Real-Time Forecasting with a (Standard) Mixed-Frequency VAR During a Pandemic*

Frank Schorfheide and Dongho Song
[a]University of Pennsylvania, CEPR, NBER, PIER
[b]Johns Hopkins University, Carey Business School

We resuscitated the mixed-frequency vector autoregression (MF-VAR) developed in Schorfheide and Song (2015) to generate macroeconomic forecasts for the United States during the COVID-19 pandemic in real time. The model combines 11 time series observed at two frequencies: quarterly and monthly. We deliberately did not modify the model specification in view of the COVID-19 outbreak, except for the exclusion of crisis observations from the estimation sample. We compare the MF-VAR forecasts to the median forecast from the Survey of Professional Forecasters (SPF). While the MF-VAR performed poorly during 2020:Q2, subsequent forecasts were at par with the SPF forecasts. We show that excluding a few months of extreme observations is a promising way of handling VAR estimation going forward, as an alternative of a sophisticated modeling of outliers.

JEL Codes: C11, C32, C53.

1. Introduction

Vector autoregressions (VARs) are widely used in empirical macroeconomics. A VAR is a multivariate time-series model that can be used to forecast individual time series, to predict co-movements of

*We thank Frank Diebold, Elmar Mertens, Michele Lenza, Giorgio Primiceri, Minchul Shin, seminar participants at the European Central Bank, the editor, and two anonymous referees for helpful comments. Schorfheide gratefully acknowledges financial support from the National Science Foundation under Grant SES 1851634. Correspondence: F. Schorfheide: Department of Economics, 133 S. 36th Street, University of Pennsylvania, Philadelphia, PA 19104-6297. E-mail: schorf@ssc.upenn.edu. D. Song: Johns Hopkins Carey Business School, 100 International Drive, Baltimore, MD 21202. E-mail: dongho.song@jhu.edu.

macroeconomic or financial variables, to analyze sources of business cycle fluctuations, or to assess the effects of monetary or fiscal policy interventions on the macro economy. The recent COVID-19 pandemic triggered long-lasting mobility restrictions in the form of stay-at-home orders across the United States and the world in 2020 and beyond. As a consequence, economic activity collapsed in many sectors and unemployment soared. The unprecedented decline of economic activity created a tremendous challenge for macroeconomic modeling and forecasting, including the use of VARs and related state-space models.

In response to this challenge we resuscitated the mixed-frequency VAR, henceforth MF-VAR, developed in Schorfheide and Song (2015). MF-VARs have been shown to be competitive with other forecasting approaches, including surveys of professional forecasters, by, for instance, Schorfheide and Song (2015), Brave, Butters, and Justiniano (2019), and McCracken, Owyang, and Sekhposyan (2020). They are preferable, in particular at shorter forecast horizons, to single-frequency VARs that use time-aggregated data. Rather than modifying the MF-VAR in real time or ex post to accommodate idiosyncrasies of the economic downturn triggered by the COVID-19 pandemic, we decided to leave the model unchanged, except for the consideration of several forms of excluding or discounting extreme observations during the estimation stage. We summarized a first set of real-time forecasts in the working paper version Schorfheide and Song (2020) and subsequently published monthly real-time forecasts at www.donghosong.com from April 30, 2020 until August 31, 2021.

The contribution of the current paper is to evaluate the 17 months of real-time MF-VAR forecasts and compare them to median forecasts from the Survey of Professional Forecasters (SPF) conducted by the Federal Reserve Bank of Philadelphia.[1] The paper focuses on the effect of outliers in the context of an MF-VAR, rather than on the advantages of a modeling strategy that mixes time series sampled at various frequencies. State-space models are more flexible in handling outliers than models in which all variables are treated as observed, such as standard VARs. We contrast two

[1]See https://www.philadelphiafed.org/surveys-and-data.

approaches of capturing outliers: inflating the scale of the measurement errors to capture observation outliers, and inflating the scale of the state-transition innovations to capture innovation outliers. The first approach downweights extreme observations without propagating them through the system. Lagged values of extreme observations are effectively replaced by forecasts from the state-transition equation. The second approach downweights the likelihood increment associated with extreme observations, but uses them as lagged values to predict subsequent "normal" observations. We use the first approach for parameter estimation and the second approach to generate forecasts conditional on parameter draws from the posterior distribution. Letting the scale tend to infinity is equivalent to dropping observations from the system.

Instead of modeling the occurrence of outliers through a fat-tailed distribution or a mixture distribution, we simply scale the innovation variances or drop observations based on an inspection of the data prior to the MF-VAR estimation. This approach is convenient and practical in situations in which observations are so extreme that they are easily recognizable as outliers, such as the COVID-19 pandemic. Our approach is also easy to implement on other model classes and less demanding in terms of time and human resources than the explicit modeling of outliers.

The SPF is a widely used benchmark in forecast comparisons. From our perspective, the SPF has two important features: first, as our MF-VAR forecasts, the SPF forecasts were made in real time as the pandemic was unfolding and they are not based on ex post model optimization. Second, the professional forecasters had the opportunity to make real-time judgmental adjustments to model-based forecasts in view of the evolving pandemic which cannot be reproduced in pseudo-out-of-sample forecast comparisons with other model classes. Our sample is too short for a formal comparison of forecasts to be informative. Numerical evaluation statistics are likely to be dominated by a few large forecast errors. Instead, we focus on visual comparisons of ex post realizations and forecast paths generated at different origins.

We draw three main conclusions from our analysis. First, during the first three months of the pandemic in the United States, in 2020:Q2, we deliberately did not exert any effort in adapting our MF-VAR to the idiosyncrasies of the pandemic, except for ending

the estimation sample on January 31, 2020, nor did we make judgmental adjustments to the model-based forecasts. We find that the MF-VAR forecasts are substantially worse than the SPF forecasts during this period. In April 2020 the data for the MF-VAR estimation did not yet contain information about the severity of the downturn and the model did not anticipate the magnitude of the recession. By June 2020, the model had identified large shocks to the economy. Propagating these large shocks through a very persistent VAR law of motion led to an overly pessimistic forecast. In short, the model performed poorly in 2020:Q2.

Second, without any modifications or adjustments the MF-VAR model generated remarkably accurate forecasts from July 2020 onward. These forecasts are at par with the median SPF forecasts. Many pundits initially expected the pandemic to be relatively short-lived and the recession to be followed by a strong recovery once the mobility restrictions were lifted. From this perspective, the cards were stacked ex ante against the MF-VAR, which is estimated based on macroeconomic time series that exhibit unit-root behavior and thereby implies that shocks tend to have long-lived effects. Ex post it turned out that mobility restrictions could only be lifted gradually and that the economic effects of the pandemic were long-lasting, just as the effects of previous recessionary shocks had been long-lasting and recoveries have often been slow.

Third, going forward, an important question for users of VARs is how to handle the extreme data points observed in the second quarter of 2020. One option is to increase the complexity of the VAR model by explicitly allowing for outliers, either specifically during the COVID-19 pandemic or in every period with some small probability. Our findings suggest that the alternative and rather simple approach of excluding observations from the first few months of the pandemic works remarkably well and provides an attractive alternative in situations in which a more sophisticated modeling of outliers is impractical. Our approach ensures that the MF-VAR performs as well after the initial downturn as it did prior to the pandemic. The real-time forecasts published at www.donghosong.com were generated by ending the estimation sample on January 31, 2020. As time has progressed, it has become desirable to start including new observations in the estimation sample. Based on our findings, we recommend dropping observations from March to June

2020 but including the subsequent data points when estimating a VAR.

Since the beginning of the pandemic, many papers have been written contemporaneously on how to adjust forecast models to cope with the unprecedented economic downturn. The two papers most closely related to our work are Lenza and Primiceri (2022), henceforth LP, and Carriero et al. (2022), henceforth CCMM. Just as in our paper, CCMM and LP propose adjustments to a baseline VAR specification. While our baseline model features mixed-frequency observations, LP use a homoskedastic VAR as a starting point, and CCMM start from a VAR with stochastic volatility (SV).

LP propose to deterministically scale the innovation covariance for the duration of the pandemic to capture the increased shock sizes. Specifically, the authors recommend estimating separate scale factors for the first months of the pandemic and then letting the last of these scale factors decay geometrically at an estimated rate. To the extent that the estimated scale factors are large, the method has a similar effect on the parameter estimates as dropping observations. However, rather than abruptly including new observations after a certain period, the method gradually increases the weight of these observations.

While VARs with SV are designed to adapt to time-varying volatility and have shown to improve density and interval forecasts (see, for instance, Clark 2011), the estimated volatility processes are typically highly persistent. This implies that large COVID-19 shocks over two to three months would raise SV for multiple years to levels that are ex ante implausible and ex post counterfactual. As a remedy, CCMM modify the SV specification to allow for Student-t distributed (instead of Gaussian) innovations as well as outliers that do not trigger a persistent increase in volatility. The authors show that the outlier-augmented SV-t specification substantially improves the forecast performance of a standard VAR with SV.

Our approach of essentially dropping observation after a casual inspection of the data and CCMM's approach of explicit outlier modeling can be viewed as the two endpoints of a continuum of empirical strategies. LP's approach of inflating the error variance at a pre-specified point in time and then estimating a decay rate lies in between the two endpoints. While the CCMM approach can handle outliers in an automated way and potentially adapt to future outliers

caused by non-COVID-related economic disruptions, our approach provides a low-tech alternative in situations in which a more sophisticated modeling is impractical. We adapt the Lenza and Primiceri (2022) approach to our mixed-frequency framework. We find in our application that, compared to our approach of excluding observations, the LP approach creates similar point forecasts but ex post unreasonably large predictive intervals because the estimated scale factor implies a relatively slow decay.

Alvarez and Odendahl (2021) use an outlier modeling approach similar to CCMM, but with the feature that the reduced-form errors rather than the structural errors are stochastically rescaled, so that the contemporaneous correlation of the reduced-form errors is preserved. Antolin-Diaz, Drechsel, and Petrella (2021) and Bobeica and Hartwig (2022) use fat-tailed error distributions to discount extreme observations—the former in the context of a dynamic factor model (DFM) and the latter in a VAR. Foroni, Marcellino, and Stevanovic (2020) explore COVID-19 adjustments of econometric model forecasts that are based on the forecasting experience during the Great Recession. Either the forecast model is exclusively estimated based on observations surrounding the Great Recession (similarity-based estimation) or its forecasts are corrected by forecast errors made during the Great Recession (intercept correction).

We are not alone in exploring the behavior of a time-series models during the pandemic without explicitly modeling outliers or tailoring the model specification to the COVID observations. For instance, Diebold (2020) studies the performance of the Aruoba-Diebold-Scotti (ADS) economic activity index, which is based on a low-dimensional dynamic factor model and has been published by the Federal Reserve Bank of Philadelphia for more than a decade. Lewis, Mertens, and Stock (2020) developed a weekly economic index (WEI) to track the rapid economic developments triggered by the coronavirus pandemic. Their principal component analysis, which uses observations from 2008 onward, does not treat the observations from the second quarter of 2020 as outliers.

Ng (2021) uses COVID indicators, such as the number of current documented infections, and the number of hospitalizations and deaths, either as exogenous controls or as endogenous variables in VARs regressions to "de-COVID" the data so that economic factors and shocks can be identified, or as additional predictors to

account for the persistent nature of COVID. Because the COVID indicators are zero before the pandemic, their behavior is very non-Gaussian. Davis and Ng (2022) develop methods for the estimation of multivariate models with heavy-tailed and thin-tailed variables, using independent component analysis to identify disaster/pandemic shocks. While all of the previous approaches treat the pandemic as large unobserved shocks that propagate through the dynamic system, Primiceri and Tambalotti (2020) adapt a VAR to the COVID-19 pandemic by assuming that also the propagation (persistence and co-movements) of the COVID shocks is potentially different from typical business cycle shocks. While their approach is useful for scenario analysis, it is difficult to accurately estimate the COVID shock propagation mechanism in real time.

There is another strand of literature that has a slightly different objective, namely to introduce non-linearities into time-series models such as VARs or DFMs. A side benefit, though not the key modeling motivation, is that these models, without further adjustments, may be more robust to the occurrence of outliers than linear models. Huber et al. (2022) develop Bayesian econometric methods for posterior parametric mixed-frequency VARs using additive regression trees. Goulet Coulombe, Marcellino, and Stevanovic (2021) document that some ML methods capture non-linearities that can improve forecasts during the COVID-19 crisis. However, it is important to note that the same adjustments that we, LP, and CCMM make to a linear VAR could be made to the non-linear time-series models.[2]

The remainder of this paper is organized as follows. In Section 2 we provide an example that illustrates how outliers can be handled in a state-space model by robustifying either the measurement equation or the state-transition equation. Section 3 reviews the specification of the MF-VAR and discusses how we drop observations from the estimation sample and implement the LP approach of scaling the

[2]This includes growth-at-risk models building on Adrian, Boyarchenko, and Giannone (2019). While the assessment of tail risk is particularly important at the onset of the COVID-19 pandemic, during the pandemic and in its aftermath the same question remains: does including the pandemic observations distort the recursive estimates of these models?

innovation covariance matrix. The real-time data set is discussed in Section 4 and the empirical results are presented in Section 5. Finally, Section 6 concludes. Additional information about the construction of our data set is provided in the appendix. Real-time forecasts from April 30, 2020 to August 31, 2021, were published and remain available at www.donghosong.com.

2. Outliers in State-Space Models—An Example

In order to examine the effect of outliers on filtering and parameter estimation in state-space models, consider the following example:

$$\begin{aligned}(ME) &: y_\tau = s_\tau + u_\tau, \quad u_\tau \sim N(0, \chi_{u,\tau}) \\ (ST) &: s_\tau = \theta_{\tau-1} s_{\tau-1} + \epsilon_\tau, \quad \epsilon_\tau \sim N(0, \chi_{\epsilon,\tau}), \quad \tau = 1, 2, \ldots, T.\end{aligned} \tag{1}$$

The observables y_τ, the latent states s_τ, and the unknown parameters $\theta_\tau \in \Theta$ are scalars. We assume that there is a pandemic in period $\tau = t$ and that the sample ends one period later in period $\tau = t + 1 = T$. In normal times

$$\chi_{u,\tau} = 1, \quad \chi_{\epsilon,\tau} = 1, \quad \theta_\tau = \theta$$

and the observations are generated from (1). During the pandemic period t, nature replaces the innovations (u_t, ϵ_t) with the values $(\tilde{u}_t, \tilde{\epsilon}_t)$ to determine (y_t, s_t). Moreover, nature sets $\theta_t = \tilde{\theta}$ to generate (y_{t+1}, s_{t+1}) from (1). Following the taxonomy in Gandhi and Mili (2010), we refer to $\tilde{u}_t$ as observation outlier, $\tilde{\epsilon}_t$ as innovation outlier, and $\tilde{\theta}_t$ as structural outlier.

The econometrician uses (1) to make inference about the latent state s_τ, the parameter θ, and forecast y_τ. She considers robustifying the econometric model to the presence of outliers by inflating the scale of the error distributions in period t. Rather than considering fat-tailed distributions, this is done by simply switching the values of the constants $\chi_{u,t}$ and $\chi_{\epsilon,t}$ from one to a large value. We refer to an increase of $\chi_{u,\tau}$ as robustifying the measurement equation (ME). Similarly, an increase of $\chi_{\epsilon,t}$ is regarded as robustifying the state-transition (ST) equation. We compare the effect of both approaches on filtering and parameter estimation. If we set $\chi_{u,\tau} = 0$, then the model reduces to an AR(1).

Let $\chi_\tau = [\chi_{u,\tau}, \chi_{\epsilon,\tau}]'$. Suppose that the Kalman filter (KF) delivers the time $t-1$ state distribution $s_{t-1}|Y_{1;t-1} \sim N(s_{t-1|t-1}, P_{t-1|t-1})$. The predictive distribution for the time t observation is $y_t|Y_{1:t-1} \sim N\big(y_{t|t-1}, F_{t|t-1}(\chi_t)\big)$, where

$$y_{t|t-1} = \theta s_{t-1|t-1}, \quad F_{t|t-1}(\chi_t) = \theta^2 P_{t-1|t-1} + \chi_{\epsilon,t} + \chi_{u,t}. \tag{2}$$

Define

$$\lambda(P, \chi, \theta) = \left(\frac{\theta^2 P + \chi_\epsilon}{\theta^2 P + \chi_\epsilon + \chi_u}\right). \tag{3}$$

The updating step yields $s_t|Y_{1:t} \sim N\big(s_{t|t}(\chi_t), P_{t|t}(\chi_t)\big)$, where

$$\begin{aligned} s_{t|t}(\chi_t) &= \lambda(P_{t-1|t-1}, \chi_t, \theta) y_t + \big[1 - \lambda(P_{t-1|t-1}, \chi_t, \theta)\big]\theta s_{t-1|t-1} \\ P_{t|t}(\chi_t) &= \lambda(P_{t-1|t-1}, \chi_t, \theta)\chi_{u,t}. \end{aligned} \tag{4}$$

Thus, the filtered state in period t is a linear combination of y_t and the forecast $\theta s_{t-1|t-1}$, obtained by iterating ST one period forward. The weight $\lambda(\cdot)$ in (3) is a function of χ_t. Iterating the KF one more period forward, we obtain the following mean and variance for the predictive distribution of y_{t+1}:

$$y_{t+1|t}(\chi_t) = \theta s_{t|t}(\chi_t), \quad F_{t+1|t}(\chi_t) = \theta^2 P_{t|t}(\chi_t) + 2. \tag{5}$$

The updating step yields the moments

$$\begin{aligned} s_{t+1|t+1}(\chi_t) &= \lambda\big(P_{t|t}(\chi_t), 1, \theta\big) y_{t+1} + \big[1 - \lambda\big(P_{t|t}(\chi_t), 1, \theta\big)\big]\theta s_{t|t}(\chi_t) \\ P_{t+1|t+1}(\chi_t) &= \lambda\big(P_{t|t}(\chi_t), 1, \theta\big). \end{aligned} \tag{6}$$

Letting $\chi_{u,t}$ or $\chi_{\epsilon,t}$ tend to infinity can be viewed as an extreme way of robustifying the filtering, likelihood evaluation, and forecasting in the presence of the period t outliers. Maintaining that $\chi_\tau = 1$ for $\tau \neq t$, define the log-likelihood function as

Table 1. Limits of KF Moments

Moment	Robust ME $\chi_{u,t} \to \infty, \chi_{\epsilon,t} = 1$	Robust ST $\chi_{u,t} = 1, \chi_{\epsilon,t} \to \infty$
$y_{t\|t-1}$	$\theta s_{t-1\|t-1}$	$\theta s_{t-1\|t-1}$
$F_{t\|t-1}(\chi_t)$	∞	∞
$s_{t\|t}(\chi_t)$	$\theta s_{t-1\|t-1}$	y_t
$P_{t\|t}(\chi_t)$	$\theta^2 P_{t-1\|t-1} + 1$	1
$y_{t+1\|t}(\chi_t)$	$\theta^2 s_{t-1\|t-1}$	θy_t
$F_{t+1\|t}(\chi_t)$	$\theta^4 P_{t-1\|t-1} + \theta^2 + 2$	$\theta^2 + 2$
$s_{t+1\|t+1}(\chi_t)$	$\left(\frac{\theta^4 P_{t-1\|t-1}+\theta^2+1}{\theta^4 P_{t-1\|t-1}+\theta^2+2}\right) y_{t+1}$ $+ \left(\frac{1}{\theta^4 P_{t-1\|t-1}+\theta^2+2}\right)\theta^2 s_{t-1\|t-1}$	$\left(\frac{\theta^2+1}{\theta^2+2}\right) y_{t+1} + \left(\frac{1}{\theta^2+2}\right)\theta y_t$
$P_{t+1\|t+1}(\chi_t)$	$\frac{\theta^4 P_{t-1\|t-1}+\theta^2+1}{\theta^4 P_{t-1\|t-1}+\theta^2+2}$	$\frac{\theta^2+1}{\theta^2+2}$

$$\ln p(Y_{1:T}|\theta,\chi_t) = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{\tau=1}^{T}\left[\ln|F_{\tau|\tau-1}(\chi_t)| + \frac{(y_\tau - y_{\tau|\tau-1}(\chi_t))^2}{F_{\tau|\tau-1}(\chi_t)}\right], \tag{7}$$

with the understanding that $y_{\tau|\tau-1}(\chi_t)$ and $F_{\tau|\tau-1}(\chi_t)$ do not vary with χ_t for $\tau < t$.

Table 1 summarizes the $\chi_{u,t} \longrightarrow \infty$ (robust ME) and $\chi_{\epsilon,t} \longrightarrow \infty$ (robust ST) limits of the KF moments for periods t and $t+1$. The time t limits of the forecasts are identical. In both cases $y_{t|t-1} = \theta s_{t-1|t-1}$ and $F_{t|t-1} = \infty$, which implies that the log-likelihood increment for the outlier observation y_t drops out. While the time t likelihood increment diverges, the limit of any log-likelihood ratio $\ln p(Y_{1:T}|\theta,\chi_t) - \ln p(Y_{1:T}|\tilde{\theta},\chi_t)$ remains well defined. The robust ME and ST limits start to differ with the time t updating step. Notice from (3) that

$$\lim_{\chi_{u,t}\longrightarrow\infty} \lambda(P,\chi_t,\theta) = 0, \qquad \lim_{\chi_{u,t}\longrightarrow\infty} \lambda(P,\chi_t,\theta)\chi_{u,t} = \theta^2 P + \chi_{\epsilon,t},$$
$$\lim_{\chi_{\epsilon,t}\longrightarrow\infty} \lambda(P,\chi_t,\theta) = 1.$$

Under robust ME the time t filtered state is solely based on the forward iteration of the ST, $\theta s_{t-1|t-1}$. In turn, neither the time t forecast of y_{t+1} nor the filtered value of s_{t+1} is affected by the outlier y_t. Under robust ST, on the other hand, $y_{t+1|t}$ and $s_{t+1|t+1}$ are functions of the extreme observation y_t instead of the forward iteration $\theta s_{t-1|t-1}$.

In the empirical part of this paper we conduct Bayesian inference, which is based on the posterior distribution

$$p(\theta|Y_{1:T}, \chi_t) \propto p(Y_{1:T}|\theta, \chi_t)p(\theta), \tag{8}$$

where $\propto$ denotes proportionality and $p(\theta)$ is the prior. In large samples the log-likelihood function is approximately quadratic around the maximum likelihood estimator (MLE)

$$\hat{\theta}(\chi_t) = \text{argmax}_{\theta\in\Theta} \ \ln p(Y_{1:T}|\theta, \chi_t).$$

Denoting $\hat{V}(\chi_t)$ the negative inverse Hessian of the log-likelihood function evaluated at the posterior mode, we can approximate the posterior as

$$\begin{aligned} p(\theta|Y_{1:T}, \chi_t) \propto (2\pi)^{-1/2}|\hat{V}(\chi_t)|^{-1/2} \\ \times \exp\left\{-\frac{\left(\theta - \hat{\theta}(\chi_t)\right)^2}{2\hat{V}(\chi_t)} + \text{small}\right\} p(\theta). \end{aligned} \tag{9}$$

We proceed by examining the sensitivity of the MLE to y_t as a function of χ_t. Recall that $T = t + 1$ and let $Y_{(-t)} = (Y_{1:t-1}, y_{t+1})$. Then define

$$\ell(\theta, y_t; Y_{(-t)}, \chi_t) = \ln p(Y_{1:T}|\theta, \chi_t)$$

and its first- and second-order derivatives $\ell_\theta(\theta, y_t; \cdot)$, $\ell_y(\theta, y_t; \cdot)$, $\ell_{y\theta}(\theta, y_t; \cdot)$, $\ell_{\theta\theta}(\theta, y_t; \cdot)$. Assuming that the MLE lies in the interior, it satisfies the first-order condition $\ell_\theta(\hat{\theta}, y_t; Y_{(-t)}, \chi_t) = 0$. We deduce from the implicit function theorem that

$$\frac{\partial\hat{\theta}}{\partial y_t} = -\left[\ell_{\theta\theta}(\hat{\theta}, y_t; Y_{(-t)}, \chi_t)\right]^{-1}\ell_{\theta y}(\hat{\theta}, y_t; Y_{(-t)}, \chi_t). \tag{10}$$

We show in the appendix that

$$\lim_{\chi_{u,t}\longrightarrow\infty} \ell_{y\theta}(\hat{\theta}, y_t; Y_{(-t)}, \chi_t) = 0$$

$$\lim_{\chi_{\epsilon,t}\longrightarrow\infty} \ell_{y\theta}(\hat{\theta}, y_t; Y_{(-t)}, \chi_t) = -\frac{\hat{\theta} y_t}{\hat{\theta}^2 + 2} + \frac{y_{t+1} - \hat{\theta} y_t}{\hat{\theta}^2 + 2}\left(1 - 2\frac{\hat{\theta}^2}{\hat{\theta}^2 + 2}\right).$$

The observation y_t affects inference about θ in two ways. First, the model needs to explain the observation y_t based on $t-1$ information. Second, θ affects the error that the model makes predicting y_{t+1} based on information that includes the observation y_t. If we let $\chi_{u,t} \longrightarrow \infty$, then the influence of y_t on $\hat{\theta}$ is eliminated because the increment $\ln p(y_t|Y_{1:t-1}, \theta, \chi_t)$ is removed from the likelihood function. Moreover, as can be seen from Table 1, the forecast of y_{t+1} also no longer depends on y_t. It is well known in the state-space model literature that the $\chi_{u,t} \longrightarrow \infty$ limit corresponds to dropping time t observation from the measurement equation using the selection matrix (here just 1×1):

$$y_\tau = M_\tau(s_\tau + u_\tau), \quad M_\tau = \begin{cases} 1 & \text{for } \tau \neq t \\ \emptyset & \text{otherwise.} \end{cases}$$

On the other hand, if we let $\chi_{\epsilon,t} \longrightarrow \infty$, then y_t continues to affect the estimator $\hat{\theta}$ because it is used in the prediction of y_{t+1}. If after the pandemic, in period $t+1$, y_{t+1} is determined in the same way as before the pandemic, i.e., $\tilde{\theta}_t = \theta$, then a large value of y_t is very beneficial, because it generates, correctly, a lot of information about θ. If on the other hand, post-pandemic dynamics are somewhat different from pre-pandemic dynamics, i.e., $\tilde{\theta}_t \neq \theta$, then the information generated by y_t will distort inference about θ and forecasts for future periods.

In view of these considerations, we adopted the following strategy to generate the baseline real-time forecasts: we estimate θ using a Gibbs sampler with a simulation smoother that is based on the previously described Kalman filter algorithm. We set $M_\tau = \emptyset$ during part of the pandemic, which removes the influence of the extreme pandemic observations on the parameter estimation. This generates posterior draws θ^i, $i = 1, \ldots, N$. When generating out-of-sample forecasts, in the notation of the example, we bring back observation

y_t to forecast y_{t+1}. For each posterior draw i, we use the filter to compute draws from $s_{t+1}|(Y_{1:t}, \theta^i, \chi_{u,t} = 1, \chi_{\epsilon,t} = 1)$. The VAR-based forecasts of Carriero et al. (2022) and Lenza and Primiceri (2022) can be interpreted as setting $\chi_{u,\tau} = 0$ for all τ and $\chi_{\epsilon,t}$ equal to a large estimated value. Primiceri and Tambalotti (2020) try to estimate $\tilde{\theta}_t$.

3. MF-VAR Specification and Estimation

We consider an MF-VAR that utilizes monthly and quarterly observations. The MF-VAR can be conveniently represented as a state-space model, in which the state-transition equations are given by a VAR at monthly frequency and the measurement equations relate the observed series to the underlying, potentially unobserved, monthly variables that are stacked in the state vector. To cope with the high dimensionality of the parameter space, the MF-VAR is equipped with a Minnesota prior and estimated using Bayesian methods. In Section 3.1 we reproduce the model description and estimation strategy from Schorfheide and Song (2015), referring the reader to our original paper for a detailed discussion of the Bayesian computations. In Section 3.2 we discuss two modifications that we consider in the empirical application: (i) dropping of observations and (ii) a break in volatility as in LP.

3.1 Baseline Version

Model Specification. We assume that the economy evolves at monthly frequency according to the following VAR(p) dynamics:

$$x_t = \Phi_1 x_{t-1} + \ldots + \Phi_p x_{t-p} + \Phi_c + u_t, \quad u_t \sim iidN\big(0, \Sigma\big). \tag{11}$$

The $n \times 1$ vector of macroeconomic variables x_t can be composed into $x_t = [x'_{m,t}, x'_{q,t}]'$, where the $n_m \times 1$ vector $x_{m,t}$ collects variables that are observed at monthly frequency, e.g., the consumer price index and the unemployment rate, and the $n_q \times 1$ vector $x_{q,t}$ comprises the unobserved monthly variables that are published only at quarterly frequency, e.g., GDP. Define $z_t = [x'_t, \ldots, x'_{t-p+1}]'$ and $\Phi = [\Phi_1, \ldots, \Phi_p, \Phi_c]'$. Write the VAR in (11) in companion form as

$$z_t = F_1(\Phi) z_{t-1} + F_c(\Phi) + v_t, \quad v_t \sim iidN\big(0, \Omega(\Sigma)\big), \tag{12}$$

where the first n rows of $F_1(\Phi)$, $F_c(\Phi)$, and v_t are defined to reproduce (11) and the remaining rows are defined to deliver the identities $x_{q,t-l} = x_{q,t-l}$ for $l = 1, \ldots, p-1$. The $n \times n$ upper-left submatrix of Ω equals Σ and all other elements are zero. Equation (12) is the state-transition equation of the MF-VAR.

We proceed by describing the measurement equation. To handle the unobserved variables, we vary the dimension of the vector of observables as a function of time t (e.g., Durbin and Koopman 2001). Let T denote the forecast origin and let $T_b \leq T$ be the last period that corresponds to the last month of the quarter for which all quarterly observations are available. The subscript b stands for *balanced* sample. Up until period T_b the vector of monthly series $x_{m,t}$ is observed every month. We denote the actual observations by $y_{m,t}$ and write

$$y_{m,t} = x_{m,t}, \quad t = 1, \ldots, T_b. \tag{13}$$

Assuming that the underlying monthly VAR has at least three lags, that is, $p \geq 3$, we express the three-month average of $x_{q,t}$ as

$$\tilde{y}_{q,t} = \frac{1}{3}(x_{q,t} + x_{q,t-1} + x_{q,t-2}) = \Lambda_{qz} z_t. \tag{14}$$

For variables measured in logs, e.g., $\ln GDP$, the formula can be interpreted as a log-linear approximation to an arithmetic average of GDP that preserves the linear structure of the state-space model. For flow variables such as GDP, we adopt the national income and product accounts (NIPA) convention and annualize high-frequency flows. As a consequence, quarterly flows are the average and not the sum of monthly flows. This three-month average, however, is only observed for every third month, which is why we use a tilde superscript. Let $M_{q,t}$ be a selection matrix that equals the identity matrix if t corresponds to the last month of a quarter and is empty otherwise. Adopting the convention that the dimension of the vector $y_{q,t}$ is n_q in periods in which quarterly averages are observed and empty otherwise, we write

$$y_{q,t} = M_{q,t}\tilde{y}_{q,t} = M_{q,t}\Lambda_{qz} z_t, \quad t = 1, \ldots, T_b. \tag{15}$$

For periods $t = T_b + 1, \ldots, T$ no additional observations of the quarterly time series are available. Thus, for these periods the dimension of $y_{q,t}$ is zero and the selection matrix $M_{q,t}$ in (15) is empty.

However, the forecaster might observe additional monthly variables. Let $y_{m,t}$ denote the subset of monthly variables for which period t observations are reported by the statistical agency after period T, and let $M_{m,t}$ be a deterministic sequence of selection matrices such that (13) can be extended to

$$y_{m,t} = M_{m,t}x_{m,t}, \quad t = T_b + 1, \ldots, T. \tag{16}$$

Notice that the dimension of the vector $y_{m,t}$ is potentially time varying and less than n_m. The measurement equations (13) to (16) can be written more compactly as

$$y_t = M_t \Lambda_z z_t, \quad t = 1, \ldots, T. \tag{17}$$

Here, M_t is a sequence of selection matrices that selects the time t variables that have been observed by period T and are part of the forecaster's information set. In sum, the state-space representation of the MF-VAR is given by (12) and (17).

Bayesian Estimation. The starting point of Bayesian inference for the MF-VAR is a joint distribution of observables $Y_{1:T}$, latent states $Z_{0:T}$, and parameters (Φ, Σ), conditional on a presample $Y_{-p+1:0}$ to initialize lags. The distribution of observables and latent states conditional on the parameters is implied by the above state-space representation of the MF-VAR. For the marginal distribution of the parameters (Φ, Σ) we use a conjugate Minnesota prior. This prior dates back to Litterman (1980) and Doan, Litterman, and Sims (1984). We use the version of the Minnesota prior described in Del Negro and Schorfheide (2011)'s handbook chapter, which in turn is based on Sims and Zha (1998). The main idea of the Minnesota prior is to center the distribution of Φ at a value that implies a random-walk behavior for each of the components of x_t in (11). We implement the Minnesota prior by mixing artificial (or *dummy*) observations into the estimation sample. The artificial observations are computationally convenient and allow us to generate plausible a priori correlations between VAR parameters. The variance of the prior distribution is controlled by a low-dimensional vector of hyperparameters λ.

We generate draws from the posterior distributions of $(\Phi, \Sigma)|Z_{0:T}$ and $Z_{0:T}|(\Phi, \Sigma)$ using a Gibbs sampler. Based on these draws, we

are able to simulate future trajectories of y_t to characterize the predictive distribution associated with the MF-VAR and to calculate point, interval, and density forecasts.

3.2 Modifications

Robustifying Estimation and Forecasting. As discussed in Section 2, we distinguish between the handling of outliers at the estimation stage and when we run the filter to infer the latent states at the forecast origin, conditional on parameter draws from the posterior distribution. Let T denote the forecast origin. For the estimation, we will consider three approaches. Approaches E1 and E2 were labeled as robustifying ME in Section 2.

- E1: Estimation with observations from $t = 1, \ldots, t_*$, where $t_* < T$ is a pre-pandemic period. In the notation of Section 2 we will set the measurement error variance to infinity, i.e., $\chi_{u,t} = \infty$, for $t > t_*$, which is equivalent to setting the selection matrix $M_t = \emptyset$ in (17); or simply ending the estimation sample in period $t = t_*$.
- E2: Rather than ending the estimation sample at the onset of the pandemic, we drop a sequence of extreme observations during the early phase of the pandemic and retain subsequent observations from $t = t_{**} + 1, \ldots, T$. This is implemented by setting $M_t = \emptyset$ for $t = t_* + 1, \ldots, t_{**}$.
- E3: Finally, we consider an estimation based on the full sample $t = 1, \ldots, T$ that includes the extreme observations during the early part of the pandemic.

To determine the latent states at the forecast origin T conditional on a draw (Φ^i, Σ^i) from the posterior distribution, we consider three different filtering strategies. In each case the filter is run from $t = 1, \ldots, T$, but we potentially use M_t (equivalently $\chi_{u,t}$) or $\chi_{\epsilon,t}$ to robustify the filter against outliers.

- F1: We keep $\chi_{u,t} = \chi_{\epsilon,t} = 1$ and leave M_t unchanged.
- F2: Set $M_t = \emptyset$ for $t > t_*$. This means that the latent state estimates for period $t = t_* + 1, \ldots, T$ are generated by simulating ST forward.

- F3: From period t_* onward, increase the innovation variance by setting $\chi_{\epsilon,t}$ to a large value.

The baseline forecasts reported in Section 5 are generated by combining estimation E1 with filtering F1. The determination of t_* is discussed in Section 4.3 below.

Volatility Breaks and Discounting. LP propose allowing for a break in volatility, which is modeled through a variable s_t that scales up the residual covariance matrix during the period of the pandemic. Following their specification, we replace (11) with

$$x_t = \Phi_1 x_{t-1} + \ldots + \Phi_p x_{t-p} + \Phi_c + s_t u_t, \quad u_t \sim iidN\big(0, \Sigma\big). \quad (18)$$

It is assumed that $s_t = 1$ before time period $t = t_*$ in which the pandemic begins. Subsequently s_t evolves according to

$$\begin{aligned} &s_{t_*} = \bar{s}_0, \quad s_{t_*+1} = \bar{s}_1, \quad s_{t_*+2} = \bar{s}_2, \quad \text{and} \\ &s_{t_*+j} = 1 + (\bar{s}_2 - 1)\rho^{j-2}. \end{aligned} \quad (19)$$

The quadruplet $\vartheta = (\bar{s}_0^2, \bar{s}_1^2, \bar{s}_2^2, \rho)$ needs to be estimated. This flexible parameterization allows for this scaling factor to take three (possibly) different values in the first three periods after the outbreak of the disease, and to decay at rate ρ after that. Note that ϑ uniquely determines the sequence $S_{1:T} = \{s_1, \ldots, s_T\}$.[3]

4. Real-Time Data

We generated and published the forecasts presented in Section 5 in real time as the pandemic unfolded.[4] Section 4.1 summarizes the monthly and quarterly series used for the MF-VAR and the timing convention for the estimates and forecasts. The timing of the real-time SPF forecasts is described in Section 4.2.

[3] One can easily modify the specification to allow for more or fewer exceptional periods.

[4] Diebold (2020) distinguishes between "pseudo-real-time" analysis, meaning the use of expanding sample estimation and vintage *data*, and "real-time" analysis, meaning the use of real-time *information* rather than hindsight. We did the latter.

4.1 Monthly and Quarterly Time Series

We consider an MF-VAR for 11 macroeconomic variables, of which 3 are observed at quarterly frequency and 8 are observed at monthly frequency. The quarterly series are GDP, fixed investment (INVFIX), and government expenditures (GOV). The monthly series are the unemployment rate (UNR), hours worked (HRS), consumer price index (CPI), industrial production index (IP), personal consumption expenditures (PCE), federal funds rate (FF), 10-year Treasury-bond yield (TB), and S&P 500 index (SP500). Precise data definitions are provided in the appendix. Series that are observed at a higher than monthly frequency are time aggregated to monthly frequency. The variables enter the MF-VAR in log levels with the exception of UNR, FF, and TB, which are divided by 100 to make them commensurable in scale to the other log-transformed variables.

Our forecasts are based on real-time data sets, assuming that the econometric analysis is conducted on the last day of each month.[5] The timing convention and the data availability for each forecast origin are summarized in Table 2. A forecaster on April 30 has access to monthly observations from March; an initial release of Q1 GDP, investment, and government spending; as well as the April observations for the average federal funds rate, the Treasury-bond yield, and the S&P 500 index. In May, monthly non-financial observations on the April unemployment rate, hours worked, inflation, industrial production, and personal consumption expenditures become available. On June 30, two monthly observations for each non-financial variable are available for the second quarter. This pattern of information repeats itself every quarter. In the remainder of the paper we will refer to the forecast origins only by month and year, with the understanding that estimates and forecasts are based on the information available on the last day of the month.

4.2 Survey of Professional Forecasters

We compare the MF-VAR forecasts to median forecasts from the SPF. The timing of the SPF is summarized in Table 3. The

[5]Due to data revisions by statistical agencies, observations of $Y_{1:T-1}$ published in period T are potentially different from the observations that have been published in period $T-1$. Moreover, some series are published with a delay of several periods.

Table 2. Information at MF-VAR Forecast Origin

April 30												
		UNR	HRS	CPI	IP	PCE	FF	TB	SP500	GDP	INVFIX	GOV
Q1	M3	X	X	X	X	X	X	X	X	QAv	QAv	QAv
Q2	M4	∅	∅	∅	∅	X	X	X	X	∅	∅	∅
May 31												
		UNR	HRS	CPI	IP	PCE	FF	TB	SP500	GDP	INVFIX	GOV
Q1	M3	X	X	X	X	X	X	X	X	QAv	QAv	QAv
Q2	M4	X	X	X	X	X	X	X	X	∅	∅	∅
Q2	M5	∅	∅	∅	∅	∅	X	X	X	∅	∅	∅
June 30												
		UNR	HRS	CPI	IP	PCE	FF	TB	SP500	GDP	INVFIX	GOV
Q1	M3	X	X	X	X	X	X	X	X	QAv	QAv	QAv
Q2	M4	X	X	X	X	X	X	X	X	∅	∅	∅
Q2	M5	X	X	X	X	X	X	X	X	∅	∅	∅
Q2	M6	∅	∅	∅	∅	∅	X	X	X	∅	∅	∅

Note: ∅ indicates that the observation is missing. X denotes monthly observation and QAv denotes quarterly average.

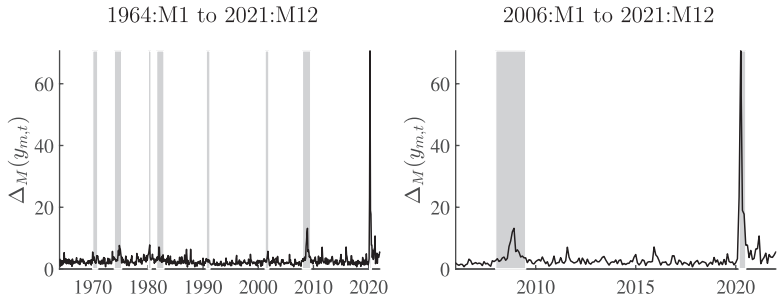
Table 3. Timing of Survey of Professional Forecasters

Survey Name	**Questionnaires Sent to Panelists**	**Submission Deadline**	**Last Quarter in Info Set**	**Quarterly Forecasts**
1st Quarter	End of January	Middle of February	Y–1:Q4	Y:Q1 to Y+1:Q1
2nd Quarter	End of April	Middle of May	Y:Q1	Y:Q2 to Y+1:Q2
3rd Quarter	End of July	Middle of August	Y:Q2	Y:Q3 to Y+1:Q3
4th Quarter	End of October	Middle of November	Y:Q3	Y:Q4 to Y+1:Q4

Note: The questionnaires are sent after the NIPA advance report. The submission deadline is in the second or third week of the month. "Y" refers to the year of the survey.

quarterly survey forecasts are comparable to our first-month-within-a-quarter forecasts (January, April, July, October). Because the survey respondents in principle have two more weeks after our end-of-month MF-VAR forecast origin, this comparison generates a slight

Figure 1. Extreme Monthly Observations: Mahalanobis Distance



Note: Based on our eight monthly series, we compute the Mahalanobis distance $\mathcal{D}(\Delta y_{m,t})$ defined in (20). The plot is based on the January 2022 vintage.

informational advantage for the SPF. On the other hand, a comparison with our third-month-within-a-quarter predictions (March, June, September, December) puts the SPF forecasts at a clear informational disadvantage against the MF-VAR because the most recent monthly data used in the MF-VAR forecasts are released well after the SPF submission deadline.

4.3 Outliers

Unlike in CCMM, we do not explicitly model the occurrence of outliers in our MF-VAR through fat-tailed error distributions. The outlier-robust estimation and forecasting methods described in Section 3.2 rely on the researcher to pre-specify the cut-off periods t_* and t_{**}. In view of the unprecedented mobility restrictions imposed by governments around the world in March 2020, there was no doubt that starting from 2020:M3 macroeconomic data would look very different from their pre-pandemic values and the COVID outliers were easily detectable in real time.

We are plotting the Mahalanobis distance for the eight monthly variables (converted into growth rates) in Figure 1. The distance is defined as

$$\mathcal{D}(\Delta y_{m,t}) = \sqrt{(\Delta y_{m,t} - \hat{\mu})' \hat{\Sigma}^{-1} (\Delta y_{m,t} - \hat{\mu})}. \tag{20}$$

Note that if $\Delta y_{m,t} \sim N(\hat{\mu}, \hat{\Sigma})$, then $\mathcal{D}^2(\Delta y_{m,t})$ has a χ^2 distribution with degrees of freedom equal to the dimension of $y_{m,t}$. Thus, $\mathcal{D}(\Delta y_{m,t})$ measures how far $\Delta y_{m,t}$ lies in the tail of its distribution if it were normally distributed. To generate the figure, we estimate $\hat{\mu}$ and $\hat{\Sigma}$ based on observations up to 2020:M1. In the plot we are simply using the January 2022 vintage, but the pattern is very similar with the real-time vintages. The left panel provides a historical perspective, plotting $\mathcal{D}(\Delta y_{m,t})$ from 1964 onward, whereas the right panel zooms into the last 16 years.

The values of $\mathcal{D}(\Delta y_{m,t})$ from 2020:M3 to 2020:M6 are unprecedented. Between February and March the distance measure jumped from 3.3 to 23.3 and it reached 70.6 in April. From June to July it dropped again from 17.8 to 7.7. For comparison, the largest value during the Great Recession was 13.2 in 2008:M11. Thus, in real time and also with hindsight, the monthly observations from 2020:M3 to 2020:M6 are clearly outliers. In the subsequent analysis we consider three types of estimation samples. For the baseline estimation, E1 in the terminology of Section 3.2, we use the January 2020 vintage which includes the 2020:M1 financial variables, the 2019:M12 monthly macroeconomic variables, and the 2019:Q4 quarterly macroeconomic variables.[6] For the estimation approach E2, we drop observations for periods in which $\mathcal{D}(\Delta y_{m,t})$ exceeds the value 16. This means that observations 2020:M1 and 2020:M2 are included in the estimation sample, the observations 2020:M3 to M6 and 2020:Q1 and Q2 are excluded from the estimation sample, and subsequent observations are included.[7] Finally, we consider the full-sample estimation E3, which includes all observations available at the forecast origin.

[6]Rather than re-estimating the MF-VAR month-by-month using the most recent vintage but keeping the estimation sample fixed, we simply freeze the parameter estimates.

[7]Maroz, Stock, and Watson (2021) estimate a dynamic factor model on a large set of macroeconomic and financial variables and show that including an additional factor for the COVID period improves the fit substantially. This new COVID factor takes on large values from 2020:M3 to 2020:M6, which is the period we drop from the estimation sample, but was less important afterward.

5. Empirical Results

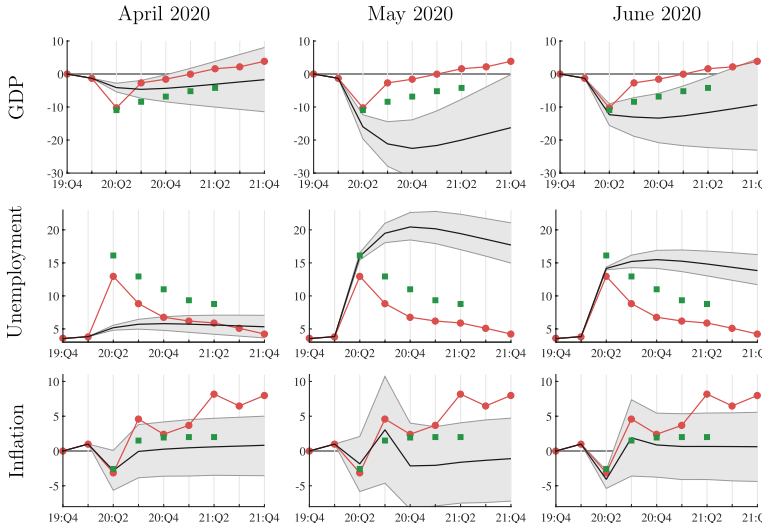
The pre-COVID forecast performance of the MF-VAR model used in this paper was documented in Schorfheide and Song (2015). We showed that the MF-VAR generates more accurate nowcasts and short-run forecasts than a VAR estimated on time-aggregated quarterly data. The improvement tempers off in the medium and long run. The short-run accuracy gain is largest in the third month of the quarter, when a lot of monthly data are available for the current quarter. We also documented that the monthly information helped the MF-VAR track the economic downturn during the 2008-09 (Great) recession period more closely in real time than a VAR estimated on quarterly data only. Similar results for other MF-VAR specifications have been obtained by Brave, Butters, and Justiniano (2019) and McCracken, Owyang, and Sekhposyan (2020).

We estimate the MF-VAR using $p = 6$ lags based on various 2020 and 2021 real-time data vintages and generate forecasts of quarterly averages of the 11 variables that appear in the VAR. All of our estimation samples start in 1964. The hyperparameter settings are the same as in Schorfheide and Song (2015). We subsequently examine the MF-VAR forecasts in chronological order: the COVID-19 outbreak in the United States in 2020:Q2 (Section 5.1), the continuation of the pandemic throughout the second half of 2020 (Section 5.2), and the first three quarters of 2021 (Section 5.3). Because our sample only spans 17 MF-VAR forecast origins and six quarters of SPF forecasts, we examine plots of forecasts and actuals, rather than computing forecast evaluation summary statistics.

5.1 The First Months of the COVID-19 Pandemic

April, May, and June forecasts for GDP, the unemployment rate, and CPI inflation are plotted in Figure 2. The panels show actual values from the January 2022 vintage (solid red), posterior median forecasts (solid black), and 90 percent posterior predictive intervals (light grey). Moreover, we also plot the median forecasts from the SPF. The MF-VAR forecasts are constructed from the real-time vintages available on the date of the forecast, as described in Section 4. Unless otherwise noted, we use the baseline approach of combining

Figure 2. Forecasts in 2020:Q2



Note: We forecast quarterly averages. Actual values (solid red, January 2022 vintage) and forecasts: median (solid black) and 90 percent bands (light grey) constructed from the posterior predictive distribution. Green squares represent median forecasts from the SPF. For GDP we depict percentage change relative to December 2019. The MF-VAR is estimated based on the January 2020 vintage; filtering uses the vintage available at the forecast origin (baseline, E1 and F1).

parameter estimation approach E1 with full-sample filtering F1; see Sections 3.2 and 4.3.

We are reporting forecasts of quarterly averages (see tick marks on the x-axes of the plots), which are obtained by averaging the within-quarter monthly values simulated from the MF-VAR. Depending on the forecast origin, actual values for some variables might be available for the first one or two months of the first quarter to be forecast. In this case, we generate the quarterly forecast by averaging actual and simulated values. While unemployment and inflation forecasts are plotted directly, we make the following adjustment for the graphical presentation of the GDP forecasts. First, we convert the level forecasts from the MF-VAR and the SPF into growth rate forecasts. Second, we add the level of GDP at the forecast origin according to the January 2022 vintage to the cumulative growth rate forecasts. This is equivalent to adjusting the level

of GDP forecasts by the difference between the GDP value at the forecast origin as measured in the January 2022 vintage and the real-time value at the forecast origin.[8]

Forecasts. In regard to the treatment of the latent states at the forecast origin conditional on the parameter estimates, the April forecasts are, with the exception of the April financial variables (federal funds rate, Treasury-bond yield, and S&P 500 index) based on Q1 and March data. The economic downturn started in the second half of March when the mobility restrictions became effective. According to the January 2022 vintage, quarter-on-quarter (Q-o-Q) GDP growth in Q1 was −1.3 percent, which is approximately 1.5 times the historical standard deviation in the estimation sample. Industrial production in March 2020 dropped by 3.9 percent and the unemployment rate increased from 3.5 percent to 4.4 percent. At an annualized rate, consumer prices fell by 3.9 percent. Recall from Figure 1 that the Mahalanobis distance for the monthly variables jumped from 3.3 to 23.3 in March.

Because the severity of the pandemic was not yet fully reflected in the observations available for the forecast origin, the April MF-VAR forecasts did not capture the unprecedented magnitude of the downturn. While the posterior median forecast for Q2 GDP growth was −2.8 percent, the actual drop was −9.4 percent. Likewise, the Q2 unemployment forecast was 5.2 percent, whereas the actual average unemployment rate in Q2 was 13.1 percent. The SPF forecasters had an additional week or two to gather information about the economic consequences of the pandemic and the freedom to make judgmental adjustments to model-based forecasts. The figure shows that the median SPF forecasts for GDP and inflation for Q2 were more pessimistic and thereby much closer to the respective actuals than the MF-VAR forecasts. In terms of inflation, the Q2 forecasts of MF-VAR and SPF essentially agree, but over the medium run the SPF forecasters correctly predicted a rise in inflation, whereas the median-run MF-VAR forecast stayed close to zero.

[8]Abstracting from publication lags, denote the time τ release of y_t by y_t^τ, $\tau = t, t+1, \ldots$. We would like to compare the forecasts $y_{t+h|t}$, $h = 1, \ldots, H$ to the end-of-sample values y_{t+h}^T. However, the real-time forecast origin value y_t^t ($h = 0$) does not match the final vintage value y_t^T. Thus, we correct the level of the forecasts by $y_t^T - y_t^t$.

We now turn to the May forecasts of GDP and unemployment, which use observations on the April monthly non-financial variables to quantify lagged values at the forecast origin. Under F1 filtering, the MF-VAR requires very large "COVID-19" shocks in the state-transition equation to be able to rationalize the April data. Because historically macroeconomic shocks have had very persistent effects on the time series included in the MF-VAR specification, the model predicts long-lasting adverse effects of the COVID-19 shocks: until the end of 2021 GDP will stay 15 percent to 20 percent below its December 2019 value and the unemployment rate will remain above 13 percent. A comparison to the actuals shows that these forecasts were overly pessimistic: over the forecast period, GDP reverts back to its 2019:Q4 level and the unemployment rate falls below 6 percent. The June information leads to slightly more favorable MF-VAR forecasts, but the model continues to predict long-lasting macroeconomic effects of the pandemic.

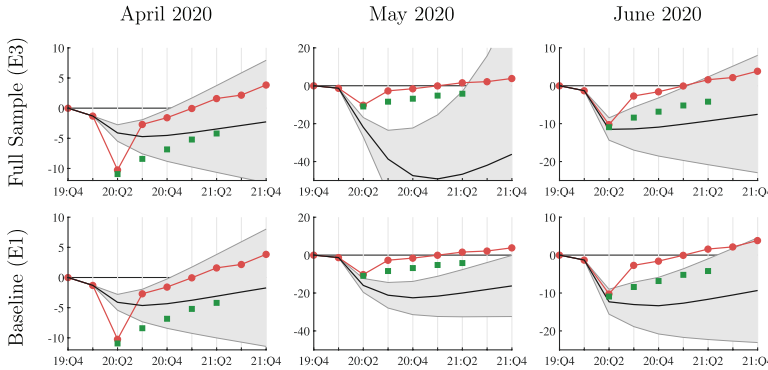
Recall that the survey underlying the SPF forecasts is only conducted quarterly. Thus, the May and June SPF forecasts plotted in Figure 2 are the same as the April forecasts.[9] As we have seen before, the SPF predicts a much faster recovery than the MF-VAR and, ex post, its median forecasts turned out to be much more accurate than the MF-VAR forecasts. Overall, the forecast performance of the MF-VAR during the first three months of the pandemic is poor, compared to the SPF, which presumably incorporates other data sources and judgment about the idiosyncratic nature of the COVID-19 recession.

Effect of Estimation Sample. We proceed by examining the effect of choosing the endpoint of the estimation sample on the forecasts. To construct the baseline forecasts, we used the approach E1 and excluded observations that became available after January 2020 from the estimation.[10] This is a sensible strategy if the pandemic was a shock to the economy that was unusually large, indeed several standard deviations in magnitude, but did not change the

[9] The plotted values can be slightly different within the quarter, because we are converting level forecasts into growth rate forecasts and add the growth rates to the level at the origin which may get revised from month to month.

[10] Recall from Section 2 that in a state-space setting this approach is equivalent to setting the measurement error variance to infinity.

Figure 3. Effect of Estimation Sample on GDP Forecasts



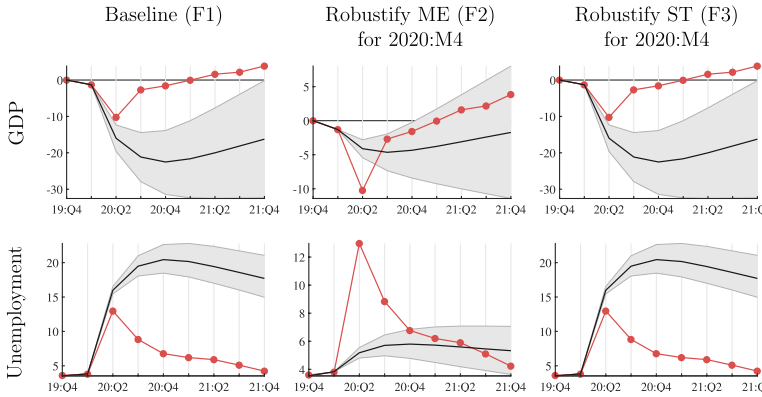
Note: We forecast quarterly averages. Actual values (solid red, January 2022 vintage) and forecasts: median (solid black) and 90 percent bands (light grey) constructed from the posterior predictive distribution. Green squares represent median forecasts from the SPF. We depict percentage change relative to December 2019. Full Sample (E3): The MF-VAR is estimated based on data up to the forecast origin. Baseline (E1): The MF-VAR is estimated based on the January 2020 vintage. Estimates are combined with filtering approach F1.

fundamental workings of the aggregate economy. Unless explicitly modeled, the COVID-19 outliers simply distort the parameter estimates. On the other hand, if the pandemic fundamentally changed macroeconomic dynamics, then the most recent observations should be included in the estimation, and earlier observations should possibly be discounted.

In Figure 3 we compare GDP forecasts from a full-sample estimation (top row) that includes data up to the forecast origin (but does not downweight pre-pandemic observations)—estimation approach E3 combined with filtering approach F1 in the terminology of Section 3.2—to the baseline forecasts. For April 2020 the two sets of forecasts are very similar, because the estimation sample ends with the Q1 and March non-financial variables which are not yet severely affected by the pandemic, as discussed previously.

The difference between the forecasts is most pronounced for the May 2020 forecast origin. Here the 90 percent bands under the full-sample estimation are considerably wider than the bands under the baseline estimation. Moreover, the median forecasts drop below

Figure 4. Effect of Filtering on Forecasts in May 2020



Note: We forecast quarterly averages. Actual values (solid red, January 2022 vintage) and forecasts: median (solid black) and 90 percent bands (light grey) constructed from the posterior predictive distribution. For GDP we depict percentage change relative to December 2019. The MF-VAR is estimated based on the January 2020 vintage (baseline estimation, E1).

−40 percent in 2020:Q4, whereas under the baseline estimation the median forecasts only fall to about −20 percent. The increase in forecast interval width is mainly driven by the estimates of Σ which increased due to the extreme observations in April 2020. The discrepancy among the forecasts shrinks again for the June 2020. In general, after the initial adjustment of the economy to the COVID-19 pandemic, we expect the magnitude of subsequent shocks to be more similar to the pre-2020 experience. For 2020:Q2 we find no upside in including post-January observations in the estimation sample.

Effect of Filtering to Extract States at Forecast Origin. In the subsequent experiment we revert to the baseline parameter estimation approach E1, excluding 2020 observations from the estimation sample. Instead, we vary the post-estimation filtering to infer the states at the forecast origin. The results are summarized in Figure 4. The plots in the first column of the figure reproduce the baseline May forecasts in Figure 2. The second column is generated using filtering approach F2. We set $M_t = \emptyset$ for 2020:M4, which means that the lagged values needed for the May forecasts are obtained by iterating the state-transition equation forward from

the March data onward. This has a drastic effect on the real activity forecasts. The F2 approach completely misses the Q2 drop in GDP and spike in unemployment. However, for Q3 onward it generates less pessimistic forecasts that turned out ex post to be closer to the actual path of the economy. Under F1 the ST propagates the large shock to rationalize the April observations, whereas F2 attributes the outliers to measurement errors and does not propagate them forward.

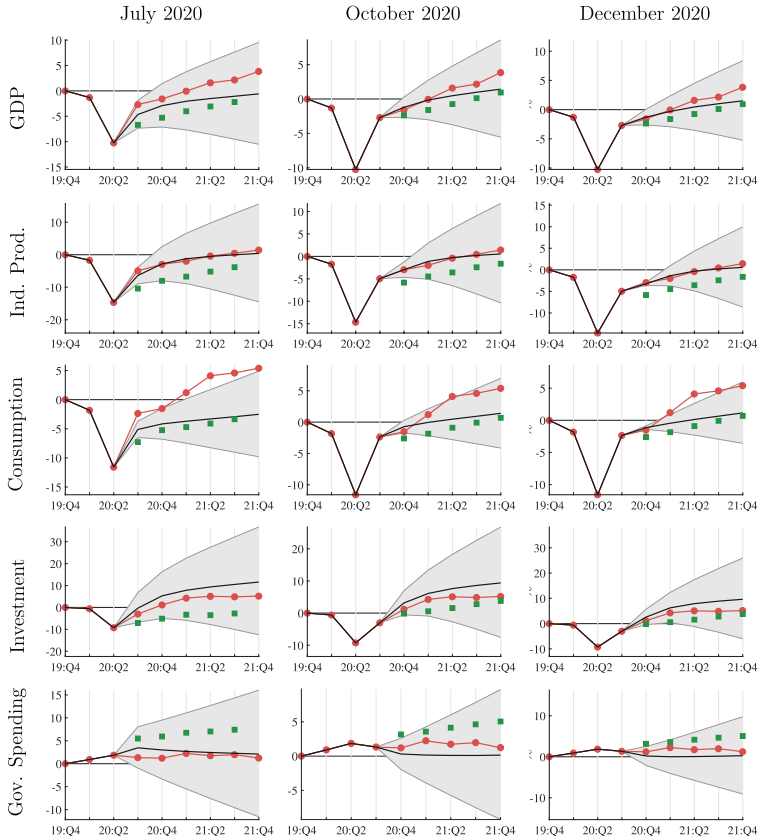
In the last column we show forecasts based on approach F3 where we robustify the ST by inflating the estimated innovation covariance matrix by a factor of 30^2. Based on the calculations summarized in Table 1, it is not surprising that the F3 forecasts look essentially identical to the F1 forecasts. Despite the inflated innovation variance, inference on the latent state continues to be driven by the April outlier.

5.2 The Second Half of 2020

We now turn to forecasts during 2020:Q3 and Q4. We proceed with the baseline approach of combining E1 and F1. MF-VAR forecasts, SPF forecasts, and actuals from the January 2022 vintage are presented in three figures: real activity variables in Figure 5, labor market variables in Figure 6, and inflation and financial variables in Figure 7. The July 2020 panels overlay the MF-VAR forecasts with SPF forecasts from the Q3 survey, whereas the October and December 2020 panels compare the MF-VAR forecasts to Q4 SPF forecasts. As discussed in Section 4.2, for the first month within a quarter, the SPF forecasters have a slight informational advantage compared to the MF-VAR, whereas for the third month within a quarter, the MF-VAR has a strong advantage.

Real Activity and Government Spending. In July 2020 the MF-VAR predicts that GDP and industrial production return to their respective 2019:Q4 values in the second half of 2021. The median forecast from the SPF is less optimistic and about 3 percent to 5 percent lower than the MF-VAR forecast. While the October and December industrial production forecasts from the MF-VAR look quite similar to the July forecast, the GDP forecasts made in 2020:Q4 imply a slightly stronger recovery than the July forecast. Compared to the July forecasts, the gaps between MF-VAR and

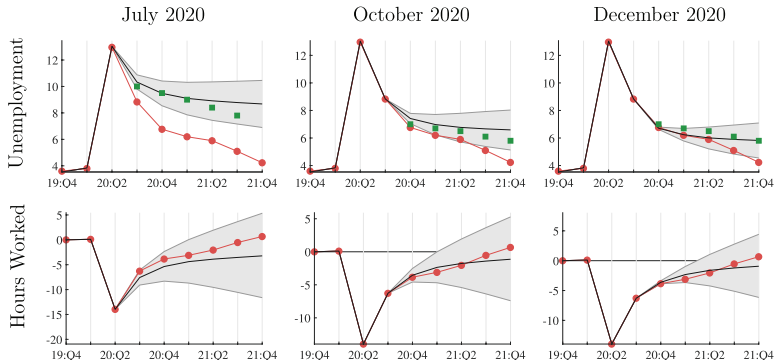
Figure 5. Real Activity Forecasts in 2020:Q3 and Q4



Note: We forecast quarterly averages. Actual values (solid red, January 2022 vintage) and forecasts: median (solid black) and 90 percent bands (light grey) constructed from the posterior predictive distribution. Green squares represent median forecasts from the SPF. We depict percentage change relative to December 2019. The MF-VAR is estimated based on the January 2020 vintage; filtering uses the vintage available at the forecast origin (baseline approach, E1 and F1).

SPF forecasts narrow in Q4 (October and December). Compared to the Q2 forecasts presented in Section 5.1, the most remarkable difference in the second half of 2020 is that the posterior median forecasts produced in Q3 and Q4 accurately predict GDP and industrial production over a one-year horizon. In fact, the MF-VAR forecasts are now more accurate than the SPF forecasts.

Figure 6. Labor Market Forecasts in 2020:Q3 and Q4

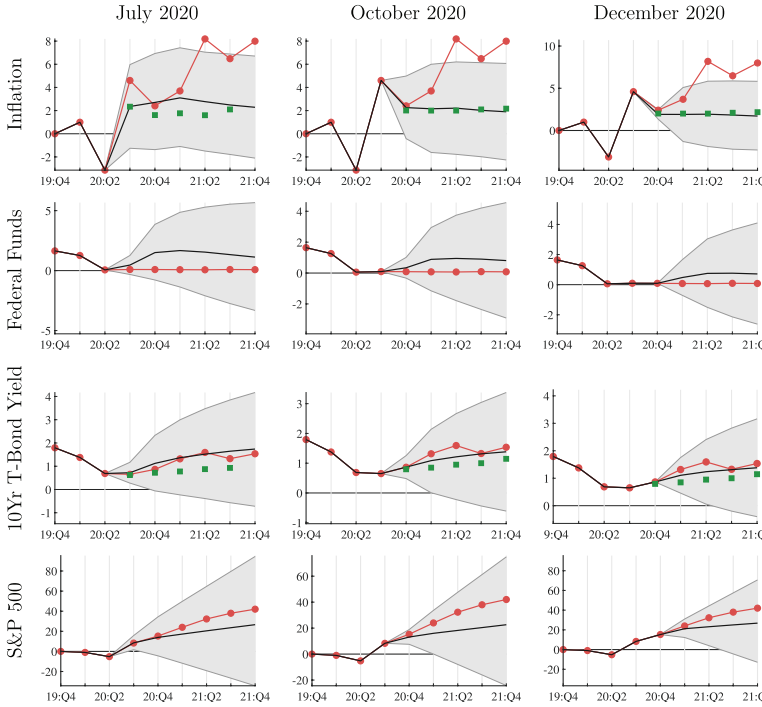


Note: We forecast quarterly averages. Actual values (solid red, January 2022 vintage) and forecasts: median (solid black) and 90 percent bands (light grey) constructed from the posterior predictive distribution. Green squares represent median forecasts from the SPF. For Hours Worked we depict percentage change relative to December 2019. The MF-VAR is estimated based on the January 2020 vintage; filtering uses the vintage available at the forecast origin (baseline, E1 and F1).

The MF-VAR predicts throughout the second half of 2020 that consumption returns to its December 2019 value by 2021:Q2 (October and December forecasts). The forecast error for the 2020:Q4 observation is close to zero. In 2021 actual consumption rises faster than the forecast, but by and large stays within the 90 percent credible intervals. Only the band from the December 31 forecast for 2021:Q2 does not cover the actual value. The posterior median forecasts for investment imply a quick recovery. By the end of 2021, investment is expected to be about 10 percent above the 2019:Q4 value. The forecasts from all three origins are quite accurate. Finally, the last row of Figure 5 shows government spending forecasts. The median MF-VAR consumption and investment forecasts are overall more optimistic than the SPF forecast in regard to the recovery from the pandemic downturn.

Labor Market. Unemployment and hours worked forecasts are presented in Figure 6. For the unemployment rate the MF-VAR and SPF forecasts are very similar. The Q3 and Q4 SPF forecasts are slightly lower than the July and October MF-VAR forecasts,

Figure 7. Inflation and Financial Forecasts in 2020:Q3 and Q4



Note: We forecast quarterly averages. Actual values (solid red, January 2022 vintage) and forecasts: median (solid black) and 90 percent bands (light grey) constructed from the posterior predictive distribution. Green squares represent median forecasts from the SPF. For S&P 500 Returns we depict percentage change relative to December 2019. The MF-VAR is estimated based on the January 2020 vintage; filtering uses the vintage available at the forecast origin (baseline, E1 and F1).

respectively, in particular over a one-year horizon. By December unemployment had fallen substantially compared to its Q2 peak and now the MF-VAR forecast that utilizes the most recent information is below the SPF forecast, at least in the short run. In absolute terms, the July forecast is too pessimistic: unemployment falls more quickly than predicted and the actual path is outside of the 90 percent band. The December MF-VAR unemployment forecast, on the other hand, is very accurate; only the 2021:Q4 actual lies slightly outside of the

predictive interval. The second row of Figure 6 demonstrates that the MF-VAR is successful in predicting the recovery of hours worked.

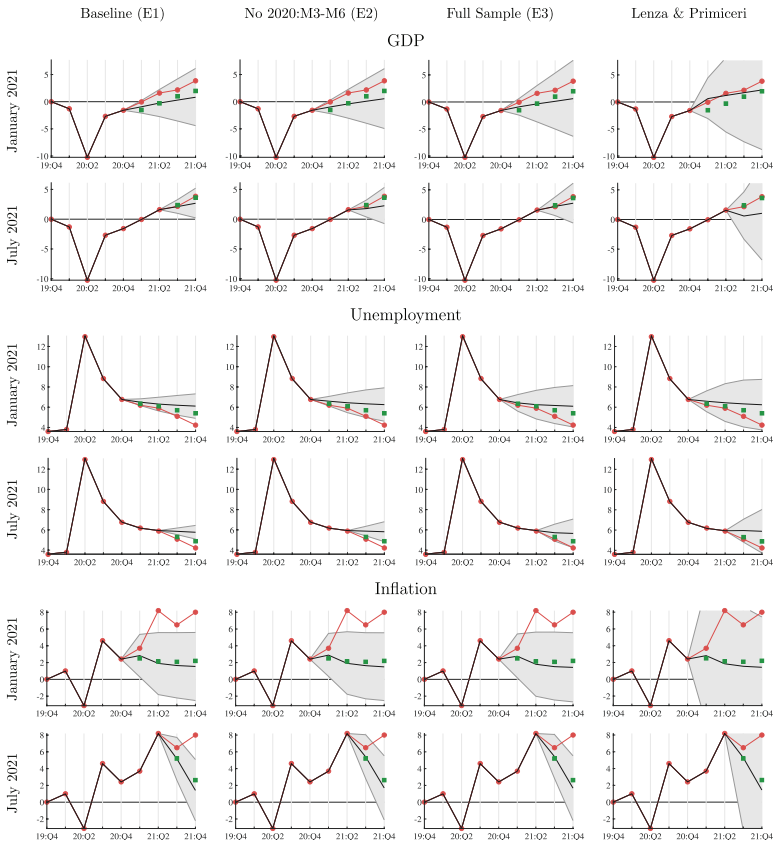
Inflation and Financial Indicators. Figure 7 shows forecasts of inflation and the financial variables. For October and December, the MF-VAR and SPF forecasts of inflation are very similar, despite the additional information used by the December MF-VAR. Both approaches capture the drop in inflation between 2020:Q3 and 2020:Q4, but they miss the subsequent rise in inflation, predicting that inflation stays around 2 percent throughout 2021. The MF-VAR predicts a lift-off from the effective lower bound on nominal interest rates which did not happen. While the 10-year Treasury-bond yield forecast in July correctly predicts the actual path, the median forecasts in October and December slightly underestimate the actual yield. For all three forecast origins, the MF-VAR yield forecasts dominate the SPF forecasts. Finally, the MF-VAR implied median forecasts for the S&P 500 slightly underpredict the stock market recovery.

Overall, we conclude that while the forecast performance of the MF-VAR was poor in the second quarter of 2020, the predictions are back on track in the third and fourth quarter. In fact, they are in general as good as or better than the SPF forecasts.

5.3 Forecasts in 2021

Forecasts of GDP, unemployment, and CPI inflation made in January and July 2021 are depicted in Figure 8. Going forward, an important question when estimating VARs and other time-series models will be how to treat the extreme observations during the pandemic. We previously compared our baseline approach E1 of ending the estimation sample in January 2020 to the full-sample estimation E3. In addition, we now consider the following two estimation strategies. First, instead of ending the estimation sample in January 2020, we only drop four months of extreme observations (from March to June, denoted by No 2020:M3–M6) from the estimation sample, denoted by E2 in Section 3.2. We thereby allow the estimates to adapt to post–June 2020 observations. Second, we implement the LP proposal of scaling the innovation variance during 2020:Q2 in a data-driven manner and then letting the scale factor decline subsequently. By inflating the innovation covariance matrix, this approach

Figure 8. January and July Forecasts for 2021



Note: We forecast quarterly averages. Actual values (solid red, January 2022 vintage) and forecasts: median (solid black) and 90 percent bands (light grey) constructed from the posterior predictive distribution. Green squares represent median forecasts from the SPF. For GDP we depict percentage change relative to December 2019. Baseline (E1): estimation based on the January 2020 vintage. No 2020:M3–M6 (E2): we treat monthly observations from 2020:M3 to M6 and quarterly observations for 2020:Q1 and Q2 as missing. Full Sample (E3): The MF-VAR is estimated based on data available at the forecast origin. Filtering uses the vintage available at the forecast origin (F1).

also discounts pandemic observations in the model estimation stage. Throughout, we use filtering approach F1.

For the forecast origins January and July 2021 baseline estimation, the No-2020:M3–M6 estimation, and full-sample estimation

yield very similar posterior median predictions, which are very close to the SPF predictions. The main difference is the width of the predictive intervals. The full estimation approach generates the widest interval among E1, E2, and E3, because the estimate of Σ is heavily influenced by the outliers in 2020:Q2. The baseline approach yields the shortest intervals because all pandemic observations are excluded from the estimation of Σ. The approach of excluding only observations from March to June 2020 generates intervals that are slightly wider than the baseline intervals but shorter than the full-sample estimation. For instance, for the January 2021 forecast origin, the widths of the predictive intervals for 2021:Q1 are 2.3, 2.5, and 3.1 under E1, E2, and E3. The interval widths for the four-quarter-ahead forecasts are 10.5, 11.0, and 14.0, respectively. Compared to the widths of the predictive intervals, the differences in the posterior median forecasts across estimation strategies are small (less than 0.3 percentage point).

The width-based ranking of the E1, E2, and E3 forecasts for unemployment and inflation is the same as for GDP. The main difference is that for the July 2021 forecast origin the wider unemployment intervals associated with the full-sample estimates contain the actual values whereas the narrower E1 and E2 intervals do not. None of the three estimation approaches generates a forecast that is able to capture the high level of inflation in 2021.

Finally, we turn to the forecasts generated with the LP approach. The key difference is that the predictive intervals are substantially wider. For instance, the widths of the January 2021 GDP forecast intervals for 2021:Q1 and Q4 GDP are 7.5 and 22.4, respectively. Thus, they are approximately twice as wide as the other intervals. The estimated decay coefficient ρ is too large for the scaling to decay sufficiently fast. Ex post, the band for the GDP forecast is unreasonably wide. The posterior median GDP point forecasts from the LP approach generated in January are slightly higher and more accurate than the other three forecasts, whereas the July forecasts are slightly lower and less accurate. The remaining panels in Figure 8 depict unemployment and inflation forecasts. The posterior median point forecasts are very similar across all four approaches, including LP. As in the case of GDP, the most striking difference between the LP interval predictions and the other three interval forecasts is that the LP intervals are considerably wider. For unemployment

and inflation this arguably works in favor of LP because the unemployment and inflation realizations in the second half of 2021 are included in the predictive intervals.

A quantitative comparison of our empirical findings to those reported by LP and CCMM is difficult because the variables included in the VAR are different, the other studies are not using real-time data, and the forecast objects—in our case, variables aggregated to quarterly frequency—are different. Moreover, LP condition their forecasts on the path of unemployment. Figure 3 of LP compares June 2020 interval forecasts from the volatility-break model to a constant-volatility model estimated with data up until February 2020 (baseline estimation E1 in our classification). LP's volatility-break intervals are considerably wider than the constant-volatility intervals. Figure 4 of LP provides the same comparison for the May 2021 forecasts. Here the intervals from the two methods have similar widths, meaning that the estimated decay rate for the volatility spike is quite large. In contrast, we find that for July 2021 the intervals generated with the LP approach remain considerably wider than our baseline intervals, as we estimate a smaller decay rate for our set of variables. CCMM plot unemployment forecasts in their Figure 5, comparing their preferred stochastic volatility (SV) model with outliers and conditional Student-t distributions to an SV model in which outliers were ex ante identified and replaced by imputed values in a way that is similar to what we labeled as "robust ME." Their December 2020 forecast is comparable to our January 2020 forecast. For that particular origin, dropping versus modeling the outliers delivers very similar point and interval forecasts which appear to match our unemployment forecasts.

Overall, we conclude that excluding the March to June 2020 observations is a promising way of handling the estimation problem of vector autoregressive time-series models going forward. This approach trades off ease of implementation against some elegance and accuracy and is desirable in settings in which a more sophisticated modeling approach appears to be overly costly. Our state-space approach treats outliers as missing observations and imputes values based on the estimated VAR law of motion. In regular VARs the analysis could be further simplified by also dropping periods for which right-hand-side variables are contaminated by extreme outliers, instead of imputing missing lags.

6. Conclusion

We resuscitated the mixed-frequency vector autoregression (MF-VAR) developed in Schorfheide and Song (2015) to generate macroeconomic forecasts for the United States during the COVID-19 pandemic from April 2020 to August 2021 in real time. While the forecasting performance of the MF-VAR was quite poor compared to the SPF from April to June 2020, the MF-VAR produced forecasts that are of similar accuracy as the SPF forecasts from July 2020 onward. The only adjustment that we made relative to our pre-pandemic MF-VAR specification was to exclude observations from the early part of the pandemic from the estimation sample. Importantly, we did not modify the model ex post to be able to generate good forecasts in the March to June period retrospectively. This finding is remarkable because it implies that going forward, in applications in which a careful modeling of outliers is impractical, VARs can simply be estimated by excluding observations from the first half of 2020. Our results also suggest that it is prudent for the assessment of the forecasting performance of time-series models to separate the first months of the pandemic from later periods. Because the period from March to June 2020 was highly unusual in many dimensions, the forecast performance in the subsequent months is likely to be more indicative of future forecast performance.

Appendix

This appendix consists of the following sections:

- A.1 Derivations for Section 2
- A.2 Computational Details
- A.3 Data Set

A.1 Derivations for Section 2

We calculate

$$\begin{aligned}\ell_y(\hat{\theta}, y_t; Y_{(-t)}, \chi_t) &= \frac{\partial}{\partial y_t}\big[\ln p(y_t|Y_{1:t-1}, \hat{\theta}, \chi_t) + \ln p(y_{t+1}|Y_{1:t}, \hat{\theta}, \chi_t)\big] \\ &= -\frac{y_t - y_{t|t-1}}{F_{t|t-1}} + \frac{y_{t+1} - y_{t+1|t}}{F_{t+1|t}}\frac{\partial}{\partial y_t}y_{t+1|t} \\ &= -\frac{y_t - y_{t|t-1}}{F_{t|t-1}} + \frac{y_{t+1} - y_{t+1|t}}{F_{t+1|t}}\theta\lambda(P_{t-1|t-1}, \chi_t, \theta).\end{aligned}$$

In turn,

$$\begin{aligned}
&\ell_{y\theta}(\hat{\theta}, y; Y_{(-t)}) \\
&= \frac{1}{F_{t|t-1}} \frac{\partial}{\partial \theta} y_{t|t-1} + \frac{y_t - y_{t|t-1}}{F_{t|t-1}^2} \frac{\partial}{\partial \theta} F_{t|t-1} \\
&\quad - \frac{1}{F_{t+1|t}} \left(\frac{\partial}{\partial \theta} y_{t+1|t} \right) \theta \lambda(P_{t-1|t-1}, \chi_t, \theta) \\
&\quad - \frac{y_{t+1} - y_{t+1|t}}{F_{t+1|t}^2} \left(\frac{\partial}{\partial \theta} F_{t+1|t} \right) \theta \lambda(P_{t-1|t-1}, \chi_t, \theta) \\
&\quad + \frac{y_{t+1} - y_{t+1|t}}{F_{t+1|t}} \Big[\lambda(P_{t-1|t-1}, \chi_t, \theta) + \theta \lambda_\theta(P_{t-1|t-1}, \chi_t, \theta) \Big].
\end{aligned}$$

We obtain

$$\begin{aligned}
\frac{\partial}{\partial \theta} y_{t|t-1} &= s_{t-1|t-1} \\
\frac{\partial}{\partial \theta} F_{t|t-1} &= 2\theta P_{t-1|t-1} \\
\frac{\partial}{\partial \theta} y_{t+1|t} &= \lambda(P_{t-1|t-1}, \chi_t, \theta) y_t + \theta \lambda_\theta(P_{t-1|t-1}, \chi_t, \theta) y_t \\
&\quad + 2\theta \big[\lambda(P_{t-1|t-1}, \chi_t, \theta) - 1 \big] s_{t-1|t-1} \\
&\quad + \theta^2 \lambda_\theta(P_{t-1|t-1}, \chi_t, \theta) s_{t-1|t-1} \\
\frac{\partial}{\partial \theta} F_{t+1|t} &= 2\theta \lambda(P_{t-1|t-1}, \chi_t, \theta) \chi_{u,t} + \theta^2 \lambda_\theta\big(P_{t-1|t-1}, \chi_t, \theta\big) \chi_{u,t}.
\end{aligned}$$

Moreover,

$$\begin{aligned}
\lambda_\theta(P, \chi, \theta) &= \frac{2\theta P(\theta^2 P + \chi_{\epsilon,t} + \chi_{u,t}) - 2\theta P(\theta^2 P + \chi_{\epsilon,t})}{(\theta^2 P + \chi_{\epsilon,t} + \chi_{u,t})^2} \\
&= \frac{2\theta P \chi_{u,t}}{(\theta^2 P + \chi_{\epsilon,t} + \chi_{u,t})^2}.
\end{aligned}$$

We can now take limits

$$\lim_{\chi_{u,t} \longrightarrow \infty} F_{t|t-1} = 0$$

$$\lim_{\chi_{u,t}\longrightarrow\infty} \lambda(P,\chi_t,\theta) = 0$$

$$\lim_{\chi_{u,t}\longrightarrow\infty} \lambda(P,\chi_t,\theta)\chi_{u,t} = \theta^2 P_{t-1|t-1} + 1$$

$$\lim_{\chi_{u,t}\longrightarrow\infty} \lambda_\theta(P,\chi_t,\theta) = 0$$

$$\lim_{\chi_{u,t}\longrightarrow\infty} \lambda_\theta(P,\chi_t,\theta)\chi_{u,t} = 2\theta$$

$$\lim_{\chi_{u,t}\longrightarrow\infty} \frac{\partial y_{t+1|t}}{\partial\theta} = 2\theta s_{t-1|t-1}$$

$$\lim_{\chi_{u,t}\longrightarrow\infty} \frac{\partial F_{t+1|t}}{\partial\theta} = 4\theta^3 + 2\theta$$

$$\lim_{\chi_{\epsilon,t}\longrightarrow\infty} F_{t|t-1} = 0$$

$$\lim_{\chi_{\epsilon,t}\longrightarrow\infty} \lambda(P,\chi_t,\theta) = 1$$

$$\lim_{\chi_{\epsilon,t}\longrightarrow\infty} \lambda(P,\chi_t,\theta)\chi_{u,t} = 1$$

$$\lim_{\chi_{\epsilon,t}\longrightarrow\infty} \lambda_\theta(P,\chi_t,\theta) = 0$$

$$\lim_{\chi_{\epsilon,t}\longrightarrow\infty} \lambda_\theta(P,\chi_t,\theta)\chi_{u,t} = 0$$

$$\lim_{\chi_{\epsilon,t}\longrightarrow\infty} \frac{\partial y_{t+1|t}}{\partial\theta} = y_t$$

$$\lim_{\chi_{\epsilon,t}\longrightarrow\infty} \frac{\partial F_{t+1|t}}{\partial\theta} = 2\theta.$$

In turn,

$$\lim_{\chi_{u,t}\longrightarrow\infty} \ell_{y\theta}(\hat{\theta}, y_t; Y_{(-t)}, \chi_t) = 0$$

$$\lim_{\chi_{\epsilon,t}\longrightarrow\infty} \ell_{y\theta}(\hat{\theta}, y_t; Y_{(-t)}, \chi_t) = -\frac{\theta y_t}{\theta^2+2} - 2\theta^2\frac{y_{t+1}-\theta y_t}{(\theta^2+2)^2} + \frac{y_{t+1}-\theta y_t}{\theta^2+2}.$$

A.2 Computation Details

We modify the posterior sampler developed in Schorfheide and Song (2015) to account for the latent scale sequence s_t defined in (19) and its associated parameter vector ϑ. The Bayesian computations are implemented with a Metropolis-within-Gibbs sampler that iterates over the following three conditional distributions:[11]

$$\begin{aligned} p(\Phi, \Sigma | Z_{0:T}, Y_{-p+1:T}, \vartheta) &\propto p(Z_{1:T} | z_0, \Phi, \Sigma, \vartheta) p(\Phi, \Sigma | \lambda) \\ p(Z_{0:T} | \Phi, \Sigma, Y_{-p+1:T}, \vartheta) &\propto p(Y_{1:T} | Z_{1:T}) p(Z_{1:T} | z_0, \Phi, \Sigma, \vartheta) \\ &\quad \times p(z_0 | Y_{-p+1}) \\ p(\vartheta | \Phi, \Sigma, Z_{0:T}, Y_{-p+1:T}) &\propto p(Z_{1:T} | z_0, \Phi, \Sigma, \vartheta) p(\vartheta), \end{aligned} \tag{A.1}$$

where the third distribution is new. The modifications are as follows:

Step 1: Conditional on $Z_{0:T}$ the MF-VAR reduces to a standard linear Gaussian VAR with a conjugate prior. To sample from $p(\Phi, \Sigma | Z_{0:T}, Y_{-p+1:T}, \vartheta)$, write the VAR in slight abuse of notation as

$$x_t' = z_{t-1}' \Phi + s_t u_t', \tag{A.2}$$

where we treat x_t as observed and incorporate a constant term in the definition of z_{t-1}. Recall that the sequence $\{s_t\}$ can be generated from ϑ. The likelihood function is given by

$$\begin{aligned} &p(X | \Phi, \Sigma, S) \\ &\propto \left(\prod_{t=1}^{T} |s_t^2 \Sigma|^{-1/2} \right) \\ &\quad \times \exp \left\{ -\frac{1}{2} \sum_{t=1}^{T} \mathrm{tr} \left[s_t^{-2} \Sigma^{-1} (x_t' - z_{t-1}' \Phi)' (x_t' - z_{t-1}' \Phi) \right] \right\} \end{aligned}$$

[11] Lenza and Primiceri (2022) use a posterior sampler that integrates out (Φ, Σ) analytically. Because of the state-space form of the MF-VAR, this approach is not feasible in our settings. Thus, we will use a Metropolis-within-Gibbs step.

$$\propto \left(\prod_{t=1} s_t^{-n}\right) |\Sigma|^{-T/2} \exp\left\{-\frac{1}{2}\text{tr}\left[\Sigma^{-1}(\tilde{X} - \tilde{Z}_{-1}\Phi)'(\tilde{X} - \tilde{Z}_{-1}\Phi)\right]\right\}, \tag{A.3}$$

where $\tilde{X}$ has rows x_t'/s_t and $\tilde{Z}_{-1}$ has rows z_{t-1}/s_t. Thus, the posterior sampler for (Φ, Σ) only requires the transformation of X into $\tilde{X}$ and Z_{-1} into $\tilde{Z}$.

Step 2: Sampling from $p(Z_{0:T}|\Phi, \Sigma, Y_{-p+1:T}, \vartheta)$ can be easily implemented by replacing the covariance matrix Σ by $\tilde{\Sigma}_t = s_t^2\Sigma$ in every period t.

Step 3: We follow Lenza and Primiceri (2022) in using a Pareto distribution to form a prior for $\bar{s}_j^2$ and a Beta distribution for ρ. Thus,

$$p(\vartheta) = \left(\prod_{j=0}^{2} \frac{1}{\bar{s}_j^2}\mathbb{I}\{\bar{s}_j^2 \geq 1\}\right) \frac{1}{B(p,q)}\rho^{p-1}(1-\rho)^{q-1}, \tag{A.4}$$

where p and q are chosen such that the Beta distribution has mean 0.8 and standard deviation 0.2. We split ϑ into three components—$\bar{s}_0^2$, $\bar{s}_1^2$, and $(\bar{s}_2^2, \rho)$—and sample each component conditional on values for the other three components.

Sampling from the Posterior of $\bar{s}_0^2$. Note that $\bar{s}_0^2$ only affects the density for period t_*:

$$\begin{aligned} &p(\bar{s}_0^2|\Phi, \Sigma, Z_{0:T}, Y_{-p+1:T}, \vartheta_-) \\ &\quad\propto \frac{1}{\sqrt{\bar{s}_0^{2n}}} \exp\left\{-\frac{1}{2\bar{s}_0^2}\text{tr}\left[\Sigma^{-1}(x_{t_*}' - z_{t_*-1}'\Phi)'(x_{t_*}' - z_{t_*-1}'\Phi)\right]\right\} \\ &\qquad\times \mathbb{I}\{\bar{s}_0^2 \geq 1\}\frac{1}{\bar{s}_0^2}. \end{aligned}$$

Now define

$$\beta_0 = \frac{1}{2}\text{tr}\left[\Sigma^{-1}(x_{t_*}' - z_{t_*-1}'\Phi)'(x_{t_*}' - z_{t_*-1}'\Phi)\right], \quad \alpha_0 = n/2.$$

Then the distribution of $\bar{s}_0^2$ is inverse Gamma (α_0, β_0) truncated at 1.[12] Notice that the distribution of $\bar{s}_0^2$ is independent of the other ϑ elements.

Sampling from the Posterior of $\bar{s}_1^2$. Note that $\bar{s}_1^2$ only affects the density for period $t_* + 1$:

$$p(\bar{s}_1^2|\Phi, \Sigma, Z_{0:T}, Y_{-p+1:T}, \vartheta_-)$$
$$\propto \frac{1}{\sqrt{\bar{s}_1^{2n}}} \exp\left\{-\frac{1}{2\bar{s}_1^2}\mathrm{tr}\big[\Sigma^{-1}(x'_{t_*+1} - z'_{t_*}\Phi)'(x'_{t_*+1} - z'_{t_*}\Phi)\big]\right\}$$
$$\times \mathbb{I}\{\bar{s}_1^2 \geq 1\}\frac{1}{\bar{s}_1^2}.$$

Now define

$$\beta_1 = \frac{1}{2}\mathrm{tr}\big[\Sigma^{-1}(x'_{t_*+1} - z'_{t_*}\Phi)'(x'_{t_*+1} - z'_{t_*}\Phi)\big], \quad \alpha_1 = n/2.$$

Then the distribution of $\bar{s}_1^2$ is inverse Gamma (α_1, β_1) truncated at 1. Notice that the distribution of $\bar{s}_1^2$ is independent of the other ϑ elements.

Sampling from the Posterior of $(\bar{s}_2^2, \rho)$. Note that $(\bar{s}_2^2, \rho)$ affect the density for period $t_* + 2$ onwards:

$$p(\bar{s}_2^2, \rho|\Phi, \Sigma, Z_{0:T}, Y_{-p+1:T}, \vartheta_-)$$
$$\propto \left(\prod_{t=t_*+2}^{T} \frac{1}{1 + (\sqrt{\bar{s}_2^2} - 1)\rho^{t-t_*-2}}\right)^n$$
$$\times \exp\left\{-\frac{1}{2}\sum_{t=t_*+2}^{T} \frac{1}{\left(1 + (\sqrt{\bar{s}_2^2} - 1)\rho^{t-t_*-2}\right)^2}\right.$$
$$\left.\times \mathrm{tr}\big[\Sigma^{-1}(x'_{t_*} - z'_{t_*-1}\Phi)'(x'_{t_*} - z'_{t_*-1}\Phi)\big]\right\}$$
$$\times \mathbb{I}\{\bar{s}_2^2 \geq 1\}\frac{1}{\bar{s}_2^2}\frac{1}{B(p,q)}\rho^{p-1}(1-\rho)^{q-1}.$$

[12] The $IG(\alpha, \beta)$ distribution has density $(\beta^\alpha/\Gamma(\alpha))(1/x)^{\alpha+1}\exp(-\beta/x)$.

Table A.1. ALFRED Series Used in Analysis

Time Series	ALFRED Name
Gross Domestic Product (GDP)	GDPC1
Fixed Investment (INVFIX)	FPIC1
Government Expenditures (GOV)	GCEC1
Unemployment Rate (UNR)	UNRATE
Hours Worked (HRS)	AWHI
Consumer Price Index (CPI)	CPIAUCSL
Industrial Production Index (IP)	INDPRO
Personal Consumption Expenditure (PCE)	PCEC96
Federal Fund Rate (FF)	FEDFUNDS
10-Year Treasury Bond Yield (TB)	GS10
S&P 500 (SP500)	SP500

This density does not belong to a specific family of distributions from which we can sample directly. Thus, we use a random-walk Metropolis-Hastings step.

Initialization of ϑ and Proposal Covariance Matrix. (i) We initialize ρ^0 using the prior mean of 0.8. (ii) Based on parameter estimates from a run that stops estimation in, say, December 2019, we use the estimates $(\hat{\Phi}, \hat{\Sigma}, \hat{Z})$ to compute (α_j, β_j) for the conditional posteriors $p(\bar{s}_j|\cdot)$, $j = 0, 1$. We initialize $\bar{s}_j$ using the mean $\beta_j/(\alpha_j - 1)$. Assuming that $\bar{s}_2$ only affects observation $t = t_* + 2$, the same approach can be used to initialize $\bar{s}_2$. (iii) For the proposal covariance matrix in the random-walk Metropolis-Hastings step, we use a diagonal matrix. The element for ρ is a fraction of its prior variance, e.g., $0.2^2/10$, and for $\bar{s}_2$ we use $\beta_2^2/(\alpha_2-1)^2(\alpha_2-2)$, where α_2 and β_2 are defined in the same way as α_1 and β_1.

A.3 Data Set

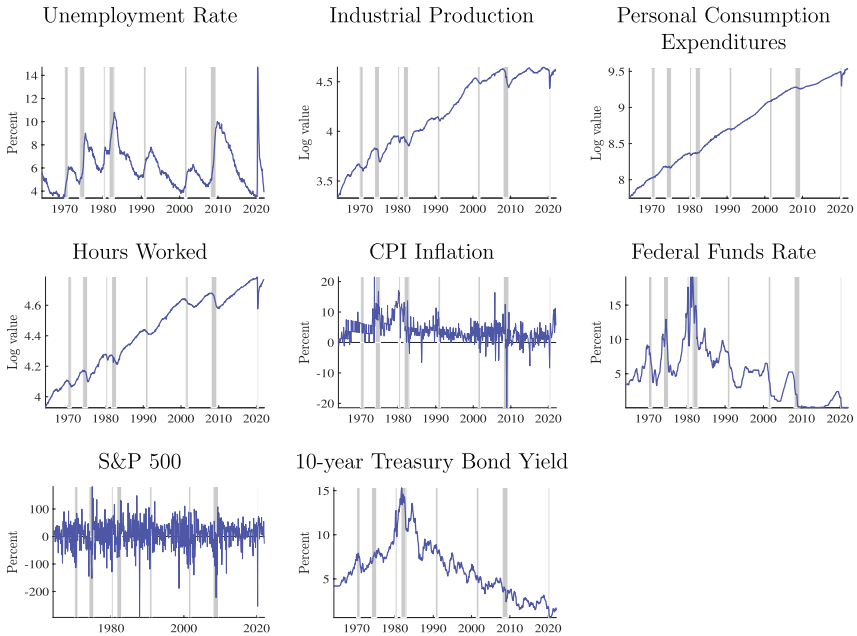
The 11 real-time macroeconomic data series are obtained from the ALFRED (Archival Federal Reserve Economic Data) database maintained by the Federal Reserve Bank of St. Louis. Table A.1 summarizes how the series used in this paper are linked to the series provided by ALFRED.

The recent vintages of PCE and INVFIX from ALFRED do not include data prior to 2002. However, the most recent data for PCE

and INVFIX can be obtained from BEA or NIPA tables. Specifically, we download "Table 2.8.3. Real Personal Consumption Expenditures by Major Type of Product, Monthly, Quantity Indexes" for PCE and "Table 5.3.3. Real Private Fixed Investment by Type, Quantity Indexes" for INVFIX, which are available from January 1, 1959 and January 1, 1948 to current periods, respectively. First, we compute the growth rates from the quantity indices. Based on the computed growth rates, we can backcast historical series up to January 1, 1964 using the January 1, 2002 data points as initializations. We think this is a reasonable way to construct the missing points for PCE and INVFIX.

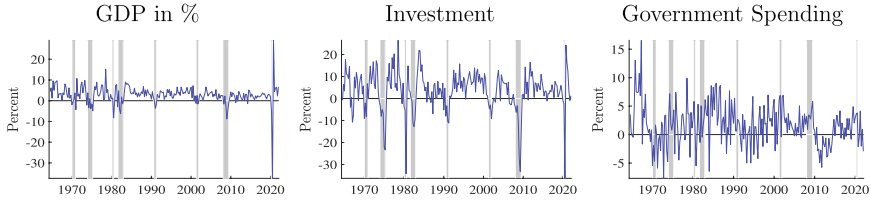
Figures A.1 and A.2 provide the time-series plot of our 11 macroeconomic variables obtained from the August 2021 vintage.

Figure A.1. Monthly Data



Note: Month-on-month percentage changes are annualized. The data are obtained from the January 2022 vintage, starting from 1964. The shaded bars indicate the NBER recession dates.

Figure A.2. Quarterly Data, Q-o-Q Growth Rates in Annualized %



Note: The data are obtained from the January 2022 vintage, starting from 1964. The shaded bars indicate the NBER recession dates.

References

Adrian, T., N. Boyarchenko, and D. Giannone. 2019. "Vulnerable Growth." *American Economic Review* 109 (4): 1263–89.

Alvarez, L. J., and F. Odendahl. 2021. "Covid-19 and Bayesian VARs in the Euro Area." Manuscript.

Antolin-Diaz, J., T. Drechsel, and I. Petrella. 2021. "Advances in Nowcasting Economic Activity: Secular Trends, Large Shocks and New Data." CEPR Discussion Paper No. 15926.

Bobeica, E., and B. Hartwig. 2023. "The Covid-19 Shock and Challenges for Inflation Modelling." *International Journal of Forecasting* 39 (1): 519–39.

Brave, S. A., R. A. Butters, and A. Justiniano. 2019. "Forecasting Economic Activity with Mixed Frequency BVARs." *International Journal of Forecasting* 35 (4): 1692–1707.

Carriero, A., T. E. Clark, M. Marcellino, and E. Mertens. 2022. "Addressing COVID-19 Outliers in BVARs with Stochastic Volatility." Forthcoming in *Review of Economics and Statistics.*

Clark, T. E. 2011. "Real-Time Density Forecasts from Bayesian Vector Autoregressions with Stochastic Volatility." *Journal of Business Economics and Statistics* 29 (3): 327–41.

Davis, R., and S. Ng. 2023. "Time Series Estimation of the Dynamic Effects of Disastertype Shocks." *Journal of Econometrics* 235 (1): 180–201.

Del Negro, M., and F. Schorfheide. 2011. "Bayesian Macroeconometrics." In *The Oxford Handbook of Bayesian Econometrics*, ed.

J. Geweke, G. Koop, and H. van Dijk, 293–389. Oxford University Press.

Diebold, F. X. 2020. "Real-Time Real Economic Activity: Exiting the Great Recession and Entering the Pandemic Recession." *Covid Economics* 62 (December): 1–19.

Doan, T., R. Litterman, and C. A. Sims. 1984. "Forecasting and Conditional Projection Using Realistic Prior Distributions." *Econometric Reviews* 3 (1): 1–100.

Durbin, J., and S. J. Koopman. 2001. *Time Series Analysis by State Space Methods.* Oxford University Press.

Foroni, C., M. Marcellino, and D. Stevanovic. 2020. "Forecasting the Covid-19 Recession and Recovery: Lessons from the Financial Crisis." CIRANO Working Paper No. 2020S-32.

Gandhi, M. A., and L. Mili. 2010. "Robust Kalman Filter Based on a Generalized Maximum-Likelihood-Type Estimator." *IEEE Transactions on Signal Processing* 58 (5): 2509–20.

Goulet Coulombe, P., M. Marcellino, and D. Stevanovic. 2021. "Can Machine Learning Catch the COVID-19 Recession?" CIRANO Working Paper No. 2021S-09.

Huber, F., G. Koop, L. Onorate, M. Pfarrhofer, and J. Schreiner. 2022. "Nowcasting in a Pandemic Using Non-parametric Mixed Frequency VARs." *Journal of Econometrics* 232 (1): 52–69.

Lenza, M., and G. E. Primiceri. 2022. "How to Estimate a VAR after March 2020." *Journal of Applied Econometrics* 37 (4): 688–99.

Lewis, D., K. Mertens, and J. Stock. 2020. "U.S. Economic Activity During the Early Weeks of the SARS-Cov-2 Outbreak." Working Paper No. 20-11, Federal Reserve Bank of Dallas.

Litterman, R. B. 1980. "Techniques for Forecasting with Vector Autoregressions." Ph.D. thesis, University of Minnesota.

Maroz, D., J. H. Stock, and M. W. Watson. 2021. "Comovement of Economic Activity During the Covid Recession." Manuscript.

McCracken, M. W., M. T. Owyang, and T. Sekhposyan. 2020. "Real-Time Forecasting and Scenario Analysis Using a Large Mixed-Frequency Bayesian VAR." Manuscript, Federal Reserve Bank of St. Louis.

Ng, S. 2021. "Modeling Macroeconomic Variations After Covid-19." NBER Working Paper No. 29060.

Primiceri, G. E., and A. Tambalotti. 2020. "Macroeconomic Forecasting in the Time of COVID-19." Manuscript, Northwestern University.

Schorfheide, F., and D. Song. 2015. "Real-Time Forecasting with a Mixed-Frequency VAR." *Journal of Business and Economic Statistics* 33 (3): 366–80.

———. 2020. "Real-Time Forecasting with a (Standard) Mixed-Frequency VAR During a Pandemic." Working Paper No. 20-26, Federal Reserve Bank of Philadelphia.

Sims, C. A., and T. Zha. 1998. "Bayesian Methods for Dynamic Multivariate Models." *International Economic Review* 39 (4): 949–68.