most accurate forecast overall.

Though, this forecast evaluation is somewhat inconclusive about the best error extension and prior calibration for a BVAR during the pandemic period, it can still provide

some valuable lessons for forecasters. Going forward, the robust BVARs with robust prior calibration may have an edge over the standard method in forecasting both real and nominal variables at the current juncture. This is because the VAR forecasting information set does no longer include the very large shocks since the start of the pandemic, which led to an overshooting forecast with the robust prior. Therefore, it may become more important to exploit cross-sectional information rather than a-priori muting transmission of real variables in order to obtain accurate forecasts in the future.

4 Conclusion

Estimation of many standard macroeconomic models has become a challenge with the outbreak of the COVID-19 pandemic and calls for an appropriate treatment of these extreme observations. For a Bayesian VAR, this paper documents that not only is choosing an appropriate generalized error structure important, but also a careful calibration of the Minnesota prior matters for inference and forecasting purposes during the pandemic. Though model diagnostics prefer a combined error structure (heavy tails and time-varying volatility) and interpret the COVID-19 shock as a rare event, the choice among outlier-robust error structures becomes less important in forecasting when a large cross-section of information is used. Overall, the robust BVARs have a similar predictive accuracy during the pandemic, with the BVAR of [Lenza and Primiceri \(2022\)](#) being particularly promising.

Besides the error structure, this paper shows that the standard calibration method for the Minnesota prior is another important source of changing macroeconomic transmission channels during the pandemic as it mutes the propagation of real variables. This prior recalibration may be beneficial for predicting real variables at the beginning of the pandemic, yet it is less attractive when inflation forecasts are of interest. In fact, this strongly revised prior prevents the large VAR from capturing the mounting inflationary pressure in 2021. To provide a flexible and outlier-robust calibration, this paper proposes the MAD(q Reg) Minnesota prior as an alternative.

The off-the-shelf robust BVARs of [Chan \(2020\)](#) can be readily used during the pandemic and remain largely competitive with the newly developed BVAR of [Lenza and Primiceri \(2022\)](#). However, the short forecast evaluation suggests that a more explicit treatment of the pandemic observations has some merits. Thus, extending the generalized BVAR with a more flexible tail and volatility distribution might be a fruitful area of research. Going forward, the abruptly revised prior may become less attractive for forecasting both real and nominal variables at the current juncture as exploiting cross-sectional information may become more important than muting the internal propagation mechanism in a VAR. Furthermore, the proposed outlier-robust calibration method is not limited to the Minnesota prior, but may also be used for calibrating other scale-dependent prior distributions such as the sum of coefficient (co-integration) and sum of initial conditions prior ([Sims and Zha, 1998](#)), steady-state VAR prior ([Villani, 2009](#)) and long-run relations prior ([Giannone et al., 2019](#)).

References

- Antolín-Díaz, J., T. Drechsel, and I. Petrella (2020). Advances in Nowcasting Economic Activity: Secular Trends, Large Shocks and New Data. Working Paper DP15926, Centre for Economic Policy Research.
- Arias, J. E., J. F. Rubio-Ramírez, and M. Shin (2022). Macroeconomic Forecasting and Variable Ordering in Multivariate Stochastic Volatility Models. *Journal of Econometrics*.
- Bañbura, M., D. Giannone, and L. Reichlin (2010). Large Bayesian Vector Auto Regressions. *Journal of Applied Econometrics* 25(1), 71–92.
- Bobeica, E. and B. Hartwig (2023). The COVID-19 Shock and Challenges for Inflation Modelling. *International Journal of Forecasting* 39(1), 519–539.
- Carriero, A., T. E. Clark, and M. Marcellino (2015). Bayesian VARs: Specification Choices and Forecast Accuracy. *Journal of Applied Econometrics* 30(1), 46–73.
- Carriero, A., T. E. Clark, and M. Marcellino (2016). Common Drifting Volatility in Large Bayesian VARs. *Journal of Business & Economic Statistics* 34(3), 375–390.
- Carriero, A., T. E. Clark, M. Marcellino, and E. Mertens (2022). Addressing COVID-19 Outliers in BVARs with Stochastic Volatility. *The Review of Economics and Statistics*, 1–38.
- Chan, J. C. C. (2017). Notes on Bayesian Macroeconometrics. Unpublished, University of Technology Sydney.
- Chan, J. C. C. (2020). Large Bayesian VARs: A Flexible Kronecker Error Covariance Structure. *Journal of Business & Economic Statistics* 38(1), 68–79.
- Chib, S. (1995). Marginal Likelihood from the Gibbs Output. *Journal of the American Statistical Association* 90(432), 1313–1321.
- Chib, S. and S. Ramamurthy (2014). DSGE Models with Student-t Errors. *Econometric Reviews* 33(1-4), 152–171.
- Chiu, C.-W. J., H. Mumtaz, and G. Pinter (2017). Forecasting with VAR Models: Fat Tails and Stochastic Volatility. *International Journal of Forecasting* 33(4), 1124–1143.
- Christiano, L. J. (2007). On the Fit of New Keynesian Models: Comment. *Journal of Business & Economic Statistics* 25(2), 143–151.
- Clark, T. E. and F. Ravazzolo (2015). Macroeconomic Forecasting Performance under Alternative Specifications of Time-Varying Volatility. *Journal of Applied Econometrics* 30(4), 551–575.
- Cogley, T. and T. J. Sargent (2005). Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S. *Review of Economic Dynamics* 8(2), 262–302.

- Cúrdia, V., M. Del Negro, and D. L. Greenwald (2014). Rare Shocks, Great Recessions. *Journal of Applied Econometrics* 29(7), 1031–1052.
- Del Negro, M., M. Lenza, G. E. Primiceri, and A. Tambalotti (2020). What’s up with the Phillips Curve? Working Paper 27003, National Bureau of Economic Research.
- Eltoft, T., T. Kim, and T.-W. Lee (2006). On the Multivariate Laplace Distribution. *IEEE Signal Processing Letters* 13(5), 300–303.
- Geweke, J. (1993). Bayesian Treatment of the Independent Student- t Linear Model. *Journal of Applied Econometrics* 8(1), 19–40.
- Giannone, D., M. Lenza, and G. E. Primiceri (2015). Prior Selection for Vector Autoregressions. *The Review of Economics and Statistics* 97(2), 436–451.
- Giannone, D., M. Lenza, and G. E. Primiceri (2019). Priors for the Long Run. *Journal of the American Statistical Association* 114(526), 565–580.
- Hartwig, B. (2020). Robust Inference in Time-Varying Structural VAR Models: The DC-Cholesky Multivariate Stochastic Volatility Model. Working Paper 34/2020, Deutsche Bundesbank.
- Huber, F., G. Koop, L. Onorante, M. Pfarrhofer, and J. Schreiner (2023). Nowcasting in a Pandemic using Non-Parametric Mixed Frequency VARs. *Journal of Econometrics* 232(1), 52–69.
- Jacquier, E., N. G. Polson, and P. E. Rossi (2004). Bayesian Analysis of Stochastic Volatility Models with Fat-Tails and Correlated Errors. *Journal of Econometrics* 122(1), 185–212.
- Karlsson, S. (2013). "Chapter 15 - Forecasting with Bayesian Vector Autoregression". In G. Elliott and A. Timmermann (Eds.), *Handbook of Economic Forecasting*, Volume 2 of *Handbook of Economic Forecasting*, pp. 791–897. Elsevier.
- Koenker, R. and G. Bassett (1978). Regression Quantiles. *Econometrica* 46(1), 33–50.
- Lenza, M. and G. E. Primiceri (2022). How to Estimate a Vector Autoregression after March 2020. *Journal of Applied Econometrics* 37(7), 688–699.
- Litterman, R. B. (1986). Forecasting with Bayesian Vector Autoregressions: Five Years of Experience. *Journal of Business & Economic Statistics* 4(1), 25–38.
- McCracken, M. and S. Ng (2020). FRED-QD: A Quarterly Database for Macroeconomic Research. Working Paper 26872, National Bureau of Economic Research.
- Ng, S. (2021). Modeling Macroeconomic Variations After COVID-19. Working Paper 29060, National Bureau of Economic Research.
- Rousseeuw, P. J. and C. Croux (1993). Alternatives to the Median Absolute Deviation. *Journal of the American Statistical Association* 88(424), 1273–1283.

- Schorfheide, F. and D. Song (2021). Real-Time Forecasting with a (Standard) Mixed-Frequency VAR during a Pandemic. Working Paper 29535, National Bureau of Economic Research.
- Sims, C. A. and T. Zha (1998). Bayesian Methods for Dynamic Multivariate Models. *International Economic Review* 39(4), 949–968.
- Stock, J. H. and M. W. Watson (2016). Core Inflation and Trend Inflation. *The Review of Economics and Statistics* 98(4), 770–784.
- Villani, M. (2009). Steady-State Priors for Vector Autoregressions. *Journal of Applied Econometrics* 24(4), 630–650.

A Supplementary Figures and Tables

A.1 Model and prior comparison

Table A.1 presents log Bayes factors against the BVAR- \mathcal{N} with weakly informative prior before and during the pandemic for the large BVAR. Qualitatively, the results between the medium and large BVAR are similar except that the large BVAR with weakly informative prior always remains the worst performing model and that the prior revision has a more adverse impact on the model fit in the large BVAR model.

Table A.1: Bayes factor in large BVAR

Prior	BVAR- \mathcal{N}	BVAR- t	BVAR- \mathcal{N} -CSV	BVAR- t -CSV
Weak	•	198.29 (2.25)	205.36 (3.07)	216.67 (3.54)
RMSD (<i>oReg</i>)	1178.44 –	1217.21 (1.30)	1267.03 (2.29)	1269.17 (2.19)
MAD (<i>oReg</i>)	1175.93 –	1216.40 (1.02)	1264.39 (2.24)	1270.89 (2.02)
S_n (<i>oReg</i>)	1172.42 –	1214.99 (1.66)	1263.23 (2.50)	1270.24 (2.08)
Q_n (<i>oReg</i>)	1176.11 –	1214.31 (1.68)	1265.80 (1.67)	1272.38 (2.36)
RMSD (<i>FD</i>)	1146.84 –	1189.97 (1.48)	1236.19 (1.89)	1241.08 (2.32)
MAD (<i>FD</i>)	1161.47 –	1203.05 (1.25)	1249.15 (2.97)	1261.98 (1.71)
S_n (<i>FD</i>)	1167.26 –	1205.36 (1.32)	1261.16 (1.12)	1264.26 (1.73)
Q_n (<i>FD</i>)	1169.25 –	1212.99 (1.36)	1261.56 (2.20)	1265.09 (2.93)
RMSD (<i>qReg</i>)	1177.79 –	1217.34 (1.10)	1265.58 (2.02)	1274.09 (2.36)
MAD (<i>qReg</i>)	1177.47 –	1220.36 (1.29)	1265.08 (2.12)	1272.01 (2.25)
S_n (<i>qReg</i>)	1177.14 –	1218.57 (1.16)	1267.98 (1.99)	1271.38 (2.28)
Q_n (<i>qReg</i>)	1177.71 –	1217.59 (1.52)	1265.42 (3.02)	1272.45 (2.62)

(a) Estimation sample until 2019:Q4

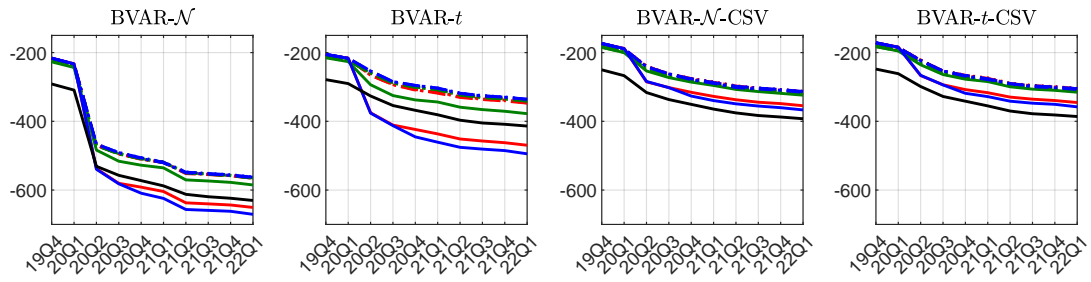
Weak	•	497.97 (3.91)	481.32 (4.34)	506.45 (4.34)
RMSD (<i>oReg</i>)	1109.22 –	1367.84 (1.66)	1435.18 (1.99)	1460.28 (1.99)
MAD (<i>oReg</i>)	1245.97 –	1545.50 (2.30)	1569.85 (2.56)	1589.37 (2.56)
S_n (<i>oReg</i>)	1246.63 –	1542.55 (1.92)	1569.78 (2.46)	1594.58 (2.46)
Q_n (<i>oReg</i>)	1246.08 –	1545.79 (1.35)	1568.20 (2.13)	1594.31 (2.13)
RMSD (<i>FD</i>)	1154.33 –	1437.05 (1.54)	1476.21 (1.61)	1502.49 (1.61)
MAD (<i>FD</i>)	1234.44 –	1537.83 (1.65)	1556.48 (1.78)	1578.09 (1.78)
S_n (<i>FD</i>)	1240.72 –	1542.11 (1.42)	1557.76 (3.02)	1583.86 (3.02)
Q_n (<i>FD</i>)	1242.10 –	1542.45 (1.86)	1562.00 (2.43)	1589.79 (2.43)
RMSD (<i>qReg</i>)	1090.49 –	1343.07 (2.05)	1413.99 (2.11)	1436.77 (2.11)
MAD (<i>qReg</i>)	1233.41 –	1539.69 (1.17)	1551.06 (3.05)	1575.71 (3.05)
S_n (<i>qReg</i>)	1240.31 –	1541.77 (1.64)	1560.06 (2.72)	1583.07 (2.72)
Q_n (<i>qReg</i>)	1240.56 –	1542.71 (1.47)	1564.44 (1.98)	1585.98 (1.98)

(b) Estimation sample until 2022:Q1

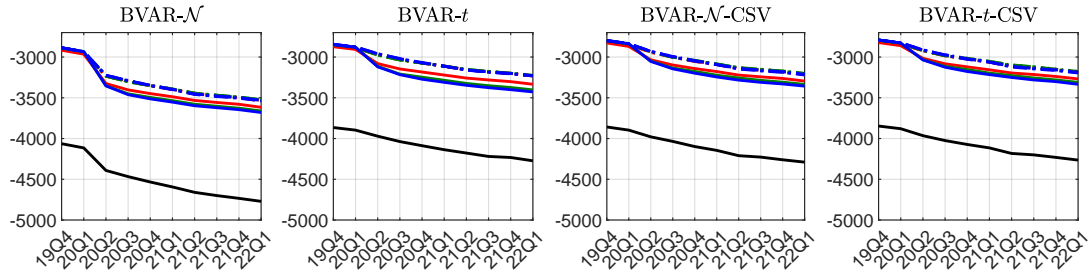
Notes: Log Bayes factors of various BVARs and prior combinations against BVAR- \mathcal{N} with weakly informative prior. Bold figures indicate maximum Bayes factor for each prior specification. Brackets report the numerical standard error.

Complementary to the tables reporting Bayes factors, Figure A.1 shows the log marginal likelihood when the models are estimated over an expanding window. The figure documents that the prior revision induced by extreme observation may worsen overall model performance, especially so in the large BVAR model.

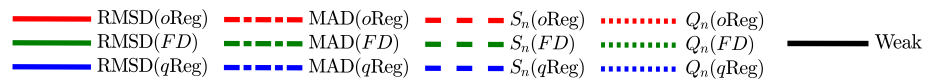
Figure A.1: Marginal likelihood with re-calibrated prior distribution



(a) Medium BVAR



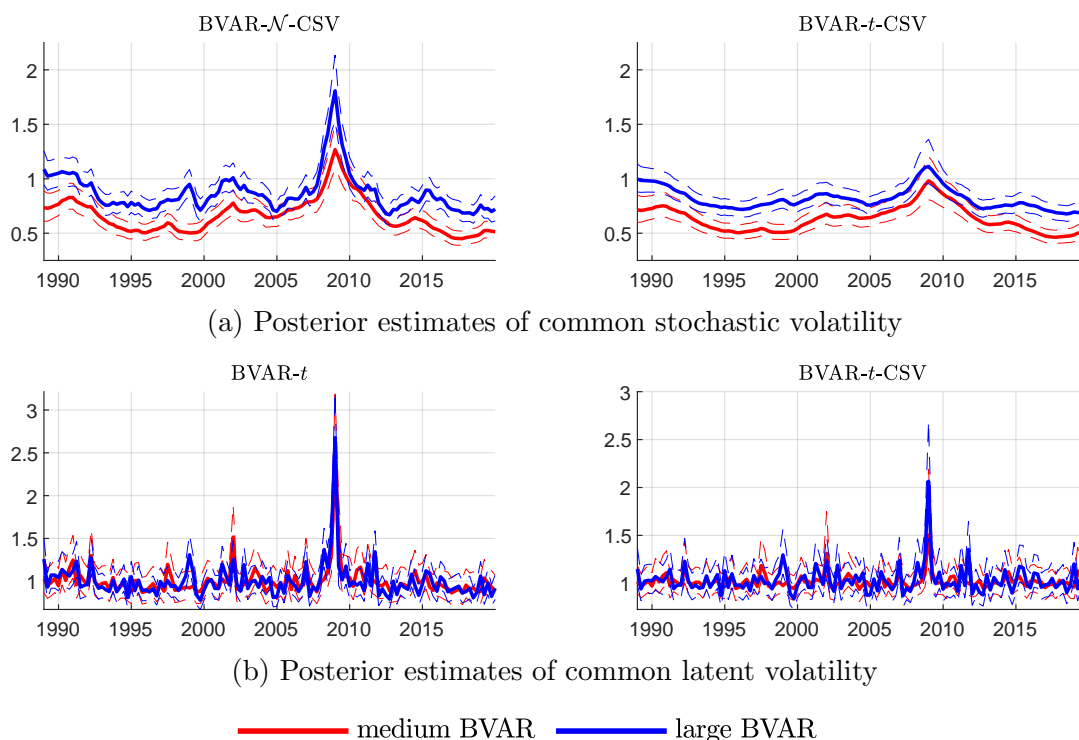
(b) Large BVAR



Notes: Log marginal likelihood of BVARs with re-calibrated prior distribution estimated over an expanding sample from 2019:Q4 to 2022:Q1.

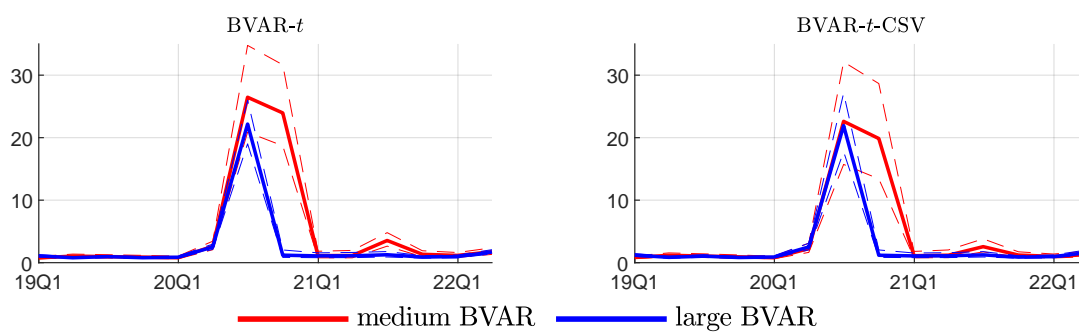
A.2 Macroeconomic tail risk and volatility

Figure A.2: Macroeconomic volatility before the pandemic



Notes: BVARs with MAD(q Reg) Minnesota prior estimated until 2019:Q4. The thick line is the posterior median and the dashed lines are the 16% to 84% credible interval.

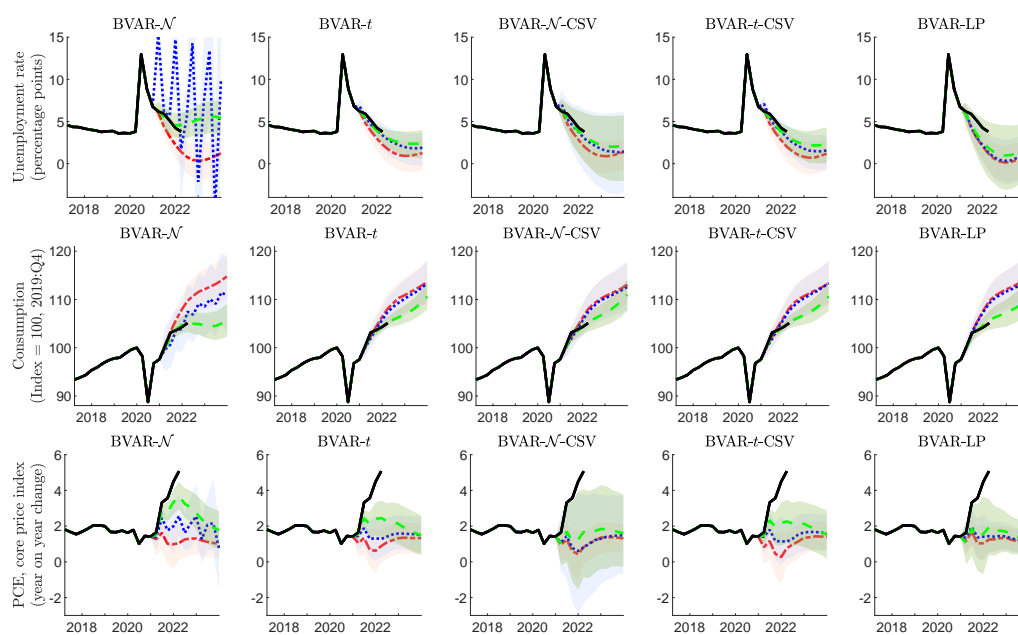
Figure A.3: Latent volatility during the pandemic



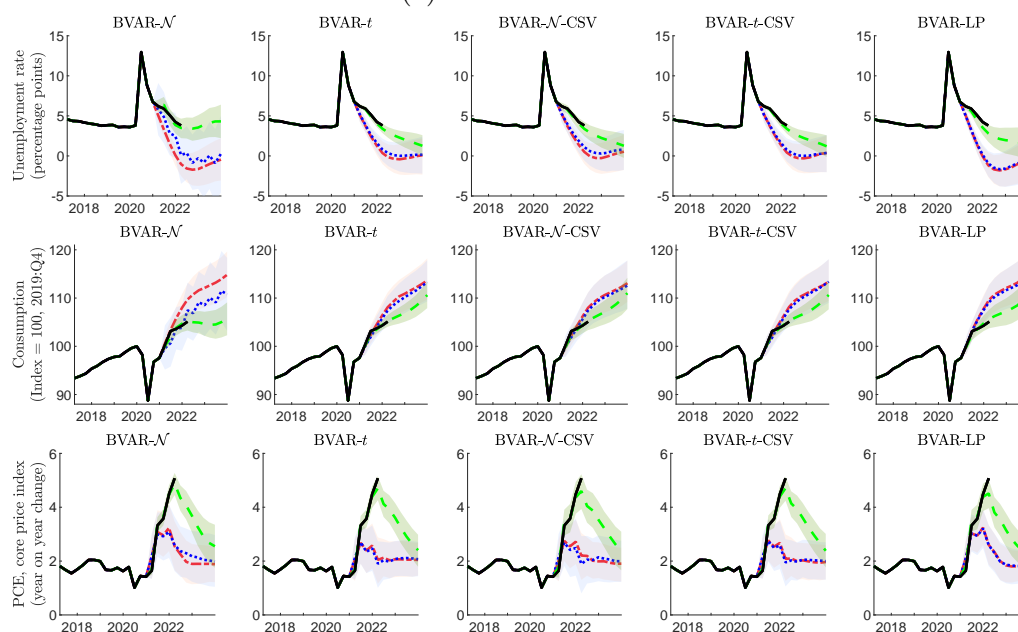
Notes: Posterior estimates of the Gaussian mixture component $\lambda_t/2$ of the t error model. BVARs with MAD(q Reg) Minnesota prior estimated until 2022:Q1. The thick line is the posterior median and the dashed lines are the 16% to 84% credible interval.

A.3 Model (in)stability and forecasting

Figure A.4: Impact of re-estimated parameters on unconditional forecasts starting 2021:Q1



(a) Medium BVAR



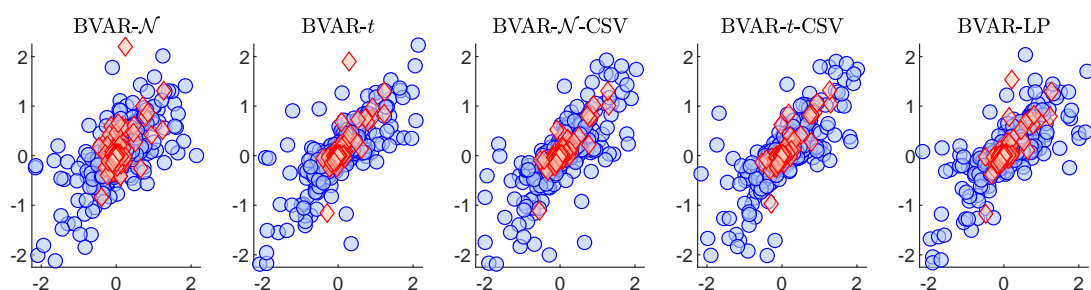
(b) Large BVAR

--- 19Q4 20Q4 - - - 22Q1 ——— actual

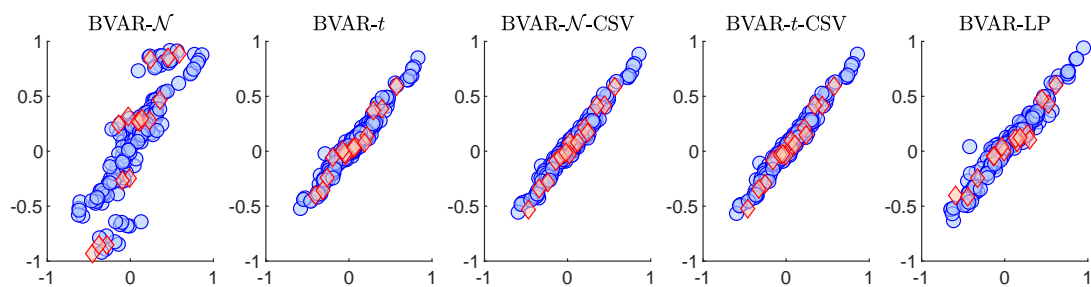
Notes: BVARs with MAD(q Reg) Minnesota prior estimated until 2019:Q4, 2020:Q4 and 2022:Q1, respectively. The thick line is the posterior median and the colored area depicts the 16% to 84% credible interval.

Figure A.5 shows scatter plots of the posterior mean VAR coefficients and residual correlation matrix obtained in 2019:Q4 against 2022:Q1 with the RMSD(*o*Reg) Minnesota prior calibration. The figure shows that the VAR coefficients are heavily revised in all BVAR models, whereas the residual correlation structure of the robust BVARs remains largely unchanged. Recall, the VAR coefficients in the robust BVARs with the MAD(*q*Reg) Minnesota prior calibration are largely resilient against the COVID-19 observations. Therefore, the COVID-19 observations coupled with RMSD(*o*Reg) Minnesota prior induce a marked revision of the VAR structure through an altered prior by shrinking the VAR higher-order and cross-dynamic lags of most real variables more strongly to zero.

Figure A.5: Change in common VAR parameters



(a) Posterior mean estimates of VAR coefficients

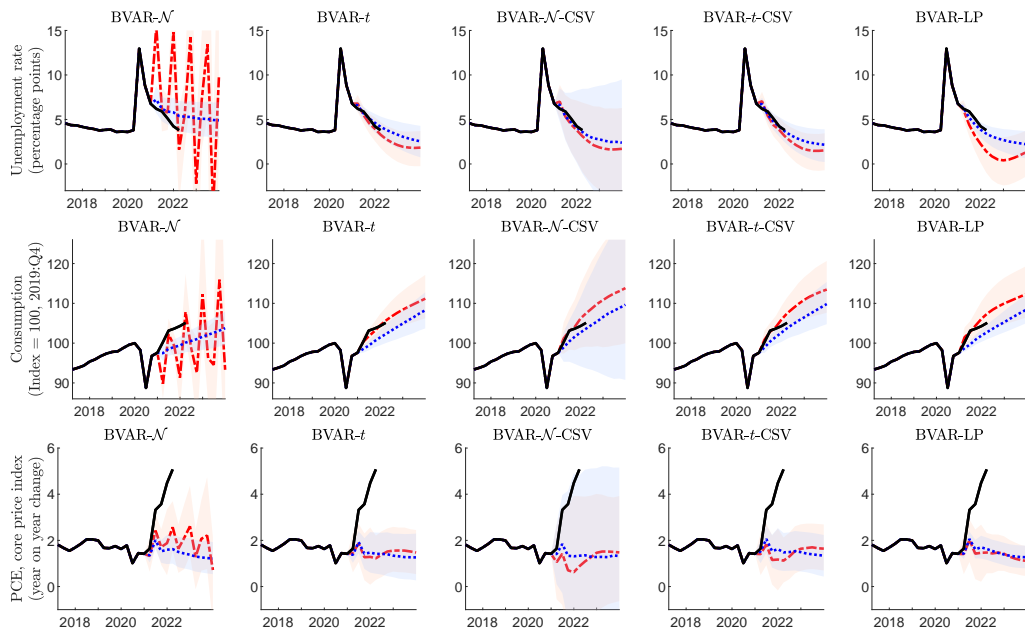


(b) Posterior mean estimates of VAR residual correlations

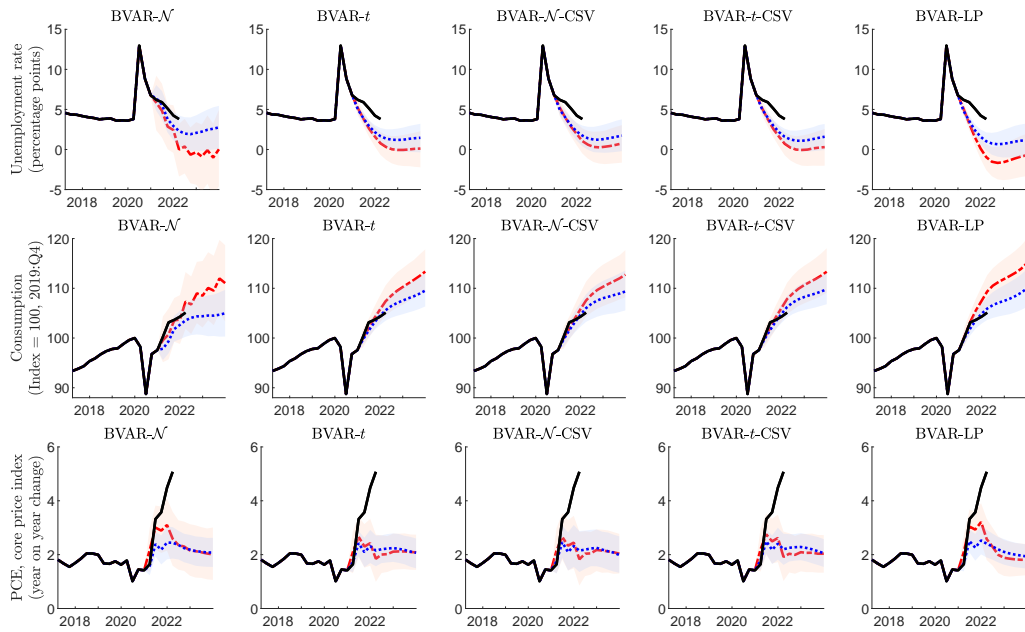
◇ medium BVAR ○ large BVAR

Notes: Common VAR parameters in 2019:Q4 (x-axis) against 2022:Q1 (y-axis) of various BVARs with RMSD(*o*Reg) Minnesota prior.

Figure A.6: Impact of re-calibrated priors on unconditional forecasts starting 2021:Q1



(a) Medium BVAR



(b) Large BVAR

--- MAD(q Reg) RMSD(o Reg) — actual

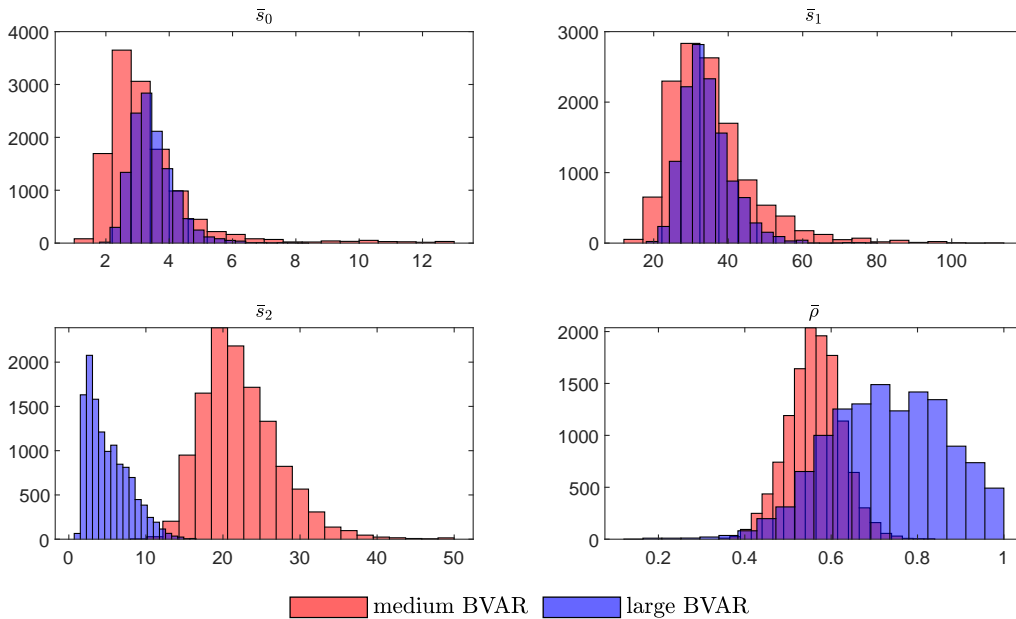
Notes: BVARs with MAD(q Reg) and RMSD(o Reg) Minnesota prior, respectively, estimated until 2020:Q4. The thick line is the posterior median and the colored area depicts the 16% to 84% credible interval.

B Estimation details of the BVAR-LP

The BVAR of [Lenza and Primiceri \(2022\)](#) is estimated on the sample from 1988:Q4 to 2022:Q1 and 2020:Q1 (t^*) is chosen as the start of the pandemic episode in which volatility is scaled by a common factor s_{t^*} , otherwise s_t is equal to one. Specifically, $s_{t^*} = \bar{s}_0$, $s_{t^*+1} = \bar{s}_1$, $s_{t^*+2} = \bar{s}_2$, and $s_{t^*+j} = 1 + (\bar{s}_2 - 1)\bar{\rho}^{j-2}$. For the VAR coefficients, the MAD(q Reg) Minnesota calibration is used and the hyperparameter $\kappa_{1,m}$ is fixed at its optimal pre-pandemic value and not re-optimized using the method of [Giannone et al. \(2015\)](#) to ensure comparability with the other BVAR models.

Figure B.1 shows the posterior distribution of the explicit volatility hyperparameters in both BVAR models. Overall, the estimated volatility burst in 2020:Q2 is about 35 standard deviations in both BVARs, but the subsequent evolution of the volatility process differs across VAR dimensions. The medium BVAR indicates a persistent increase in volatility for about two years while the large model views the COVID-19 shock as a sudden burst in 2020:Q2 and moderation afterwards in 2020:Q3. Furthermore, the posterior of $\bar{\rho}$ implies that the volatility process in the medium BVAR is less persistent than in the large model. The impact of persistence on the explicit volatility process is however overshadowed by the different starting points in 2020:Q3, see posterior of \bar{s}_2 .

Figure B.1: Posterior of volatility hyperparameters in the BVAR-LP

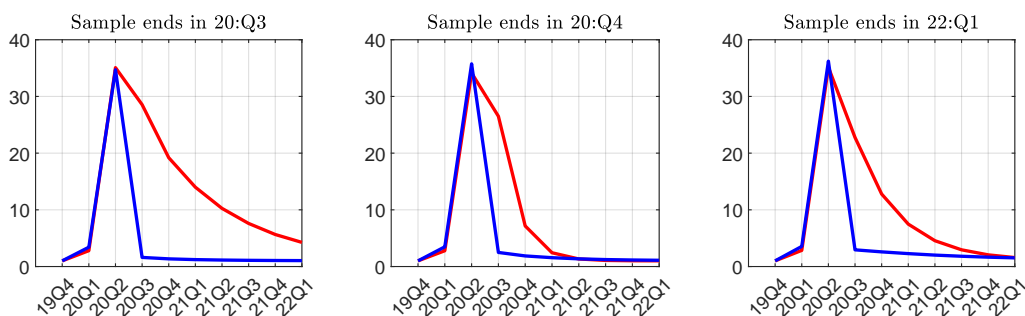


Notes: Posterior distributions of volatility scaling factors for 2020:Q1 (\bar{s}_0), 2020:Q2 (\bar{s}_1), 2020:Q3 (\bar{s}_2) and decay factor $\bar{\rho}$ of BVAR-LP with MAD(q Reg) Minnesota prior estimated until 2022:Q1.

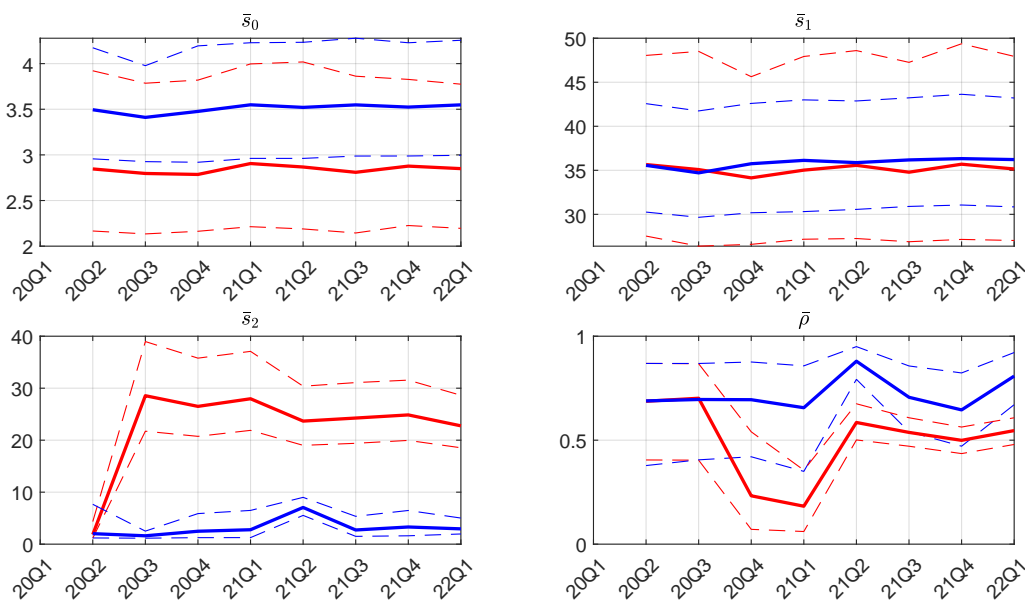
The data informs the large BVAR-LP very quickly about the nature of the COVID-19 shock while the estimate of the medium BVAR is revised several times before becoming more stable in the second half of 2021, see Figure B.2 (a) and (b). Therefore, explicit volatility hinges on the data set, being more persistent when relatively more time se-

ries that exhibit abnormal variation during the COVID-19 period are included. It also demonstrates that a larger information set helps to sharpen the identification.

Figure B.2: Revision in explicit volatility estimates in BVAR-LP



(a) Posterior estimates of explicit volatility for selected samples



(b) Posterior estimates of volatility hyperparameters over expanding sample

— medium BVAR — large BVAR

Notes: BVAR-LP with MAD(q Reg) Minnesota prior estimated recursively over expanding sample from 2020:Q2 to 2022:Q1. The thick line is the posterior median and the dashed lines are the 16% to 84% credible interval.