Addressing COVID-19 Outliers
in BVARs with Stochastic Volatility[*]

Andrea Carriero,[1] Todd E. Clark,[2]
Massimiliano Marcellino,[3] and Elmar Mertens[4]

[1]*Queen Mary University of London, and University of Bologna,* [2]*Federal Reserve Bank of Cleveland,*
[3]*Bocconi University, CEPR, IGIER, BIDSA, and BAFFI,* [4]*Deutsche Bundesbank.*

April 15, 2022

Abstract

The COVID-19 pandemic has led to enormous data movements that strongly affect parameters and forecasts from standard Bayesian vector autoregressions (BVARs). To address these issues, we propose VAR models with outlier-augmented stochastic volatility (SV) that combine transitory and persistent changes in volatility. The resulting density forecasts are much less sensitive to outliers in the data than standard BVARs. Predictive Bayes factors indicate that our outlier-augmented SV model provides the best data fit for the pandemic period, as well as for earlier subsamples of relatively high volatility. In historical forecasting, outlier-augmented SV schemes fare at least as well as a conventional SV model.

Keywords: Bayesian VARs, stochastic volatility, outliers, pandemics, forecasts

JEL classification codes: C53, E17, E37, F47

---

1 Introduction

Bayesian vector autoregressions (BVARs) have a successful track record in macroeconomic fore-
casting and structural analysis. However, economic turbulence created by the ongoing COVID-19
pandemic has posed some basic challenges for estimation and inference with BVARs. As exam-
ples of the extreme variability, payroll employment plummeted by about 15 percent from March
to April 2020, a decline nearly 16 times as large as the previous largest monthly decline, and
real income rose by about 12 percent in the month, an increase 3 times larger than the previous
record growth rate. These extreme realizations can have strong effects on parameter estimates
and forecasts generated by conventional constant-parameter BVARs. In response, Schorfheide and
Song (2021) suggest ignoring the recent data in estimating BVAR parameters, whereas Lenza and
Primiceri (2022) propose a specific form of heteroskedasticity, tuned to the COVID-19 data, to
down-weight observations since March 2020 in the estimation.

Prior to the pandemic, BVARs with stochastic volatility (SV) provided more accurate point and
density forecasts than constant-parameter models (see, e.g., Clark (2011), Clark and Ravazzolo
(2015), and D'Agostino, Gambetti, and Giannone (2013)). SV models generate time variation
in predictive densities through changes in the variance-covariance matrix of the BVAR's forecast
errors over time. The heteroskedasticity in the form of time-varying error variances also affects
the estimation of slope coefficients in the BVAR (at least in finite samples). As an application of
generalized least squares, when extreme realizations are modeled as sudden increases in volatility,
heteroskedastic BVARs will down-weight the associated observations when estimating parameters.

A typical SV model assumes changes in volatility to be highly persistent. However, by defi-
nition, extreme observations are more reflective of short-lived spikes, not permanent increases, in
volatility. Like Lenza and Primiceri (2022) and Schorfheide and Song (2021), we view the extreme
observations of the COVID-19 period as possible outliers that are characterized by transient and
infrequent increases in volatility, in which case it may be desirable to reduce their influence on
model estimates and forecast distributions.

This paper develops BVAR models with SV that feature combinations of (1) large but infre-

quent volatility outliers and (2) fat-tailed errors. For the infrequent, large volatility outlier, we adopt a discrete mixture representation that Stock and Watson (2016) used in unobserved component models of inflation to accommodate extreme volatility during the global financial crisis. The Stock-Watson model augments the standard SV specification of a highly persistent volatility state with an outlier state, acting as scale factor for volatility, that infrequently and temporarily jumps to values above 1. For the treatment of fat-tailed (rather than Gaussian) errors in SV, we adopt the Jacquier, Polson, and Rossi (2004) specification of t-distributed innovations.

With these building blocks, we consider one BVAR with SV specification (SVO) that features the infrequent volatility outliers but Gaussian errors. We also consider — and prefer, for reasons indicated below — a specification (SVO-t) that has both the infrequent volatility outliers and fat-tailed errors. We emphasize that our approach is data-based: Our models provide probabilistic assessments of the timing and scale of realized outliers in the data; we are not simply restricting recent observations to be outliers. In empirical analysis of the models, we use a medium-sized data set of 16 monthly variables, in keeping with existing evidence of the forecast accuracy advantages of medium-sized models (e.g., Carriero, Clark, and Marcellino (2019) and Koop (2013)).

Our empirical results with common macroeconomic time series show the efficacy of our proposed SVO and SVO-t specifications for mitigating the influence of COVID-induced outliers on estimates and forecasts, fitting the data of not only the pandemic period but also earlier periods, and forecasting out-of-sample in a long period preceding the pandemic. Although both models succeed in mitigating the influence of COVID-induced outliers and the SVO model fits historical data better than the slightly more complicated SVO-t specification, the advantages of SVO-t over SVO in forecasting lead us to favor it in our analysis and recommendation.

More specifically, as a starting point for our empirical work, we confirm the findings of Lenza and Primiceri (2022) and Schorfheide and Song (2021) that forecasts generated since March 2020 from homoskedastic BVARs are often distorted.[1] In general, the recent outliers cause the forecast paths of some variables to become extreme by historical standards. Instead, BVARs with time-varying volatility generated better-behaved forecasts. SV, SVO, and SVO-t estimates all register

increases in forecast uncertainty. But while the SV specification sees all shocks to forecast uncertainty as permanent, the SVO and SVO-t models explicitly allow for one-off spikes in volatility, resulting in estimates of forecast uncertainty that are still elevated but, in our subjective assessment, appear less extreme and more reasonable.

As an alternative, we also consider relying on a standard BVAR-SV model but treating as missing data those observations identified ex-ante as extreme. The methods discussed so far adjust parameters (including the volatility states) but not the data vector used at the forecast origin in forming a prediction; treating observations as missing data also alters the jumping-off point of the forecasts. To identify extreme observations as outliers, we use an ex-ante criterion known from the literature on dynamic factor models that is based on the distance of a given data point from the time-series median.[2] This approach differs from the SVO and SVO-t approaches, which estimate the occurrence of outliers jointly with the BVAR, by treating the dates of outliers as known ex-ante. In addition, the missing-data treatment remains agnostic about the specific stochastic properties of those observations that are pre-selected as outliers. In the COVID-19 period, this approach also produces much better-behaved forecasts than a constant-variance BVAR. In forecasting, the biggest difference with the outlier-augmented SV procedures is that conditioning on the incidence of outliers, while otherwise ignoring any signal from their specific realization, leads to predictive densities that can be considerably tighter than those from SVO and SVO-t.

To evaluate which model best characterizes the data in the COVID-19 period and earlier, we employ predictive Bayes factors (which are based on sums of predictive likelihoods). By this measure, our SVO specification fits the COVID-19 sample the best, with SVO-t next. In earlier samples, the SVO and SVO-t specifications also fare well in model fit (with SVO ahead). The advantages of these models are driven by the subsamples of relatively high volatility; the baseline SV model fits best in the Great Moderation years of 1985 through 2007.

Although to this point we have focused on efficacy in reducing outlier-induced distortions to forecast distributions, to be broadly effective, it is important that a given method not only helps reduce such distortions but also forecasts effectively over long periods less affected by outliers.

Accordingly, we conduct an historical assessment of out-of-sample forecast performance with an evaluation window starting in 1975 and ending in 2017, comparing the accuracy of point and density forecasts. It turns out that pre-COVID data include outliers; indeed, SVO and SVO-t detect pre-COVID-19 outliers in macroeconomic and financial time series, whose existence had been noted before by, among others, Stock and Watson (2002). In forecast accuracy for 1975-2017, the SVO-t approach modestly outperforms SV. In results presented in a supplementary online appendix, the SVO-t specification also has one advantage over the SVO model: At longer horizons, the SVO-t model is modestly better than the SVO in density forecast accuracy. It is this advantage that leads us to recommend and focus on the SVO-t model over the SVO specification.

The remainder of this paper proceeds as follows. Section 2 briefly reviews the related literature not covered above. Section 3 introduces the models and describes their estimation. Section 4 describes the data used. Section 5 provides our results. Section 6 summarizes robustness checks provided in our supplementary online appendix. Section 7 concludes.

2 Related Literature

As noted above, the arrival of COVID-19 has prompted a number of studies to consider treating the extreme observations of the COVID-19 period as outliers. A particular contribution of our paper is the comprehensive analysis of forecast performance and model fit over a wide set of macroeconomic and financial variables of BVARs with and without outlier-augmented SV. By studying model performance over a relatively long sample of post-war US data, we can also document the recurring benefits of outlier treatments at times of crisis or other economic upheavals.

Antolín-Díaz, Drechsel, and Petrella (2021) develop a dynamic factor model for nowcasting, with outliers modeled as additive measurement errors that have student-t distributions. Focusing on euro area inflation, Bobeica and Hartwig (2022) document that pandemic observations can shift parameter estimates and find some benefits to allowing fat tails in a BVAR for the euro area. In another application to euro area data, Alvarez and Odendahl (2021) find that the pandemic's

outliers distort BVAR estimates and consider alternative approaches to modeling volatility outliers.

Prior to the arrival of COVID-19, some studies had already considered BVAR specifications with fat-tailed error distributions. For example, t-distributed shocks were used in BVAR-SV models by Chiu, Mumtaz, and Pintér (2017) and Clark and Ravazzolo (2015) and estimated DSGE models, with and without SV, by Cúrdia, Del Negro, and Greenwald (2014) and Chib, Shin, and Tan (2021). Karlsson and Mazur (2020) and Chan (2020) provide general treatments of heteroskedasticity in BVAR models with and without SV and fat-tailed error distributions.

Other recent analyses have proposed approaches more geared to specific circumstances of the pandemic and the estimation of causal (or structural) dependencies. For example, Primiceri and Tambalotti (2020) and Ng (2021) argue for seeing the COVID-19 period as adding a new type of shock to the dynamic system of the economy. Assuming that the new COVID-19 shock has been the dominant source of variation since early 2020, Primiceri and Tambalotti (2020) derive a set of conditional forecasts for different scenarios of future developments. Instead, Ng (2021) uses pandemic indicators to "de-covid" data prior to estimation of time series models. Specifically, in application to a structural VAR, Ng (2021) shows that after accounting for exogenous COVID-19-related indicators, dynamic responses to other shocks appear similar pre- and post-COVID-19.

3   BVAR Models

We study BVAR models of the following form:

$$y_t = \Pi_0 + \Pi(L)y_{t-1} + v_t\,, \qquad\qquad v_t \sim N(0, \Sigma_t), \qquad\qquad (1)$$

where y_t is a vector of N observables, $\Pi(L) = \sum_{i=1}^{p} \Pi_i \ L^{i-1}$ is a pth-order lag polynomial of VAR coefficients, and v_t denotes the VAR's residuals. We denote the vector of stacked coefficients contained in $\{\Pi_i\}_{i=0}^{p}$ as Π. Throughout, we maintain the assumption of time-invariant transition coefficients Π, which is commonly (and so far successfully) used in forecasting.[3] All models are specified with non-conjugate priors for Π and Σ_t. The models differ mainly in whether the residuals

are homoskedastic, or in the form of their heteroskedasticity. Note that, in the context of BVAR models, homoskedasticity refers to treating Σ_t as constant over time, whereas heteroskedasticity refers to treating Σ_t as varying over time, with particular stochastic structures so that v_t can be seen as mixed Gaussian instead of Gaussian (in our context, heteroskedasticity does not refer to the conditional variance of v_t depending on regressors).

As we show, our models featuring outlier states share a latent state representation in which residuals are the product of Gaussian shocks and outlier states, but differ in the assumed densities for the outlier states. One model (SVO) puts more mass on outliers being large events that increase volatility by more than twofold, whereas another (SVO-t) sees fewer large outliers and more frequent small outliers. Conventional Markov chain Monte Carlo (MCMC) estimation procedures for BVAR-SV models can easily be extended to handle the SVO and SVO-t models, with two extra steps. First, realized outlier states are drawn from their posteriors, conditional on draws for each variable's outlier probability. Second, the outlier probability for each variable is drawn from a (conditional posterior) distribution conditional on the draws of the time series of outlier states.

As noted in the introduction, time-varying volatility in the BVAR residuals, v_t, can help to insulate estimation of the transition coefficients Π from the effects of extreme outliers. Intuitively, observations with higher residual volatility receive less weight in the estimation of BVAR coefficients. However, down-weighting extreme observations in the estimation of Π will not completely insulate the resulting forecasts from outliers. Consider the simple case of an AR(1) model without intercept, where the h-step-ahead forecast is given by $y_{t+h|t} = \pi^h y_t$ and y_t was an outlier. Even if the outlier were excluded from estimation of π, it would still have a direct effect on the forecast $y_{t+h|t}$. To address these concerns, we also consider a variant of the SV model that treats pre-specified outliers as missing values, in a way described below.

3.1 Model Specification

We consider the following five variants of the BVAR model (1). The first four differ in the specified process for the residuals v_t, whereas the last variant treats pre-specified outliers as missing data.

Density forecasts will crucially depend on the assumed dynamics of the variances in Σ_t, and we consider different forms of persistence in variance changes, detailed below.

1) CONST: A homoskedastic BVAR with $v_t \sim N(0, \Sigma)$.

2) SV: In this baseline SV model, the VAR residuals can be written as

$$v_t = A^{-1} \Lambda_t^{0.5} \varepsilon_t, \qquad \text{with} \quad \varepsilon_t \sim N(0, I), \qquad (2)$$

where A^{-1} is a unit-lower-triangular matrix, $\Lambda_t^{0.5}$ is a diagonal matrix of stochastic volatilities, and $\Sigma_t = A^{-1} \Lambda_t (A^{-1})'$. The vector of logs of the diagonal elements of Λ_t, denoted $\log \lambda_t$, evolves as a random walk with correlated errors:

$$\log \lambda_t = \log \lambda_{t-1} + e_t, \qquad \text{with} \quad e_t \sim N(0, \Phi). \qquad (3)$$

3) SVO: The SVO model is intended to capture outliers as rare, transitory, and large changes in volatility. The outliers enter the model in a diagonal matrix of scale factors, denoted O_t, with diagonal elements $o_{j,t}$ that are mutually $i.i.d.$ over all j and t. The outlier $o_{j,t}$ has a two-part distribution that distinguishes regular observations with $o_{j,t} = 1$ and outliers for which $o_{j,t} \geq 2$.[4] Outliers in variable j, $j = 1, \ldots, N$, occur with probability p_j and the distribution:

$$o_{j,t} = \begin{cases} 1 & \text{with probability} \quad 1 - p_j \\ U(2, 20) & \text{with probability} \quad p_j, \end{cases}$$

where $U(2, 20)$ denotes a uniform distribution with support between 2 and 20. Conditional on O_t and Λ_t, the VAR residuals are Gaussian in the SVO model. With A^{-1} and $\Lambda_t^{0.5}$ specified as before, the vector of residuals and its covariance matrix take the forms:

$$v_t = A^{-1} \Lambda_t^{0.5} O_t \varepsilon_t, \quad \Sigma_t = A^{-1} O_t \Lambda_t O_t' (A^{-1})'.$$

Of course, variations on this specification of outliers are also possible; some of the choices underlying the specification are convenient but unlikely to be essential. For example, we view the discrete grid for outlier values as a convenient and tractable approximation for feasibility. Perhaps more important is treatment of outliers as rare and large, so that the distribution for the outlier state has a mode at 1 but a large tail with values substantially above 1. As to the assumed independence over time, one might model outliers as having some serial correlation (e.g., as in the outlier treatment of Lenza and Primiceri (2022)), but to do so could create challenges — especially with the modest sample sizes in macro applications — in distinguishing persistent changes in volatility through the SV state from less-persistent but serially correlated changes through an outlier state.

4) SVO-t: The SVO-t model extends the SVO specification to include one state capturing rare jumps in volatility and a second state that captures transitory changes in volatility that are more frequent but less extreme in impact (consistent with draws from the tails of a fat-tailed distribution).[5] Each kind of outlier enters the model in a diagonal matrix of scale factors, denoted O_t and Q_t, with diagonal elements $o_{j,t}$ and $q_{j,t}$, respectively, that are mutually $i.i.d.$ over all j and t.

The first kind of outlier, $o_{j,t}$, is the same as that of the SVO model, with a two-part distribution that distinguishes regular observations and outliers. The second, less extreme, type of outlier in the SVO-t model is equivalent to having t-distributed VAR residuals (conditional on Λ_t and O_t). Following Jacquier, Polson, and Rossi (2004), we let the squares of the diagonal elements of Q_t, $q_{j,t}$, have inverse-gamma distributions:

$$q_{j,t}^2 \sim IG\left(\frac{d_j}{2}, \frac{d_j}{2}\right).$$

The vector of VAR residuals in the SVO-t model and its covariance matrix take the forms:

$$v_t = A^{-1}\,\Lambda_t^{0.5}\,O_t\,Q_t\,\varepsilon_t,\ \ \Sigma_t = A^{-1}\,O_t\,Q_t\Lambda_t\,Q_t'\,O_t'(A^{-1})',$$

with A^{-1}, $\Lambda_t^{0.5}$, and O_t specified as before. The jth residual $q_{j,t}\cdot\varepsilon_{j,t}$ (adjusted for the rotation

by A^{-1} and scaling by $\Lambda_t^{0.5} O_t$), has a student-t distribution with d_j degrees of freedom, since $\varepsilon_{j,t} \sim N(0,1)$ and $d_j/q_{j,t} \sim \chi^2_{d_j}$.

5) SV-OutMiss: This model applies the standard SV specification for Σ_t, but ignores a given set of outlier observations in the BVAR estimation altogether by treating them as missing data. The approach builds on a practice from the literature on dynamic factor models, in which input data are pruned of extreme observations that are multiples times the inter-quartile range away from the series median (by replacement with a moment of central tendency). We adopt the same ex-ante criterion for the identification of outliers — implemented using a threshold factor of 5 (with similar results for a factor of 10) — and treat these observations as missing data in estimation and forecasting. Treating pre-identified outlier observations as missing data avoids specification of their exact stochastic distribution.[6] For each missing value, our Bayesian methods generate a posterior distribution that informs the resulting forecasts. Formally, denote the history of y_t after pruning outliers as z^t, and continue the AR(1) example introduced above: Forecasts are then generated by $y_{t+h|t} = \pi^h E(y_t|z^t)$, where $E(y_t|z^t)$ is identical to y_t in the no-outlier case. Similarly, forecast uncertainty is generated based on estimates of SV that condition only on z^t, not potential outliers in the history of y_t.

3.2 Model Estimation

Throughout, our BVARs include $p = 12$ lags in a monthly data set, which is described in further detail in Section 4. Each of our models is estimated with a MCMC sampler, based on the methods of Carriero, Clark, and Marcellino (2019) (henceforth "CCM") for estimating large BVARs, but as corrected in Carriero, et al. (2022). As in CCM, we use a Minnesota prior for the VAR coefficients Π and follow their other choices for priors as far as applicable, too.

For the infrequent outlier components of the SVO and SVO-t models, we follow Stock and Watson (2016) in placing a beta prior on the outlier probability p_j. The prior is set to imply a mean outlier frequency of once every 4 years in monthly data for SVO estimates and once every 10 years

for SVO-t estimates, with precision set to be consistent with 10 years' worth of prior observations. For the t-distributed component of the SVO-t model, we follow Jacquier, Polson, and Rossi (2004) and estimate the degrees of freedom d_j for each variable using a uniform discrete prior with a range of 3 to 40.[7]

Here we briefly explain the algorithm adjustments needed for the version of the model with constant variance and the alternative with outlier volatility states. The algorithm includes all of the same steps given in CCM (as corrected in Carriero, et al. (2022)), except for necessary adjustments to account for the two alternative cases. For the constant-volatility model, an inverse-Wishart prior for Σ, with a (conditionally) conjugate inverse-Wishart updating step for the MCMC sampler, replaces the SV block of the model.[8]

For the SVO-t variant, the following extra steps are added to the original BVAR-SV setup: Realized outlier states $o_{j,t}$ and $q_{j,t}$ need to be drawn from their posteriors. The step for $o_{j,t}$ conditions on draws for the outlier probability p_j and proceeds analogously to the sampling of the mixture states needed with the Kim, Shephard, and Chib (1998) approach to the stochastic volatility states $\log \lambda_t$. The step for $q_{j,t}$ takes a draw from an inverse Gamma distribution. A further additional step draws the outlier probability p_j for each variable from a (conditional posterior) beta distribution conditional on the draws of the time series of outlier states. The algorithm for SVO is a simplified version of that for SVO-t.[9]

For the SV-OutMiss model, which treats pre-specified outliers as missing values, the MCMC sampler for the standard SV model is augmented by an additional step that draws the missing values from a state-space representation of the BVAR system using the disturbance smoothing algorithm of Durbin and Koopman (2002). Computational cost increases substantially with the SV-OutMiss model, as it requires an additional sequence of Kalman filtering and smoothing steps. In contrast, the added cost of computing SVO-t or SVO over standard SV is small, since this model adds only steps for sampling the $i.i.d.$ outlier states.

All results in the paper are based on 1,000 retained draws, obtained by sampling a total of 1,200 draws with 200 burn-in draws. Unreported comparisons of posteriors obtained under different

starting values indicate satisfactory convergence of the MCMC algorithms.

4 Data

Our data set consists of monthly observations for 16 macroeconomic and financial variables for the sample March 1959 to March 2021, taken from the April 2021 vintage of the FRED-MD database maintained by the Federal Reserve Bank of St. Louis. The variables and their transformation to logs or log-differences are listed in Table 1. To avoid issues related to the effective lower bound (ELB) on nominal interest rates, the data set includes only longer-term interest rates and omits a policy rate measure, like the federal funds rate, which was constrained by the ELB from late 2008 to 2016, and then again starting in March 2020.[10]

[Table 1 about here.]

In keeping with some work in the factor model literature, the prevalence of outliers can be roughly gauged by defining an outlier as an observation with distance from the series median exceeding 5 times the inter-quartile range. As detailed in the supplementary online appendix, real personal income has regularly displayed outliers over the post-war sample. Many other series, like payroll growth, exhibit such outliers only over the COVID-19 period, whereas a few others, like returns on the S&P500, inflation, or the exchange rate between the US dollar and pound sterling, displayed large outliers only on earlier occasions. Some variables, like the unemployment rate, have registered outstanding changes since the pandemic's outbreak, but without registering explicit outliers by this metric. In some cases, outliers may be attributed to specific unusual events. For example, industrial production registered a positive outlier in December 1959, when production bounced back following a strike in the steel industry. More recently, income transfers from the CARES Act caused growth in personal income to surge in April 2020.

11

5 Results

This section presents results on outlier estimates, forecast performance over the pandemic period of 2020-21, model fit, and forecast accuracy pre-COVID-19.

5.1 Outlier Estimates In 2020-21 and Before

As described in Section 3, the SVO-t approach extends the baseline SV model by adding latent outlier states $o_{j,t}$ and $q_{j,t}$ for each variable $j = 1, \ldots, N$, with the former uniformly distributed and squares of the latter having an inverse Gamma distribution. The outlier states enrich the dynamics of the time-varying variance-covariance matrix, Σ_t, so that volatility can change due to transitory changes in $o_{j,t}$ and $q_{j,t}$, as well as the persistent variations induced through the log-SV terms $\log \lambda_t$. The SVO model adds just the state $o_{j,t}$ to a SV model.

The supplementary online appendix reports posterior estimates of the probabilities of large outliers in the SVO and SVO-t models and for the degrees of freedom for the fat-tail components of the SVO-t specification. In the SVO model, the posterior mean probability of a large outlier is greatest for real income, at 3.19 percent, and ranges from about 0.3 percent (housing starts) to 1.1 percent (nonfarm payrolls and hours) for other variables. In the SVO-t specification that allows for both small and large outliers, the posterior mean estimate of the degrees of freedom is 3 for about one-half of the model's variables — implying frequent small outliers — but above 20 (near-Gaussian) for six other variables. In all cases, the estimated probabilities of large outliers are sharply lower than in the SVO model.

We can also provide a closer comparison of the volatility and outlier estimates obtained from SVO-t and SVO. Focusing on just real income and S&P500 returns in the interest of chart readability, Figure 1 displays posterior medians of the SV component (i.e., $\lambda_{j,t}^{0.5}$) and outlier estimates ($o_{j,t}$ and $q_{j,t}$) obtained over the full sample, with dark solid lines depicting the actual forecast error volatility, including outlier components and shaded areas showing the persistent SV component. Echoing our discussion of each model's properties in Section 3, these results show that the SVO-t

specification tends to see outliers as being more moderately sized but occurring also more regularly than SVO. For example, in the real income estimates, SVO-t shows a relatively large number of outliers in the 1970s and 1980s, whereas SVO shows fewer outliers that are larger in size. With S&P500 returns, SVO shows few outliers before the COVID-19 period, whereas the SVO-t estimates yield relatively regular, small outliers, with more variability in the SV estimate ($\lambda_{j,t}^{0.5}$) in the SVO case than the SVO-t case.

[Figure 1 about here.]

Time variation in Σ_t affects forecasts through two channels: first, the estimation of BVAR coefficients Π as discussed in Section 3; and second, the projection of uncertainty about future shocks v_t that arises when simulating forward the dynamics of $\log(\lambda_t)$, as given in (2) and (3), to construct predictive densities. Historical forecast results for 1975 to 2017, discussed below, suggest that the latter channel is more relevant than the former, as the point forecast accuracy differences between SV and SVO-t are very small, while the density accuracy differences are sometimes larger. The outlier states in SVO-t (as well as SVO) allow for volatility spikes to occur without having to project a persistent increase in uncertainty into the future as SV would be required to do. To illustrate the effects of this feature, we compare trajectories of volatility as estimated in quasi-real time over the course of 2020 and early 2021.[11]

[Figure 2 about here.]

Focusing on the example of payroll growth to limit charts, Figure 2 reports estimates of time variation in the volatility of forecast errors generated by SV and SVO-t, as well as the persistent components of Σ_t imputed from SVO-t when the effects of the outlier states $o_{j,t}$ and $q_{j,t}$ are ignored. (The online appendix provides results for other variables.) For this counterfactual, we compute $\tilde{\Sigma}_t = A^{-1} \Lambda_t (A^{-1})'$ based on the SVO-t estimates for Λ_t and A^{-1}. In addition, we consider the corresponding measures of residual volatility obtained from the SV-OutMiss model that treats pre-specified outliers as missing data. These estimates show that, over the COVID-19 period, the

SVO-t model clearly differentiates between increases in uncertainty that are short- and longer-lived, which the SV model cannot do. In early 2020, prior to the impact of COVID-19, volatility estimates from all models were hovering below 10. By April, volatility estimates from the SV model increased strongly to a peak near 60, but leveled off only somewhat over the summer, and remained substantially elevated in the fall, near values around 20 in estimates using data through September 2020. Crucially, at each point in time, the SV model expects these levels to persist.

In contrast, SVO-t proves both more nimble and more discerning in accounting for the extreme data seen in the spring with a big jump in overall volatility in April, to a peak of about 90, as shown in Panel (b) of the figure. However, as revealed by comparison with Panel (d), this jump is largely seen as a transitory result of an outlier (both as it occurred in the spring and with the hindsight of estimates constructed based on data for the fall). In contrast, in Panel (d) the persistent component of volatility in the case of SVO-t is seen to have risen no more than 8-fold over the course of the year, to a peak of roughly 12 before declining. That is, the SVO-t estimates yield a much smaller rise in the persistent component of volatility than do the estimates from the SV model. The SV-OutMiss model yields an even smaller increase in the persistent component of volatility (the only component of volatility in that model); the estimates from SV-OutMiss shown in Panel (c) have risen by less than 5 times their level at the beginning of the year, peaking at a variance of about 8 in April 2020.

The more moderate rise in estimates of the persistent volatility component obtained with the SVO-t specification yields noticeably narrower (and arguably less extreme) uncertainty bands around forecasts compared to the SV model. In contrast, forecasts that condition on knowledge of when outliers occurred, but otherwise ignore any further information from their realization (as in the SV-OutMiss case), lead to particularly narrow uncertainty bands.

5.2 Forecasts Made In 2020-21

In the months immediately preceding the COVID-19 outbreak, such as January 2020, predictive densities generated from the CONST, SV, SVO, and SVO-t models differ a little, but not markedly

14

so for most variables. As we now detail, the picture changed significantly in subsequent months.

[Figure 3 about here.]

Over the course of March and April, the COVID-19 pandemic sharply affected the economy, most visibly with the introduction of lockdown measures in the second half of March 2020, resulting in strong swings among measures of real activity in subsequent months. Figure 3 displays the evolution of forecasts for real income and payroll growth over the months of March, April, and June generated from our CONST, SV, and SVO-t specifications.[12] As noted by Lenza and Primiceri (2022) and Schorfheide and Song (2021), forecasts generated by homoskedastic BVARs, like our CONST specification, can display extreme behavior.[13] For example, Panel (d) shows that, following the drop in payroll growth in March and April, the CONST model's posterior median forecast for May is about -136 percent (at an annualized rate) and between -64 and -124 percent for the next few months. The model's estimated forecast uncertainty is immense, with a 68 percent uncertainty band that widens to 100 percentage points or more by the 12-months-ahead horizon.

In contrast, the reaction of point and density forecasts generated by the SV and SVO-t specifications to the incoming data in spring 2020 is better behaved, particularly with SVO-t. Considering again the payroll growth forecasts shown in Figure 3, the SV model yields very negative point forecasts for May and the next few months, but not nearly as negative as those from the CONST model (e.g., the posterior median forecast for May is -17.8 percent and -20.1 percent for the SV and SVO-t models, respectively). The SVO-t model yields point forecasts fairly similar to those of the SV model, for most variables and forecast origins. That said, the SV model is prone to some distortion of its estimated forecast uncertainty, particularly early in the COVID-19 period. In March, April, and June of 2020, the uncertainty bands of the predictive densities obtained with SV are typically wider than those of not only the SVO-t but also the CONST specifications. In keeping with the volatility comparisons provided above, while the observations of 2020 widen the predictive densities of both SV and SVO-t forecasts, their impact is much greater for the former than for the latter; SVO-t generates much narrower bands than SV.

Supplementary results in the online appendix compare our preferred SVO-t results to those for the more restrictive SVO specification.[14] While the point forecasts of these specifications are difficult to distinguish, bigger differences are evident in the predictive densities. The predictive densities are generally the narrowest with the SVO-t forecasts. The SVO model generally yields wider densities, although in most cases the differences are less stark in June than March and April.

[Figure 4 about here.]

In additional forecast results for the pandemic period, we compare results from the SVO-t specification (which treats outliers as unknown and estimates them) to results from the SV-OutMiss approach that conditions on knowledge of when and which outliers occurred in the data. As described above, outliers are observations that are more than 5 times the inter-quartile range away from their sample median. SV-OutMiss treats these observations as missing data in estimation of the parameters and volatility states of an otherwise standard BVAR-SV model and also replaces the outliers in the data vectors used to simulate predictive densities at every forecast origin.

For these specifications, Figure 4 provides predictive densities for more recent forecast origins, ranging from September 2020 to March 2021, for growth in payrolls and the unemployment rate.[15] Even almost a year after the onset of the COVID-19 pandemic impacted economic data, uncertainty bands from SVO-t remain noticeably wider than before the pandemic (results omitted in the interest of brevity). In most cases, forecast densities obtained from SV-OutMiss, which treats the timing of outliers as known, are relatively tight. However, exceptions are evident in the unemployment rate forecasts, with the SV-OutMiss bands wider than those of SVO-t for forecasts made with data in September and December 2020. Although harder to discern in the wide scales of the charts necessitated by the extreme realizations of actual data, the point forecasts produced by the alternative methods tend to be broadly similar at longer forecast horizons, although more sizable differences can occur at shorter horizons.

5.3 Model Fit

So which model best characterizes the data in the COVID-19 period? The COVID-19 sample is too short to permit meaningful inference on the average accuracy of out-of-sample forecasts. Drawing on precedents such as Geweke and Amisano (2010), we instead consider the basic metric of predictive Bayes factors: the sums of 1-step-ahead predictive likelihoods. In these comparisons, we take the SV specification as the baseline and report sums of differences in predictive likelihoods, such that the more positive (negative) the number, the better (worse) the fit of a given specification compared to SV (to facilitate comparisons over time, Table 2 includes in parentheses average score differences across time). Particularly with unusual observations, some care in computing predictive scores is warranted. We follow the recommendations of Krüger, et al. (2021) and use what they characterize as a mixture-of-parameters approach. As an instance of Rao-Blackwellization, the approach relies as far as possible on the availability of analytical expressions for predictive likelihoods conditional on parameter values and latent SV states at each MCMC draw. In computational accuracy, we find it to be particularly important to integrate out future values of the transitory outlier states, instead of characterizing their arrival via Monte Carlo simulation. The supplementary online appendix provides further details on the calculations for each model.

A first issue is how the models compare by this measure of model fit over the COVID-19 sample of March 2020 through February 2021. These estimates are provided in the last row of Table 2. Over this sample, the best fitting model is SVO, followed by the SVO-t specification. In an overall fit sense, the data seem to favor a specification allowing infrequent, large outliers, and the data imply that the fit gain over the SV baseline is large. The SV-OutMiss approach that rests on identifying outliers ex-ante fits the data of the COVID-19 period much worse, with a score difference on the order of -950 log points. Perhaps not surprisingly, the CONST specification fares the worst over this volatile period. By design, the large advantages of the SVO and SVO-t specifications over the COVID-19 period are primarily driven by the first few, most dramatic months of the pandemic; as shown in the last row of entries, when March through June 2020 are omitted, the fits of these models are very similar to those of the SV baseline. In earlier periods,

these models also gained sizable fit advantages with large data movements.

The consideration of the COVID-19 period of course raises the question of how, earlier in time, the specifications compare in model fit. For the sample running from 1975 (when our out-of-sample forecast evaluation of Section 5.4 begins) into 2021, the patterns in model fit line up with those for the COVID-19 period, but with a bigger advantage of the SVO model. The SVO model also fares best in two other periods known for relatively high economic volatility: the 1975-1984 period coinciding with what some have referred to as the Great Inflation and the 2008-2014 sample of the Great Recession and ensuing slow recovery. The SVO-t model again has the second best score in the 2008-2014 period, but slips to third best in the 1975-1984 sample. On a per-period basis, the biggest fit advantage of the SVO model over the SVO-t specification occurs in the COVID-19 sample, when the per-period fit advantage was 0.60 (6 percent), whereas in the Great Inflation, the per-period fit advantage of SVO-t over SVO was 0.21 (2.1 percent). In contrast, over the relatively tranquil period of 1985-2007, key years of the Great Moderation, the benchmark SV specification fits best. SV-OutMiss fits the data next-best, because there are few outliers, so that this approach is a small departure from SV. Among the models featuring some form of SV, allowing frequent, small outliers in the SVO-t model fit the data worst, with SVO and its large, infrequent outliers not as far off the SV benchmark. Overall, our approach of extending a SV model to allow infrequent outliers works well by the metric underlying predictive Bayes factors, achieving its gains in the several historical subsamples that have featured high volatility.

[Table 2 about here.]

5.4 Forecast Performance pre-COVID-19

Although our focus is on models that successfully mitigate the influence of enormous data movements in the COVID-19 pandemic on parameters and forecasts obtained from standard BVARs, applicability of the outlier-augmented SVO and SVO-t models is not necessarily specific to data from the pandemic. As noted above, individual data series have exhibited occasional outliers

18

before, leading to some earlier studies of the potential benefits of modeling fat-tailed error distributions and other forms of outliers. Importantly, the preceding results on model fit show that the SVO and SVO-t models have advantages over other models in earlier periods. But the model fit measure is based on 1-step ahead predictive likelihoods, which leaves open the question of how the models compare in historical forecast accuracy at longer horizons.

Accordingly, this section provides an evaluation of out-of-sample forecasts made from January 1975 through December 2017. For brevity, we focus on the forecast accuracy of our SVO-t specification compared to a conventional SV model; the supplementary online appendix provides additional comparisons. Specifically, for every forecast origin, each model is re-estimated based on growing samples of data that start in March 1959. All data are taken from the April 2021 vintage of FRED-MD; we abstract from issues related to real-time data collection. The forecast horizons considered extend from 1 to 24 months. We evaluate point and density forecasts based on root-mean-squared errors (RMSE) and continuous ranked probability scores (CRPS), respectively, as described in, among others, Clark and Ravazzolo (2015) and Krüger, et al. (2021). Statistical significance of differences in loss functions is evaluated using the Diebold and Mariano (1995) and West (1996) test.

[Table 3 about here.]

Table 3 compares point and density forecasts generated by BVARs with SV and SVO-t specifications, taking the SV model as the benchmark (see the supplementary online appendix for RMSE and CRPS levels for the baseline model). Point forecasts generated by the SVO-t model over the post-war period (and pre-COVID) are generally on par with those from the SV model, with RMSE ratios in some cases a little below or above 1 but often very close to 1. With density forecast accuracy as gauged by the CRPS, at shorter horizons the SVO-t specification performs very similarly to the SV baseline, with CRPS ratios very close to 1, occasionally a bit lower. At the 12 months horizon, SVO-t yields larger gains over SV, ranging from 2 to 6 percent. Bigger gains in accuracy occur at the horizon of 24 months, with improvements as large as 15 percent. At this horizon,

19

SVO-t improves forecast accuracy for variables including consumption, industrial production, employment, hours, and stock returns. The SVO-t gains are largest for real income, the variable most prone to outliers. Overall, consistent use of SVO-t over the post-war sample improves on the commonly used SV specification, in particular in terms of density forecasts and for those variables more subject to frequent outliers, such as personal income.

Although these 1975-2017 forecast results are favorable to our proposed specifications, they are not necessarily as sharp as Section 5.3's results on model fit. In addition, the forecast results favor the SVO-t model over SVO, whereas the model fits favor SVO. Such a finding is not necessarily uncommon: Even though model fit as assessed through predictive likelihoods is elemental to Bayesian evaluation of models, results on fit can differ from some results on out-of-sample forecast accuracy. One explanation is that, due to the strong curvature of the predictive likelihood's log score loss function, the predictive likelihoods are more responsive to outcomes in the tails; the forecast metrics we use are relatively insensitive to outcomes in the tails. Our SVO estimates appear to assign a little more predictive mass in the tails compared to other models. Another factor, noted above, is the forecast horizon. In view of this paper's focus on the use of BVAR models for forecasting, the advantages of SVO-t over SVO in forecasting leads us to favor it in our analysis and recommendation.

6   Robustness Checks

This section provides a brief overview of a few model robustness checks. The supplementary online appendix provides additional detail and results on these and some other checks.

Common outlier: With the COVID-19 pandemic inducing extreme volatility in a number of variables, some may view it as plausible that the outlier is common to all variables, rather than independent across variables as in the SVO specification. Some other work, such as Lenza and Primiceri (2022), has developed models in which the pandemic induces a common shift in volatility in an otherwise homoskedastic BVAR. Accordingly, we have also considered a specification in which

the outlier state is common to all variables, in which case the time-varying variance-covariance matrix of the VAR residuals is given by $\Sigma_t = \bar{o}_t^2 A^{-1} \Lambda_t (A^{-1})'$, where $\bar{o}_t$ denotes a scalar outlier state. Our estimates indicate that making the outlier common seems to have no advantages. In historical estimates, the common-outlier specification registers virtually no outliers prior to the COVID-19 pandemic. Instead, the common-outlier specification sees outliers only in the early stages of the pandemic period, from March through June 2020, when a good number of variables experienced enormous realizations at the same time, but none in late 2020 or early 2021.

Capturing the pandemic period with dummy variables: As another simple approach to conditioning on knowledge of when and which outliers occurred in the data, particularly the timing of the COVID-19 pandemic, we consider an otherwise standard BVAR-SV model with separate dummy variables (with wide priors assigned to each dummy coefficient) added to represent each month of the sample since COVID's outbreak in March 2020. By soaking up all information contained in data since the onset of the pandemic, the dummy approach generates point forecasts comparable to our outlier-augmented SV models. But because the dummy approach is conditioned on ex-ante knowledge that all COVID-19-related data points are highly unusual, its forecast densities are much tighter than those derived from our more agnostic outlier-augmented SV models or the SV-OutMiss specification.

Variable ordering: In BVARs with stochastic volatility specified as in equations (1) through (3), variable ordering affects estimates. Recent work by Arias, Rubio-Ramirez, and Shin (2021) has shown that ordering choices in VARs with time-varying parameters and SV can affect out-of-sample forecasts. In particular, in their results, ordering has little effect on point forecasts but measurable effects on density-related measures, including the standard deviation of the predictive density and the length of prediction intervals.

The relatively large number of variables in our model means a very large number of possible orderings. Accordingly, we have investigated sensitivity to variable ordering with an approach meant to be broad but streamlined to be computationally tractable (if still demanding). Our basic

metric for sensitivity is the distance between predictive densities obtained in one ordering versus another. We assess the distance and its significance with the potential scale reduction factor (PSRF) of Gelman and Rubin (1992). In particular, we compare predictive densities generated from the SVO model at different forecast origins around and during the onset of the COVID-19 pandemic: December 2019, March and April 2020, and March 2021. For each of these origins, we randomly draw 640 different orderings of the model's 16 variables, estimate each model, and form forecast densities. We then compute a Gelman-Rubin scale reduction test for each variable at each horizon (1 to 24 months ahead). Overall, these results suggest small ordering effects in our forecasts: The vast majority of Gelman-Rubin statistics are under 1.2.

Model stability: The unusual developments of the pandemic inevitably raise a question as to whether it represents a break in conventional business cycle dynamics and time series models. Our results treat the BVARs as stable, taking various steps to limit the influence that extreme observations can have on model estimates. Although it would be ideal to formally test model stability, the sample since March 2020 is too short to permit formal inference with conventional tests or metrics.

As a simple and feasible alternative, we examine the stability of recursive estimates of the BVARs from January 2020 through the end of our sample in 2021. To assess the significance of a change in each coefficient, we take the January 2020 posterior for each coefficient as a reference point. For the sake of comparability, we standardize the change in the posterior means obtained at subsequent forecast origins, by dividing the changes by January's posterior standard deviation. Broadly, these results indicate that — except for the CONST case — there are at most only fairly limited changes in some coefficients, while the vast majority of coefficients show little change. By our simple measures of significance, the CONST specification is quite prone to some coefficient change, most sizably for some economic activity indicators. In the SV specifications, coefficient change appears much less significant. The SVO and SVO-t models show changes in intercepts for some variables, but otherwise, estimates look to be broadly stable over the period.

7 Conclusion

We study an outlier-augmented stochastic volatility specification for Bayesian vector autoregressions. Our work is prompted by the enormous 2020-21 movements in many macroeconomic time series due to the COVID-19 pandemic. As recognized by recent studies such as Lenza and Primiceri (2022) and Schorfheide and Song (2021), these outliers have strong, and sometimes outsized, effects on forecasts made with standard constant-variance BVARs. Our proposed specifications extend to BVARs the earlier work of Stock and Watson (2016) in the context of unobserved component models of inflation, and it is related to SV models with t-distributed errors developed by Jacquier, Polson, and Rossi (2004). The SVO model features stochastic volatility and an outlier-state treatment, and our preferred SVO-t specification augments SVO with fat-tailed shocks. Although we focus on BVARs, our outlier treatments can be readily applied to other time series models estimated with Bayesian methods, including unobserved component models, factor models, factor-augmented VARs, and dynamic stochastic general equilibrium models that feature SV.

Although estimates of BVARs with time-varying volatility tend to down-weight high-volatility observations, the resulting forecasts can be better insulated from outliers. As shown in Section 5.2, BVARs with time-varying volatility generate point forecasts that are less distorted than in the constant-variance case. But a conventional SV model expects all changes in volatility to be persistent, so that it extrapolates huge forecast uncertainty from the initial COVID-19 shocks. In contrast, SVO and SVO-t models fit sharp spikes in current volatility while adapting their uncertainty forecasts more moderately.

An alternative approach could be to pre-screen the data to identify outliers based on a simple measure of historical norms, and then treat these variable-specific outliers as missing observations in an otherwise conventional BVAR with SV. Forecasts from this missing-data approach (SV-OutMiss) neglect the possible arrival of future outliers. In contrast, our outlier-augmented SV models provide a coherent treatment of extremes in the data by modeling the occurrence of outliers as stochastic events, with unknown timing. Accordingly, the resulting forecast densities fully reflect the uncertainty emanating from the presence of outliers in the data.

To evaluate which model best characterizes the data in the COVID-19 period, forecast accuracy could, of course, be a natural metric. However, the sample is too short to support formal inference on the basis of average forecast accuracy. Instead, we employ predictive Bayes factors. By this measure, our SVO specification fits the COVID-19 sample the best, with SVO-t next. The neglected arrival of future outliers in the SV-OutMiss model incurs a sizable penalty in the predictive Bayes factors. Over the entire evaluation sample since 1975, the SVO specification again fares best. The gains of the outlier-augmented SV model are driven by various episodes of relatively high volatility in the data; in contrast, the baseline SV model fits well only in the Great Moderation years of 1985 through 2007. We also conduct an evaluation of out-of-sample forecast performance for a pre-pandemic sample starting in 1975 and ending in 2017. We compare the accuracy of point and density forecasts, as measured by RMSE and CRPS, from standard BVARs against our proposed SVO-t specification. Even in the pre-COVID-19 period, our outlier-augmented SVO-t model forecasts the data, on balance, a little better than a conventional BVAR-SV model.

Notes

[1]For example, suppose one uses monthly data through April 2020 to estimate a medium-sized BVAR and forecast payroll employment growth starting in May 2020. In light of the suggestion of Schorfheide and Song (2021), we also consider forecasts for the same period but using parameter estimates based on data ending in February. The forecasts turn out to be strikingly different.

[2]Following Stock and Watson (2002), applications of dynamic factor models such as Mc-Cracken and Ng (2016) have considered observations to be outliers when they are some multiple of the inter-quartile range away from the series median.

[3]These linear models remain the workhorse of applied forecasting in policy analysis and a benchmark for research. Beyond linear VARs, Guerrón-Quintana and Zhong (2021) and Huber, et al. (2020) employ semi- and non-parametric methods to better allow forecasting relationships to adapt to changing conditions, in particular at times of crisis. Our proposed approach to outliers could also be incorporated into VARs that feature time-varying regression parameters.

[4]The lower bound of 2 on the scale shift in outliers is motivated by seeing outliers as events firmly outside the typical mass of their otherwise Gaussian distribution (conditional on $o_{j,t}$).

[5]The supplementary online appendix includes results for another specification that is nested within the SVO-t model: a SV specification with fat-tail errors but without the infrequent, large outliers of SVO.

[6]In the limit, the missing data approach corresponds to a version of attaching additive measurement error to specific observations, but with infinite variance, whereas the remaining observations are observed without error.

[7]The prior mean of $p_j = 1/(4 \cdot 12)$ implies about the same variance of $o_{j,t}$ in the SVO model as do our prior means of p_j and d_j in the SVO-t model for the combined outlier states $o_{j,t} \cdot q_{j,t}$ (see the supplementary online appendix).

[8]The prior for Σ in the constant-variance model is uninformative; that is, we use an improper Wishart with zero degrees of freedom and scale matrix equal to zero.

[9]The ordering of steps in our MCMC sampler reflects the recommendations of Del Negro

and Primiceri (2015) as also implemented by Cúrdia, Del Negro, and Greenwald (2014) (for SV specifications with fat tails) and Stock and Watson (2016) (for SVO). Specifically, the t-error states, $q_{j,t}$, are sampled before the SV mixture states, while draws from $o_{j,t}$ condition on those mixture states so that $o_{j,t}$ and p_j are sampled after the SV steps of Kim, Shephard, and Chib (1998).

[10]The related paper by Lenza and Primiceri (2022) does not include any interest rates in its BVAR setup. When simulating forecasts for our longer-rate measures, the 5- and 10-year Treasury yields, individual draws have fallen below the ELB as well, and the predictive densities were truncated at the ELB in these cases. Due to the dynamic nature of the forecast simulation, this truncation also has indirect effects on the predictive densities of other variables.

[11]The reported trajectories of volatilities in the VAR residuals, v_t, reflect smoothed estimates of the square roots of the diagonal elements of Σ_t computed from MCMC estimates for different end-points of the data (that correspond to different forecast origins in our out-of-sample forecast evaluation).

[12]For brevity, our discussion will abstract from nuances of the real-time data flow, and simply refer to forecasts being "made" at (or even "in" the month of) a particular forecast origin, even though the underlying data would have been available in FRED-MD only in a subsequent month.

[13]Lenza and Primiceri (2022) consider a slightly smaller BVAR system (with six variables covering mostly employment and price data and observations starting only in 1988) where problems related to COVID-19 already become apparent with data for March 2020; in our 16-variable system case estimated from data starting in 1959, the effects of outliers become most apparent by April.

[14]These additional comparisons also include the SV model with fat tails (SV-t), for which estimates are more varied. In some cases, the SV-t forecast intervals are very similar to the SVO-t estimates, but in others, the SV-t intervals are wider than the SVO-t estimates.

[15]These figures also report realized data and imputed values for lagged outliers obtained from SV-OutMiss. For better scale, we are showing here results for payroll growth and unemployment rate instead of real income growth.

References

[1] Alvarez, Luis J., and Florens Odendahl (2021), "COVID-19 and Bayesian VARs in the Euro Area," *mimeo*, July.

[2] Antolín-Díaz, Juan, Thomas Drechsel, and Ivan Petrella (2021), "Advances in nowcasting economic activity: Secular trends, large shocks and new data," Discussion Paper 15926, CEPR.

[3] Arias, Jonas E., Juan F. Rubio-Ramirez, and Minchul Shin (2021), "Macroeconomic forecasting and variable ordering in multivariate stochastic volatility models," Working Paper 21-21, Federal Reserve Bank of Philadelphia.

[4] Bobeica, Elena, and Benny Hartwig (2022), "The COVID-19 shock and challenges for inflation modelling," *International Journal of Forecasting*, forthcoming.

[5] Carriero, Andrea, Joshua Chan, Todd E. Clark, and Massimiliano Marcellino (2022), "Corrigendum to: large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors," *Journal of Econometrics*, 227, 506–512.

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

[7] Chan, Joshua C.C. (2020), "Large Bayesian VARs: A flexible Kronecker error covariance structure," *Journal of Business & Economic Statistics*, 38, 68–79.

[8] Chib, Siddhartha, Minchul Shin, and Fei Tan (2021), "DSGE-SVt: An econometric toolkit for high-dimensional DSGE models with SV and t errors," *Computational Economics*.

[9] Chiu, Ching-Wai (Jeremy), Haroon Mumtaz, and Gábor Pintér (2017), "Forecasting with VAR models: Fat tails and stochastic volatility," *International Journal of Forecasting*, 33, 1124–1143.

[10] Clark, Todd E. (2011), "Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility," *Journal of Business and Economic Statistics*, 29, 327–341.

[11] Clark, Todd E., and Francesco Ravazzolo (2015), "Macroeconomic forecasting performance under alternative specifications of time-varying volatility," *Journal of Applied Econometrics*, 30, 551–575.

[12] Cúrdia, Vasco, Marco Del Negro, and Daniel L. Greenwald (2014), "Rare shocks, great recessions," *Journal of Applied Econometrics*, 29, 1031–1052.

[13] D'Agostino, Antonello, Luca Gambetti, and Domenico Giannone (2013), "Macroeconomic forecasting and structural change," *Journal of Applied Econometrics*, 28, 82–101.

[14] Del Negro, Marco, and Giorgio E. Primiceri (2015), "Time varying structural vector autoregressions and monetary policy: A corrigendum," *Review of Economic Studies*, 82, 1342–1345.

[15] Diebold, Francis X., and Roberto S. Mariano (1995), "Comparing predictive accuracy," *Journal of Business and Economic Statistics*, 13, 253–63.

[16] Durbin, J., and S.J. Koopman (2002), "A simple and efficient simulation smoother for state space time series analysis," *Biometrika*, 89, 603–615.

[17] Gelman, Andrew, and Donald B. Rubin (1992), "Inference from iterative simulation using multiple sequences," *Statistical Science*, 7, 457–472.

[18] Geweke, John, and Gianni Amisano (2010), "Comparing and evaluating Bayesian predictive distributions of asset returns," *International Journal of Forecasting*, 26, 216 – 230, Special Issue: Bayesian Forecasting in Economics.

[19] Guerrón-Quintana, Pablo, and Molin Zhong (2021), "Macroeconomic forecasting in times of crises," *Journal of Applied Econometrics*, forthcoming.

[20] Huber, Florian, Gary Koop, Luca Onorante, Michael Pfarrhofer, and Josef Schreiner (2020), "Nowcasting in a pandemic using non-parametric mixed frequency VARs," *Journal of Econometrics*, forthcoming.

[21] Jacquier, Eric, Nicholas G. Polson, and Peter E. Rossi (2004), "Bayesian analysis of stochastic volatility models with fat-tails and correlated errors," *Journal of Econometrics*, 122, 185 − 212.

[22] Karlsson, Sune, and Stepan Mazur (2020), "Flexible fat-tailed vector autoregression," Working Paper 2020:5, Örebro University, School of Business.

[23] Kim, Sangjoon, Neil Shephard, and Siddhartha Chib (1998), "Stochastic volatility: Likelihood inference and comparison with ARCH models," *The Review of Economic Studies*, 65, 361–393.

[24] Koop, Gary M. (2013), "Forecasting with medium and large Bayesian VARs," *Journal of Applied Econometrics*, 28, 177–203.

[25] Krüger, Fabian, Sebastian Lerch, Thordis L. Thorarinsdottir, and Tilmann Gneiting (2021), "Predictive inference based on Markov Chain Monte Carlo output," *International Statistical Review*, 89, 274–301.

[26] Lenza, Michele, and Giorgio E. Primiceri (2022), "How to estimate a VAR after March 2020," *Journal of Applied Econometrics*, forthcoming.

[27] McCracken, Michael W., and Serena Ng (2016), "FRED-MD: A monthly database for macroeconomic research," *Journal of Business & Economic Statistics*, 34, 574–589.

[28] Ng, Serena (2021), "Modeling macroeconomic variations after COVID-19," Working Paper 29060, National Bureau of Economic Research.

[29] Primiceri, Giorgio E., and Andrea Tambalotti (2020), "Macroeconomic forecasting in the time of COVID-19," June, *mimeo*.

[30] Schorfheide, Frank, and Dongho Song (2021), "Real-time forecasting with a (standard) mixed-frequency VAR during a pandemic," NBER Working Papers 29535, National Bureau of Economic Research, Inc.

[31] Stock, James H., and Mark W. Watson (2002), "Macroeconomic forecasting using diffusion indexes," *Journal of Business & Economic Statistics*, 20, 147–162.

[32] ——— (2016), "Core inflation and trend inflation," *The Review of Economics and Statistics*, 98, 770–784.

[33] West, Kenneth D. (1996), "Asymptotic inference about predictive ability," *Econometrica*, 64, 1067–1084.

Table 1: List of Variables

Variable	FRED-MD code	Transformation	Minnesota Prior
Real income	RPI	$\Delta \log(x_t) \cdot 1200$	0
Real consumption	DPCERA3M086SBEA	$\Delta \log(x_t) \cdot 1200$	0
Industrial production	INDPRO	$\Delta \log(x_t) \cdot 1200$	0
Capacity utilization	CUMFNS		1
Unemployment rate	UNRATE		1
Nonfarm payrolls	PAYEMS	$\Delta \log(x_t) \cdot 1200$	0
Hours	CES0600000007		0
Hourly earnings	CES0600000008	$\Delta \log(x_t) \cdot 1200$	0
PPI (fin. goods)	WPSFD49207	$\Delta \log(x_t) \cdot 1200$	1
PCE price index	PCEPI	$\Delta \log(x_t) \cdot 1200$	1
Housing starts	HOUST	$\log(x_t)$	1
S&P 500	SP500	$\Delta \log(x_t) \cdot 1200$	0
USD/GBP ex. rate	EXUSUKx	$\Delta \log(x_t) \cdot 1200$	0
5-Year Treasury yield	GS5		1
10-Year Treasury yield	GS10		1
Baa spread	BAAFFM		1

Note: Data obtained from the April 2021 vintage of FRED-MD. Monthly observations from March 1959 to March 2021. Entries in the column "Minnesota Prior" report the prior mean on the first own-lag coefficient of the corresponding variable in each BVAR. Prior means on all other VAR coefficients are set to zero.

Table 2: Log Bayes Factors Relative to SV

| Samples	Models			
	SVO-t	**SVO**	**SV-OutMiss**	**CONST**
Full sample				
1975:01-2021:02	244.11	334.84	-782.79	-9200.01
	(0.44)	(0.60)	(-1.41)	(-16.61)
Great Inflation				
1975-1984	8.25	33.22	17.38	-250.02
	(0.07)	(0.28)	(0.14)	(-2.08)
Great Moderation				
1985-2007	-41.94	-9.69	-6.64	-385.43
	(-0.15)	(-0.04)	(-0.02)	(-1.40)
GFC				
2008-2014	21.53	29.50	-56.28	-236.40
	(0.26)	(0.35)	(-0.67)	(-2.81)
COVID-19				
2020:03-2021:02	225.33	232.59	-739.52	-8167.44
	(18.78)	(19.38)	(-61.63)	(-680.62)
COVID-19 since July 2020				
2020:07-2021:02	3.23	-1.40	-66.19	-617.90
	(0.40)	(-0.18)	(-8.27)	(-77.24)

Note: Differences in cumulative log Bayes factors, $\log L(\mathcal{M}_i) - \log L(\mathcal{M}_0)$, where $\log L(\mathcal{M}_i) = \sum_{t=T_0}^{T_1} \log p(y_{t+1}|y^t, \mathcal{M}_i)$ between the different models listed above ($\mathcal{M}_i$) and the SV model ($\mathcal{M}_0$), measured over different subsamples of forecast origins, t. Unless stated otherwise, samples extend from January to December of the years given. Figures in parentheses provide average score differences over the indicated samples.

Table 3: Historical Forecast Accuracy Comparison, SVO-t vs. SV

| Variable / Horizons	RMSE				CRPS			
	1	3	12	24	1	3	12	24
Real income	1.01	1.00	1.01**	0.94	0.99	0.96***	0.94***	0.85***
Real consumption	1.00	1.01	1.00	1.00	1.00	1.00	0.97***	0.90***
Industrial production	1.00	0.99	0.99	0.97**	1.00	0.99*	0.96***	0.88***
Capacity utilization	1.01	1.00	0.97	0.96	1.00	1.00	0.99	0.95**
Unemployment rate	1.00	0.99	0.98	0.98	1.00	1.00	1.00	0.98
Nonfarm payrolls	0.99	0.99	1.00	0.99	0.99	0.99	0.97***	0.91***
Hours	1.00	0.99	0.96*	0.98	0.99**	0.98**	0.96***	0.90***
Hourly earnings	1.00	1.00	1.01*	1.02**	1.00	0.99**	0.98***	0.92***
PPI (fin. goods)	1.00	1.00	1.00	1.01	1.00	0.99	0.98***	0.94***
PCE price index	1.01**	1.01	1.01*	1.03**	1.01**	1.01	1.00	0.97***
Housing starts	1.00	1.00	1.00	1.02***	1.00	1.01	1.01**	1.00
S&P 500	1.01	1.00	1.01	1.01***	1.00	0.99**	0.97***	0.90***
USD/GBP ex. rate	1.00	1.00	1.00	0.88	0.99*	0.99***	0.96***	0.89***
5-Year Treasury yield	1.01**	1.00	1.00	0.98*	1.01**	1.01	1.00	1.00
10-Year Treasury yield	1.00	1.00	1.00	0.99	1.01**	1.01	1.00	1.01
Baa Spread	0.99**	1.00	1.00	0.99	0.99**	1.00	1.00	0.97***

Note: Comparison of "SVO-t" against "SV" (baseline, in denominator of relative comparisons). Values below 1 indicate improvement over baseline. Evaluation window from January 1975 through December 2017. Significance assessed by Diebold-Mariano-West test using Newey-West standard errors with $h + 1$ lags.

Figure 1: Contributions of Outlier Adjustments to Forecast Error Volatilities

Real income

(a) SVO

(b) SVO-t


S&P 500 returns

(c) SVO

(d) SVO-t
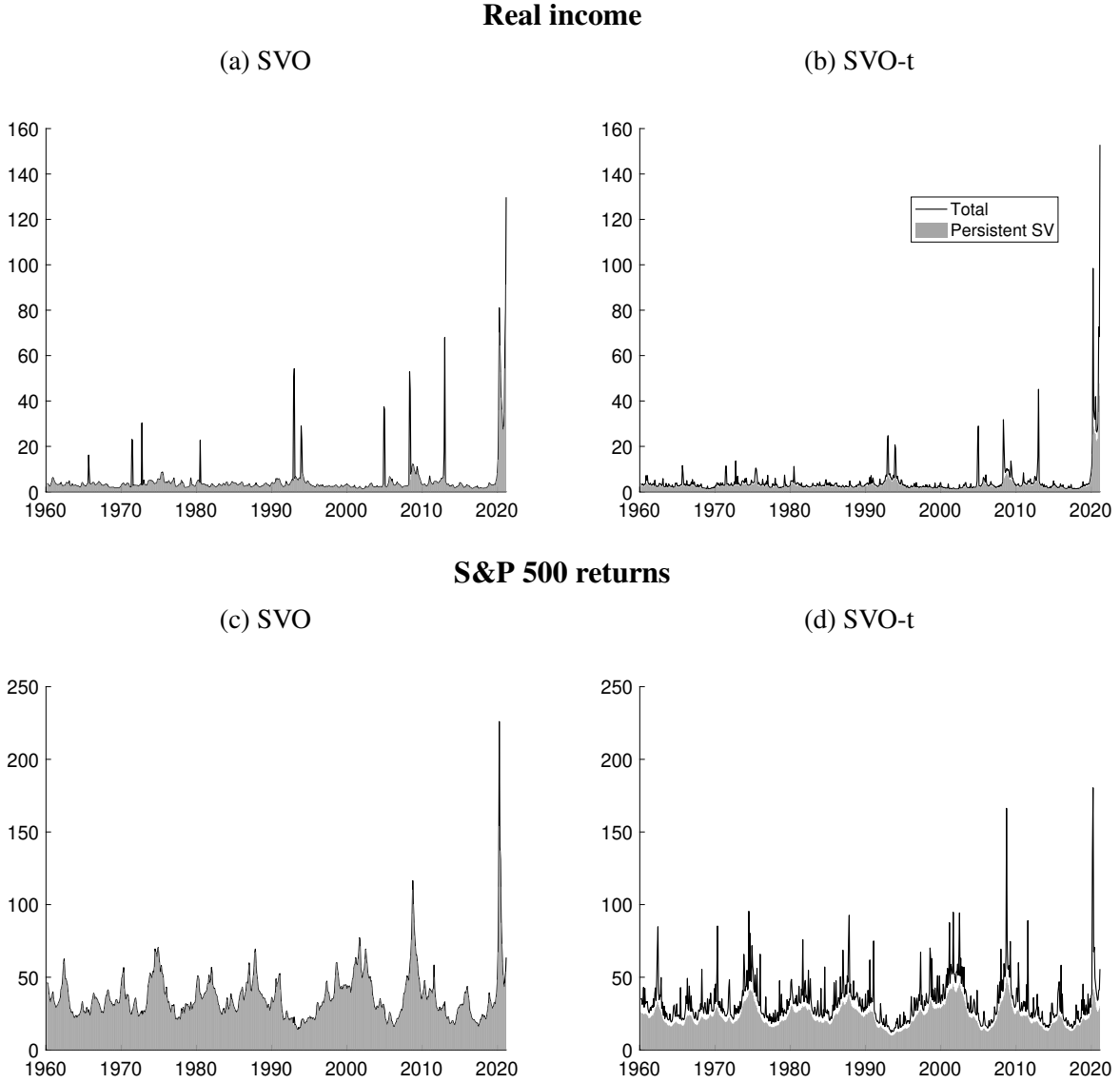
Note: Posterior median estimates per March 2021, of time-varying volatilities in forecast errors of the indicated variables in three outlier-augmented versions of the BVAR-SV model. Dark solid lines depict the actual forecast error volatility, including the effects of O_t and Q_t as applicable in each model. The shaded areas depict the component of each variable's forecast error volatility due to the persistent SV component. Specifically, for the SVO-t model, the forecast error volatility is given by the square root of diagonal elements of $\Sigma_t = A^{-1} O_t Q_t \Lambda_t Q_t' O_t A^{-T}$, whereas the contribution from the persistent SV component follows from $\tilde{\Sigma}_t = A^{-1} \Lambda_t A^{-T}$. For SVO, corresponding computations are performed using only O_t, respectively. These calculations are performed for every MCMC draw, with the resulting medians reported in the figure.

Figure 2: Time-Varying Volatilities of Payroll Growth Since 2020
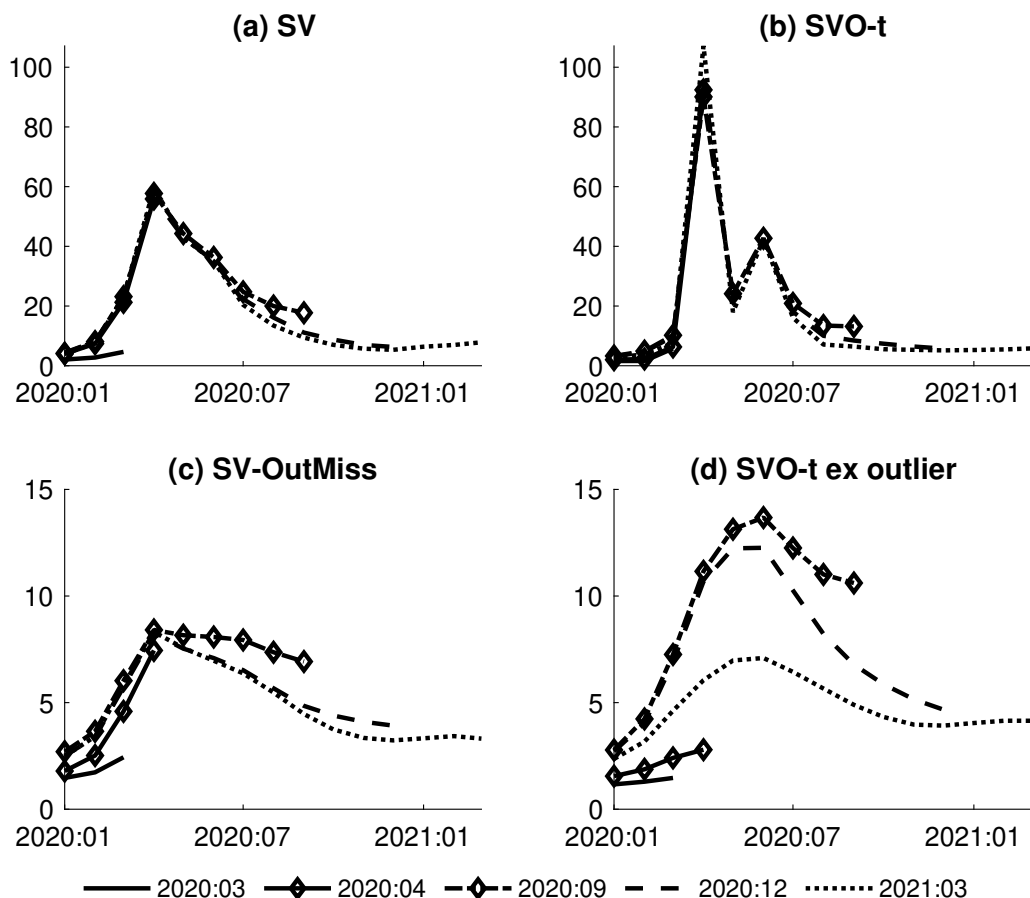
Note: Quasi-real-time trajectories of time-varying volatility in BVAR residuals, measured by (the square roots of) the diagonal elements of $\mathrm{Var}_t\left(v_t\right) = \Sigma_t$ as implied by each model. Medians of (smoothed) posterior obtained from different data samples ending at forecast origins as indicated in the figure legend. Panel (d) displays estimates of stochastic volatility for SVO-t that ignore the contributions from outliers and that are computed from $\tilde{\Sigma}_t = A^{-1}\Lambda_t A^{-T}$ (i.e., neglecting the O_t and Q_t components in the computation of the uncertainty measures shown here, while including these outliers in estimation of A^{-1}, Λ_t, etc.). Reflecting the sizable differences in the size of estimates resulting with and without outlier treatment, different scales are used in upper- and lower-row panels.
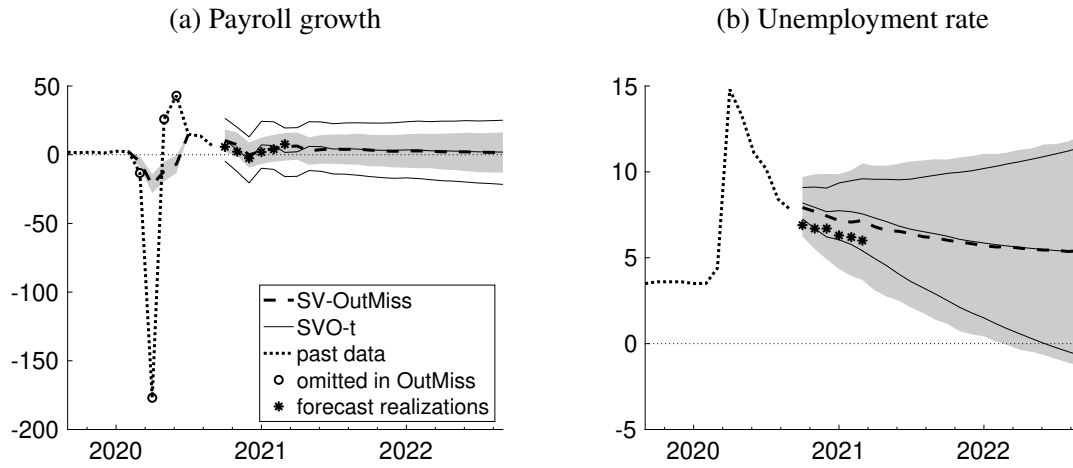
Figure 3: Predictive Densities Since March 2020 from CONST, SV, and SVO-t
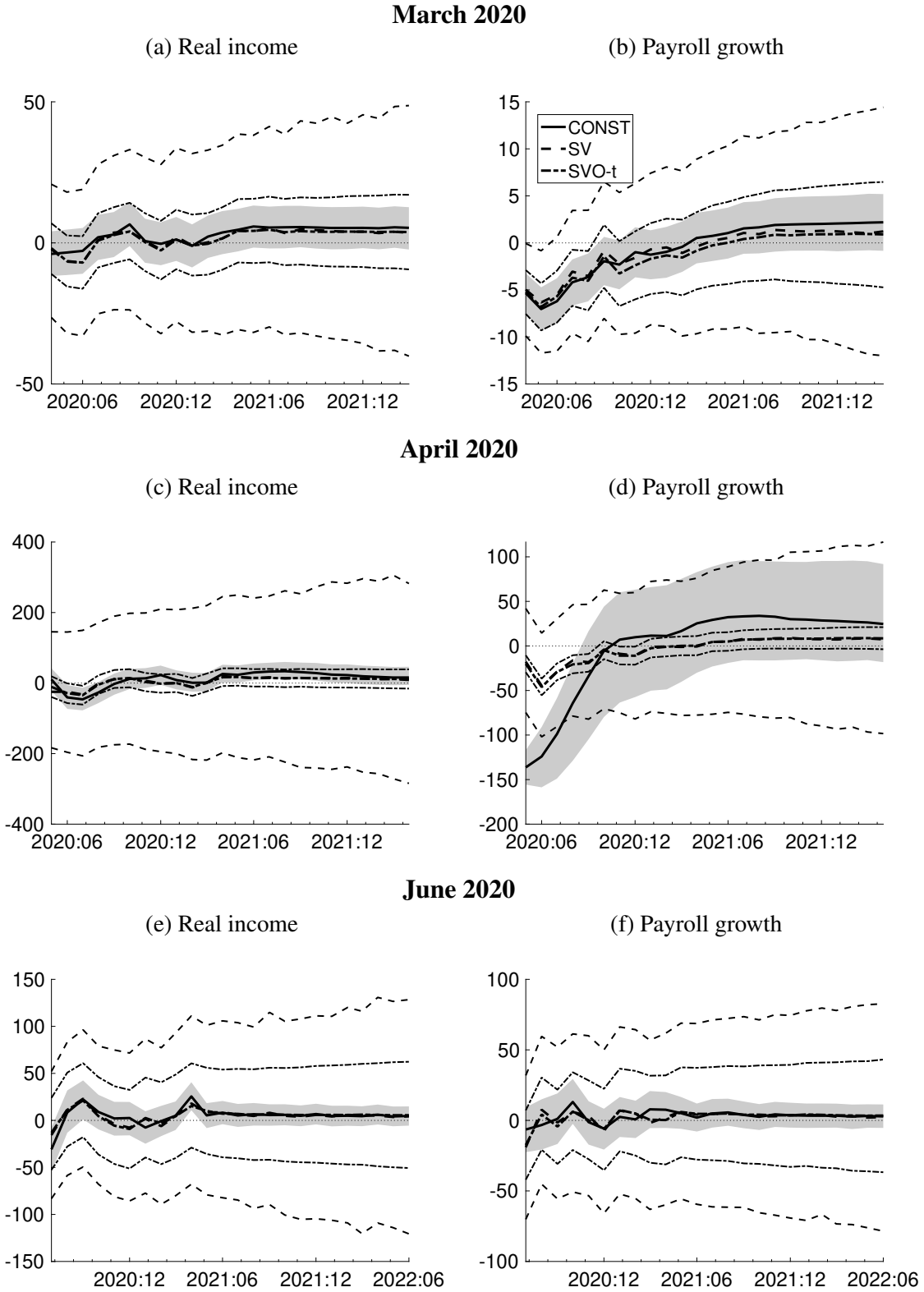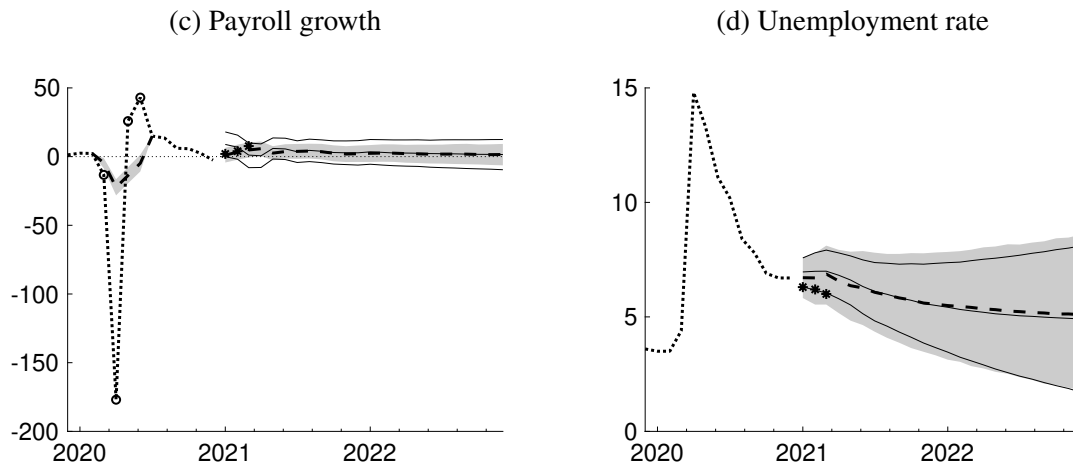
Note: Medians and 68% uncertainty bands of predictive densities (shaded regions give the bands for the CONST forecast), simulated out-of-sample at various forecast origins as indicated in each panel.
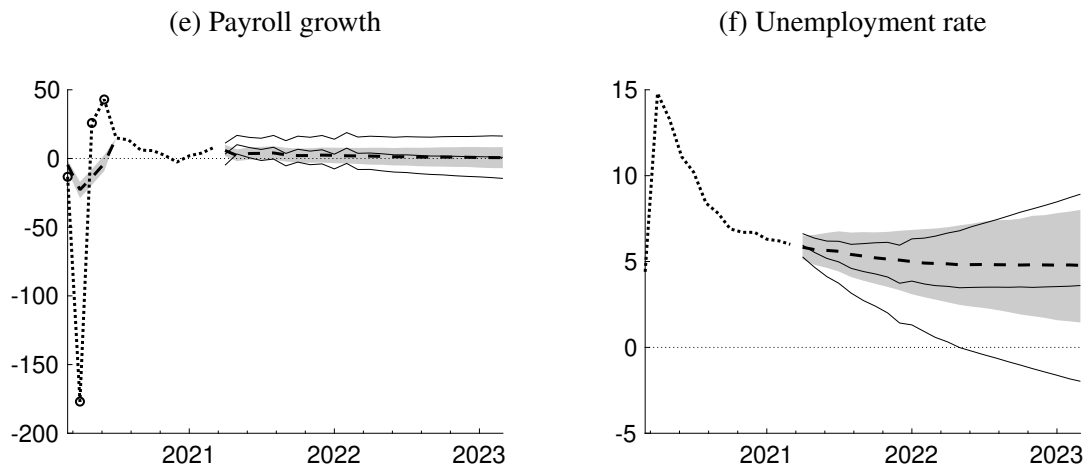
36

Figure 4: Predictive densities since mid-2020

September 2020

(a) Payroll growth

(b) Unemployment rate

December 2020

(c) Payroll growth

(d) Unemployment rate

March 2021

(e) Payroll growth

(f) Unemployment rate

Note: Medians and 68% uncertainty bands of predictive densities (shaded regions give the bands for the SV-OutMiss forecast), simulated out-of-sample at various forecast origins as indicated in each panel.