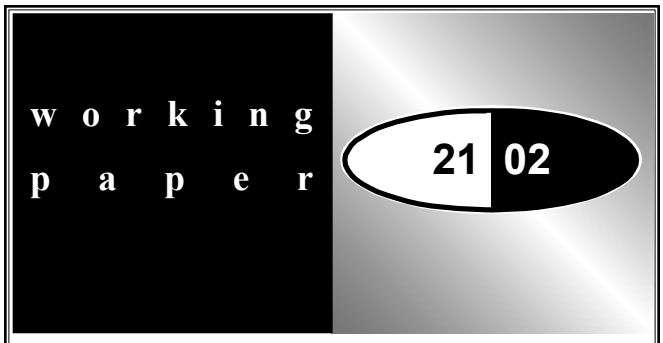


Addressing COVID-19 Outliers in BVARs with Stochastic Volatility

Andrea Carriero, Todd E. Clark,
Massimiliano Marcellino, and Elmar Mertens

FEDERAL RESERVE BANK OF CLEVELAND

ISSN: 2573-7953

Working papers of the Federal Reserve Bank of Cleveland are preliminary materials circulated to stimulate discussion and critical comment on research in progress. They may not have been subject to the formal editorial review accorded official Federal Reserve Bank of Cleveland publications. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Cleveland, the Federal Reserve System, the Eurosystem, or the Deutsche Bundesbank.

Working papers are available on the Cleveland Fed's website at:

www.clevelandfed.org/research.

Working Paper 21-02 February 2021

Addressing COVID-19 Outliers in BVARs with Stochastic Volatility

Andrea Carriero, Todd E. Clark, Massimiliano Marcellino, and Elmar Mertens

Incoming data in 2020 posed sizable challenges for the use of VARs in economic analysis: Enormous movements in a number of series have had strong effects on parameters and forecasts constructed with standard VAR methods. We propose the use of VAR models with time-varying volatility that include a treatment of the COVID extremes as outlier observations. Typical VARs with time-varying volatility assume changes in uncertainty to be highly persistent. Instead, we adopt an outlier-adjusted stochastic volatility (SV) model for VAR residuals that combines transitory and persistent changes in volatility. In addition, we consider the treatment of outliers as missing data. Evaluating forecast performance over the last few decades in quasi-real time, we find that the outlier-augmented SV scheme does at least as well as a conventional SV model, while both outperform standard homoskedastic VARs. Point forecasts made in 2020 from heteroskedastic VARs are much less sensitive to outliers in the data, and the outlier-adjusted SV model generates more reasonable gauges of forecast uncertainty than a standard SV model. At least pre-COVID, a close alternative to the outlier-adjusted model is an SV model with *t*-distributed shocks. Treating outliers as missing data also generates better-behaved forecasts than the conventional SV model. However, since uncertainty about the incidence of outliers is ignored in that approach, it leads to strikingly tight predictive densities.

Keywords: Bayesian VARs, stochastic volatility, outliers, pandemics, forecasts.
JEL classification codes: C53, E17, E37, F47.

Suggested citation: Carriero, Andrea, Todd E. Clark, Massimiliano Marcellino, and Elmar Mertens. 2021. "Addressing COVID-19 Outliers in BVARs with Stochastic Volatility." Federal Reserve Bank of Cleveland, Working Paper No. 21-02. https://doi.org/10.26509/frbc-wp-202102.

Andrea Carriero is at Queen Mary University of London (a.carriero@qmul.ac.uk). Todd E. Clark (corresponding author) is at the Federal Reserve Bank of Cleveland (todd.clark@clev.frb.org). Massimiliano Marcellino is at Bocconi University, IGIER, and CEPR (massimiliano.marcellino@unibocconi.it). Elmar Mertens is at the Deutsche Bundesbank (elmar.mertens@bundesbank.de). The authors gratefully acknowledge helpful suggestions and comments received from seminar participants at the ECB, Bocconi University, Bank of England, Deutsche Bundesbank, and the seventh annual meeting on New Developments in Business Cycle Analysis, organized by Danmarks Nationalbank, Deutsche Bundesbank, and Norges Bank. Additional results are provided in a supplementary online appendix.

1 Introduction

Bayesian VARs have a successful track record in point and density forecasting, the measurement of tail risks, and structural analysis. However, incoming data in 2020 posed some basic challenges for estimation and inference with VARs. The economic turbulence created by the ongoing COVID-19 pandemic is reflected in extreme realizations for a number of macroeconomic and financial series for the US, as shown in Figure 1.[1] The period yielded unprecedented changes in many key variables. For example, payroll employment plummeted by about 15 percent from March to April, 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.[2] Measured by the business conditions index of Aruoba, Diebold, and Scotti (2009), the drop in real activity recorded in 2020 is more than 5 times as deep as in any other recession since 1960, so that the previous Great Recession of 2007-09 "appears minor by comparison" as noted by Diebold (2020). These extreme realizations have strong effects on parameter estimates and forecasts generated by conventional constant-parameter VARs. In response, Schorfheide and Song (2020) suggest ignoring the recent data in estimating VAR parameters, whereas Lenza and Primiceri (2020) propose a specific form of heteroskedasticity, tuned to the COVID data, to down-weight observations since March 2020 in the estimation.

Prior to the COVID-19 era, heteroskedastic VAR models, in particular models with stochastic volatility (SV), have been shown to provide 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 VAR's fore-

[1] Throughout this paper, we consider US data, but the pandemic led to similar turbulence in other economies around the world.

[2] These calculations use log growth rates and data from the October 2020 vintage of FRED-MD. The rise in measured income from March to April also reflects payouts of government stimulus in that month. In contrast, over the following month, real income fell by about 4.5 percent, the second-highest drop in our data (the largest drop in real income, by about 5 percent, occurred in January 2013).

cast errors over time, with potential benefits for the accuracy of density forecasts (Clark, McCracken, and Mertens (2020)). In addition, heteroskedasticity affects the estimation of slope coefficients in each VAR equation (at least in finite samples). As an application of generalized least squares, observations recorded at times of high volatility are down-weighted in the estimation of VAR parameters.[3] When extreme realizations are modeled as sudden increases in volatility, heteroskedastic VARs will down-weight the associated observations when estimating parameters; in the limit, outliers associated with infinite volatility would be discarded.

A typical SV model assumes changes in volatility to be highly persistent.[4] However, almost by definition, extreme observations are more reflective of short-lived spikes in volatility, not permanent increases in forecast uncertainty. Like Schorfheide and Song (2020) and Lenza and Primiceri (2020), we view the extreme observations of the COVID period as possible outliers that are characterized by transient increases in volatility, in which case it may be desirable to reduce their influence on model estimates and forecast distribution.

In this spirit, a few prior contributions in the SV literature have taken steps to address outliers in historical data as transitory outliers. One example is the SV model with fat-tailed, instead of normal, errors by Jacquier, Polson, and Rossi (2004), henceforth denoted "SV-t."[5] Another example is the outlier-augmented SV process used by Stock and Watson (2016) with unobserved component models of inflation, henceforth denoted "SVO."

[3]For example, when applied to data samples starting in the 1960s or 1970s, VARs with SV tend to discount data points prior to the onset of the low-volatility period known as the Great Moderation that started in the mid-1980s (Perez-Quiros and McConnell (2000)). Of course, the distinction between generalized and ordinary least squares matters only in finite samples, as both converge to the same asymptotic limit (to which a Bayesian estimate would also converge). But as demonstrated by the COVID-19 episode, common samples of macroeconomic data are still sufficiently finite for (huge) outliers to matter.

[4]In typical implementations, such as those following Cogley and Sargent (2005), Stock and Watson (2007), Justiniano and Primiceri (2008), and Clark (2011), log-variances are assumed to follow random walks, or highly persistent AR(1) processes, and Clark and Ravazzolo (2015) find relatively similar forecast performance resulting from either approach in post-war US data.

[5]Following Jacquier, Polson, and Rossi (2004), t-distributed shocks have been 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 (2020). Most recently, Karlsson and Mazur (2020) provide a general treatment of heteroskedasticity in BVAR models with and without SV and fat-tailed error distributions.

In this paper, we extend the SVO approach of Stock and Watson (2016) to BVAR-SV models and show that it effectively filters the outliers associated with the unprecedented, temporary volatility induced by the COVID-19 pandemic. In addition, SVO also detects pre-COVID outliers in macroeconomic and financial time series, whose existence had been noted by, among others, Stock and Watson (2002). Conventional BVAR-SV procedures can easily be extended to include outlier state estimation via the SVO approach. We consider the effects of adding the SVO specification to a BVAR during both the recent COVID-19 period and the post-war sample of US data on macroeconomic and financial variables. Although at this point we are comfortable viewing the extreme realizations of the COVID-19 period as outliers, we should emphasize that our approach is data-based: Our model estimates outliers conditional on the data; we are not simply deeming (i.e., restricting) recent observations to be outliers.[6]

The SVO model augments the standard SV specification of a highly persistent volatility state with an outlier volatility state that infrequently and temporarily jumps to values above 1. As we demonstrate further below, both SV-t and SVO share the same latent state representation where residuals are written as the product of a normally distributed shock and an outlier state, but differ in the assumed densities for the outlier state. SVO puts more mass on outliers being large events that increase volatility by more than twofold, whereas SV-t sees outliers as more moderately sized. While the SV-t model has been studied more extensively already, the SVO approach is relatively novel, and it comes with special promise for modeling large jumps. While we are particularly interested in the performance of SVO during the COVID-19 episode, we also study its versatility outside the pandemic.

The COVID-19 pandemic visibly affected the US economy starting in March 2020. We

[6]Throughout, we stay in the class of (conditionally) linear VAR models with time-invariant transition (i.e., coefficient) matrices that remain the workhorse of applied forecasting in policy analysis and a benchmark for use in research. Beyond linear VARs, Guerrón-Quintana and Zhong (2017) and Huber et al. (forthcoming) employ semi- and non-parametric methods to better allow forecasting relationships to adapt to changing conditions, in particular at times of crisis. Antolín-Díaz, Drechsel, and Petrella (2020) consider a dynamic factor model with time-varying volatility and shifting means, where outliers are modeled as t-distributed measurement errors.

confirm the findings of Lenza and Primiceri (2020) and Schorfheide and Song (2020) that forecasts generated since then from homoskedastic BVARs are often implausible; for example, the recent outliers cause the forecast paths of some variables to become explosive. Instead, we find that BVARs with SV or SVO specifications generated better-behaved point forecasts. Both SV and SVO estimates register increases in forecast uncertainty. But, while the SV specification sees all shocks to forecast uncertainty as permanent, the SVO model explicitly allows 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. So, in our assessment, the SVO specification offers an effective approach for managing infrequent outliers with BVARs used for forecasting. That being said, other approaches may offer similar benefits in forecasting with BVARs, with one option being the SV-t model. Although we prefer the SVO specification for being nimble in adapting to extreme and rare jumps in volatility like those seen since the onset of the recent pandemic, the SV-t model yields some of the same benefits to forecast accuracy.

In addition, we consider two alternative approaches for treating outliers when their occurrence can be identified prior to the VAR's estimation. First, specifically for handling COVID-19 outliers, we estimate a BVAR-SV model with separate dummies attached to the VAR's mean equation for every month since March 2020. By construction, these dummies soak up the VAR residuals since March so that the approach is tantamount to ignoring data since March for the estimation of forecast parameters.[7] Empirically, we find that the point forecasts resulting from the dummy-augmented VAR are similar to those obtained from standard SV or the outlier-augmented SVO specification. But estimates of forecast uncertainty remain unrealistically stuck at pre-COVID levels.

[7]In a related effort, Holston, Laubach, and Williams (2020) augment a trend-cycle decomposition for output in the US and other economies with an exogenous COVID indicator based on the COVID-19 Government Response Stringency Index computer at the Oxford Blavatnik School of Government for each country or region. In most cases, the stringency index is a slow-moving variable, and the procedure corresponds to correcting mean effects from COVID with a (time-varying) dummy. Similarly, updates for the uncertainty measures from Jurado, Ludvigson, and Ng (2015) are computed by these authors based on mean-adjusted data for the COVID period.

Second, to guard against outliers affecting the jump-off data, we also consider a standard BVAR-SV that treats extreme observations as missing data. Most of 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 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.[8] This approach differs from the SVO approach, which estimates the occurrence of outliers jointly with the VAR, by treating the dates of outliers as known ex-ante. In the COVID period, this approach also produces much better-behaved forecasts than a constant-variance BVAR. Empirically, the biggest difference with the outlier-adjusted SV procedures is that conditioning on the incidence of outliers, while otherwise ignoring any signal from their specific realization, leads to predictive densities that are considerably tighter than those from SVO (or SV-t), though not quite as implausibly so as the aforementioned dummy approach.

Although to this point we have focused on the efficacy of methods for reducing distortions to forecast distributions in the presence of outliers, to be broadly effective, it is important that a given method not only helps reduce such distortions but also performs effectively in forecasting over long periods of time less affected by outliers. Accordingly, we conduct a quasi-real-time evaluation of forecast performance using monthly data with an evaluation window starting in 1985 and ending in 2017, comparing the accuracy of point and density forecasts from our proposed SVO specification and the alternatives discussed above. In all cases, we use a medium-sized data set of 16 monthly variables, motivated by research that has found that larger BVARs tend to forecast more accurately than smaller BVARs, while going beyond medium-sized models adds little gains (e.g., Banbura, Giannone, and Reichlin (2010), Carriero, Clark, and Marcellino (2019), and Koop (2013)). Considering

[8]Following Stock and Watson (2002), applications of dynamic factor models have considered observations to be outliers when they are some multiple of the inter-quartile range away from the series median; among others, see Artis, Banerjee, and Marcellino (2005) and McCracken and Ng (2016).

forecast performance over a long sample starting in 1985 (excluding COVID-19 data), the SVO approach marginally outperforms SV, and both do better than a homoskedastic BVAR in terms of point and density forecasts. The SV-t specification offers similar advantages.

All told, the use of VARs with time-varying volatility, like SV and SVO, broadly mitigates the drastic effects that outliers can have on forecasts. But only an outlier-adjusted SV specification, like SVO or SV-t, prevents the width of predictive densities from blowing up as they would in the SV case. Importantly, the added value of SVO also holds up over a longer sample outside the recent COVID-19 episode, similar to an SV model with t-distributed errors.[9]

The remainder of this paper proceeds as follows: Section 2 introduces the SVO model and alternative specifications to handle outliers, and describes their estimation. Section 3 describes the data used. Section 4 provides a forecast comparison between the various models over a long pre-COVID sample. Section 5 reports details about estimated outlier states before and during the COVID-19 episode, and Section 6 describes the evolution of forecasts made over the course of 2020. Section 7 concludes. Additional results are provided in a supplementary online appendix.

2 BVAR models

We study VAR models of the following form:

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

where y_t is a vector of N observables, $\Pi(L) = \sum_{i=1}^{p} \Pi_i \; L^{i-1}$ is a p-th 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. Building on the methods of Carriero, Clark, and

[9] In a companion paper (Carriero et al. (2020)), we document the effects of SVO for measuring uncertainty and its effects on the economy during the COVID-19 era, and we find the estimates to be more reasonable compared to standard SV.

Marcellino (2019) (henceforth "CCM") for estimating large BVARs, all models are specified with non-conjugate priors for Π and Σ_t.

The models differ mainly in whether the residuals are homoskedastic, or the form of their heteroskedasticity. We maintain the assumption of time-invariant transition coefficients Π. Such constant-parameter VARs are commonly and successfully used in forecasting.[10] Heteroskedasticity in the VAR residuals has important effects on the estimation of Π, in particular when there are outliers with large residual volatility. Intuitively, observations with higher residual volatility receive less weight in the estimation of VAR coefficients. For the sake of illustration, consider an AR(1) model without intercept: $y_t = \pi y_{t-1} + v_t$, $v_t \sim N(0, \sigma_t^2)$ with σ_t^2 known, and a prior conditional on past data $\pi | y_{t-1} \sim N(\underline{\pi}, \underline{\omega}^2)$. This is a signal extraction problem where y_t serves as a noisy signal about the unknown π, with a signal-to-noise ratio that is decreasing in σ_t^2. Accordingly, the posterior mean for π is a weighted average of the prior mean, $\underline{\pi}$, and the data-driven OLS estimate, π^{OLS}, with the weight decreasing in σ_t^2. In the case of observing a single observation y_t, these are:

$$E\left(\pi | y_t, y_{t-1}\right) = (1-\kappa) \cdot \underline{\pi} + \kappa \cdot \pi^{OLS}, \quad \text{with} \quad \pi^{OLS} = \frac{y_t \, y_{t-1}}{y_{t-1}^2}, \quad \text{and} \quad \kappa = \frac{\underline{\omega}^2}{\underline{\omega}^2 + \frac{\sigma_t^2}{y_{t-1}^2}}.$$

Recursive application of the above extends the example to multiple periods. In addition, the logic of down-weighting observations subject to high residual variance carries over to the multivariate case, as described, for example, in Koop (2003, Chapter 6).

As argued above, time-varying volatility in the VAR residuals, v_t, can help to insulate estimation of the transition coefficients Π from the effects of extreme outliers. However, density forecasts will crucially depend on the assumed dynamics of the variances in Σ_t, and we further consider different forms of persistence in variance changes below.

Down-weighting extreme observations in the estimation of Π will not completely insulate

[10] Although we leave an extension to future research, our proposed approach to outliers could easily be incorporated into VARs also featuring time-varying regression parameters in the smaller specification and estimation approach of D'Agostino, Gambetti, and Giannone (2013) and the larger specification and estimation approach of Chan (2019).

the resulting forecasts from outliers. Consider again the case of the AR(1) 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 $y_{t+h|t}$.[11] To address these concerns, we consider a variant of the SV model that treats pre-specified outliers as missing values. To identify extreme observations as outliers, we use an ex-ante criterion taken from the literature on dynamic factor models that is based on the distance between a given data point and the time-series median.[12]

2.1 Model specification

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

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

2) SV: This is the baseline SV model of CCM, where the VAR residuals can be written as

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

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

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

[11] In VAR (or AR) models with higher lag orders, the forecast would not singularly depend on the outlier y_t but also preceding values that are not necessarily outliers. Nevertheless, outliers in the "jump-off" data point, y_t, may unduly influence the forecast.

[12] In addition, Section 6 reports results for a model variant where (1) is augmented by additional dummy terms for months during the COVID-19 period.

3) SVO and 4) SV-t: The outlier-adjusted version of the SV model (SVO) and a version of the SV model with t-distributed errors share a latent-state representation where the vector of VAR residuals is written as $v_t = A^{-1} O_t \Lambda_t^{0.5} \varepsilon_t$ with A^{-1} and $\Lambda_t^{0.5}$ specified as before. O_t is a diagonal matrix of latent outlier states that are mutually and serially *iid* with typical element $o_{j,t}$. SVO and SV-t differ in their specification of the density for the outlier states, $o_{j,t}$, as described further below.

The SVO specification distinguishes between regular observations with $o_{j,t} = 1$ and outliers with $o_{j,t} > 1$. Outliers in variable j occur with probability p_j. Building on Stock and Watson's (2016) treatment of outliers in unobserved component models, the outlier states have the following density:

$$o_{j,t} = \begin{cases} 1 & \text{with probability} \quad 1 - p_j \\ U(2,20) & \text{with probability} \quad p_j \end{cases} \tag{4}$$

for $j = 1, \ldots, N$ and where $U(2, 20)$ denotes a uniform distribution with support between 2 and 20.[13] The time-varying variance-covariance matrix of the VAR residuals is then given by $\Sigma_t = A^{-1} O_t \Lambda_t O_t^T (A^{-1})'$. As in Stock and Watson (2016), we place a beta prior on the outlier probability p that corresponds to 10 years' worth of prior data with an outlier occurring once every four years.

The SV-t specification corresponds to the fat-tailed SV model of Jacquier, Polson, and Rossi (2004) where the standard-normal shocks ε_t driving the VAR residuals in (2) are replaced by t-distributed shocks. For our estimation, the degrees of freedom of the t distribution are fixed at a value of five.[14] As noted by Jacquier, Polson, and Rossi, the model with t-distributed errors also has a representation for the VAR residuals of the form $v_t = A^{-1} O_t \Lambda_t^{0.5} \varepsilon_t$ with $\varepsilon_t \sim N(0, I)$ as above, and O_t being a diagonal matrix of *iid* latent

[13]The SVO specification closely follows Stock and Watson (2016), except for their use of a $U(2, 10)$ in the context of modeling inflation. As in Stock and Watson (2016), we implement the uniform distribution for outliers on an evenly spaced grid of all integers in its support.

[14]We obtain similar results for alternative choices of degrees of freedom that are below 10, so as to preserve sufficiently fat tails compared to the normal distribution as used in the SV model.

states.

The difference with the SVO case is that, in the SV-t specification, the diagonal elements of O_t, $o_{j,t}$, have inverse-gamma distributions.[15] Figure 2 illustrates the differences in densities implied for $o_{j,t}$.[16] In both cases, the density for the outlier state $o_{j,t}$ peaks at (SVO) or near (SV-t) the value of 1 with a fat right-hand tail. In the SVO case, there is equal probability on outlier states between 2 and 20, whereas the SV-t case assigns most probability on values close to 1, albeit with some minimal measure placed also on values far above 20. Also, while the outlier states in the SVO case cannot take values below 1, the SV-t case assigns some mass also to values below 1. While we regard the SVO and SV-t approaches as comparable in some respects, SVO is more geared toward generating sizable outliers at a variable-specific rate of occurrence p_j that is directly governed by an explicit prior, which we view as conceptually preferable, particularly in the COVID context.

5) SV-OutMiss: This model applies the standard SV specification for Σ_t, but ignores a given set of outlier observations in the VAR estimation altogether by treating them as missing data. The approach builds on a practice known from the literature on dynamic factor models (DFM), in which input data are pruned of extreme observations that are multiples times the inter-quartile range away from the series median. Typical values for the multiple used in the literature vary from 5 to 10, and we adopt a threshold factor of 5 as a baseline, with very similar results based on a value of 10. Figure 3 provides an overview of which observations in our data qualify as outliers according to this criterion. Apart from readings for employment, consumption, income, and stock returns in 2020, and the fairly

[15] Specifically, let $d_j/o_{j,t} \sim \chi^2_{d_j}$ where d_j are the degrees of freedom of the resulting t distribution for $o_{j,t} \cdot \varepsilon_{j,t} \sim t_{d_j}$.

[16] To create the figure, the outlier probability p_j of the SVO model has been set equal to its prior mean, corresponding to an outlier occurring once every four years in monthly data. The degrees of freedom of the corresponding SV-t distribution have been set to match the variance of $o_{j,t} \cdot \varepsilon_{j,t}$ implied by the aforementioned choice for p_j in the SVO model. That means that with $d_j = 5.70$ and $p_j = 1/(4 \cdot 12)$, we get

$$\mathrm{Var}\left(o_{j,t} \cdot \varepsilon_{j,t}\right) = \mathrm{Var}\left(o_{j,t}\right) = (1 - p_j) + p_j \cdot \frac{(20-2)^2}{12} = \frac{d_j}{d_j - 2} = 1.54.$$

frequent occurrence of outliers in income throughout the sample seen also in Panel a of Figure 1, further outliers are recorded in industrial production, inflation, and stock returns during the recession of 2007-09, as well as exchange rates during the 1970s.

The DFM literature replaces extreme observations by the time-series median or a similar moment of central tendency. We adopt the same ex-ante criterion for the identification of outliers, but we instead treat these as missing data in estimation and forecasting. For each missing value, our Bayesian methods generate a posterior distribution that also informs the resulting forecasts. Formally, denote the history of y_t after pruning from outliers as z^t, and continue the AR(1) example introduced above: Forecasts are then generated by $y_{t+1|t} = \pi^h E(y_t|z^t)$ 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.

2.2 Model estimation

Each of our models is estimated with an MCMC sampler, based on the methods of CCM for large BVAR-SV models, with details provided therein. 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. Throughout, we use $p = 12$ lags in a monthly data set, which is described in further detail in Section 3.

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, 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.[17] For the SVO variant, two extra steps need to be added to the original BVAR-SV setup: First, realized outlier states need to be drawn from their

[17]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.

posterior, conditional on draws for each variable's outlier probability; this step 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 second additional step draws the outlier probability for each variable from a (conditional posterior) beta distribution conditional on the draws of the time series of outlier states.

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 VAR 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.[18] In contrast, the added cost of computing SVO or SV-t over standard SV is small, since both models add only steps for sampling the *iid* 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.

3 Data

Our data set consists of monthly observations for 16 macroeconomic and financial variables for the sample 1959:M3 to 2020:M9, taken from the October 2020 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.[19]

[18] In our application, and across different computational settings, the added cost of estimating SV-OutMiss was a multiple of the computational time used for the original SV model.

[19] The related paper by Lenza and Primiceri (2020) does not include any interest rates in its VAR setup. When simulating forecasts for our longer-rate measures, the 5- and 10-year Treasury yields, individual draws

A few selected series are shown in Figure 1, with potential outliers marked in red. In the figure, observations are marked as outliers if their distance from the series median exceeds 5 times the inter-quartile range (IQR), where median and IQR are computed from the pre-COVID-19 sample. As discussed in the introduction, similar definitions of outliers have been used in the literature on factor models in macroeconomics. Real personal income, shown in Panel a of the figure, has regularly displayed outliers over the post-war sample. Many other series, like payroll growth shown in Panel b, exhibit such outliers only over the recent COVID-19 period, whereas a few others, like returns on the S&P500, in Panel c, 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 shown in Panel d, have registered outstanding changes this year, but without registering explicit outliers according to this metric. In some cases, outliers may be attributed to unusual events. For example, in results not shown, industrial production registers a positive outlier in December 1959, when production bounced back following a strike in the steel industry from mid-July through early November. More recently, income transfers from the CARES Act caused growth in personal income to surge in April 2020.

4 Forecast performance pre-COVID-19

Applicability of the outlier-augmented BVAR-SVO model is not necessarily specific to data resulting from the current COVID-19 pandemic. As noted above, individual data series have exhibited occasional outliers before, leading to some earlier studies of the potential benefits of modeling fat-tailed error distributions and other forms of outliers.[20] In this section,

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. In companion work (Carriero et al. (2021)), we focus on the estimation of VARs that model nominal interest rates as censored variables based on the shadow-rate approach described by Johannsen and Mertens (forthcoming).

[20]See, for example, Chiu, Mumtaz, and Pintér (2017), Clark and Ravazzolo (2015), and Cúrdia, Del Negro, and Greenwald (2014) for the use of SV-t specifications in VARs or DSGE models and Stock and Watson (2016) for the use of SVO in unobserved component models.

we evaluate the forecast performance of the alternative BVAR specifications described in Section 2 when applied to a sample of post-war US data prior to the onset of COVID-19.

We conduct an out-of-sample forecast evaluation in quasi-real time, where we simulate forecasts made from 1985:M1 through 2017:M12.[21] For every forecast origin, each model is re-estimated based on growing samples of data that start in 1959:M3. All data are taken from the October 2020 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. (2020). Statistical significance of differences in loss functions is evaluated using the Diebold and Mariano (1995) and West (1996) test.

Tables 2 and 3 compare point and density forecasts generated by BVARs with SV and SVO specifications relative to those resulting from a homoskedastic BVAR. Confirming results known from the earlier literature on the use of BVAR-SV models (e.g., Clark (2011) and Clark and Ravazzolo (2015)), SV outperforms the CONST benchmark for many variables and forecast horizons. For example, with point forecasts, the SV model improves on the RMSEs of the CONST specification by 2 to 5 percent in the case of employment growth and as much as 21 percent in the case of the Baa spread. With density forecast accuracy as gauged by the CRPS, at shorter horizons the SV specification yields significant gains for many variables, including consumption, employment, hours, and interest rates.

The SVO specification could be expected to capture better the occasional outliers in pre-COVID-19 data, but possibly also at the expense of overfitting elsewhere. However, such concerns are not borne out by our forecast evaluation. In terms of both point and density forecasts, SVO typically performs as well as, and at times even better than, SV. SVO yields gains over the CONST specification comparable to those discussed for the SV model. Like

[21] The end of our evaluation window has been chosen to avoid overlap with COVID-19-related realizations; however, we obtain very similar results when the evaluation window is extended through the end of our data sample in 2020.

the SV model, the SVO model improves on the RMSEs of the CONST specification by 3 to 5 percent in the case of employment growth and as much as 22 percent in the case of the Baa spread. With density forecast accuracy as gauged by the CRPS, at shorter horizons the SVO specification yields gains for many variables, including consumption, employment, hours, and interest rates. For a variable subject to a number of historical outliers, such as real personal income, the SVO specification yields consistently better density forecasts than SV does, and point forecasts that perform at par with or better than SV.

Overall, the evidence suggests that consistent use of SVO over the post-war sample shares similar benefits over CONST with SV, and marginally improves forecasts even further, in particular in terms of density forecasts and for those variables more subject to frequent outliers, such as personal income.

Tables 4 and 5 compare SVO against the related SV-t model as well as the SV-OutMiss approach, which treats pre-specified outliers as missing data as described in Section 2. By and large, point and density forecasts from these alternatives are quite similar in accuracy to those from the SVO specification. Differences in relative RMSE or CRPS rarely exceed 5 percent of the performance statistics achieved under SVO, with more pronounced differences in density rather than point forecasts. For a few variables, including interest rates and employment, the RMSE of point forecasts obtained from SV-t are a little worse (by 2 to 3 percent) than under SVO. But SV-t does a little better in terms of density forecasts as measured by CRPS, in particular at longer horizons with improvements of up to 6 to 10 percent in the cases of real income and the Baa spread, respectively. While the RMSE of point forecasts generated by SV-OutMiss are largely identical to those from SVO, the CRPS of the corresponding density forecasts tend to be a little better, in particular at longer horizons, though a little less so than under SV-t.

5 Outlier estimates in 2020 and before

The key novelty of the SVO approach is the additional latent outlier states, $o_{j,t}$ as specified in (4) for each variable $j = 1, \ldots, N$, and the associated probability, p_j, of observing an outlier of scale larger than unity. 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 $o_{j,t}$, as well as the persistent variations induced through the log-SV terms $\log \lambda_t$. As described in Section 2, the SV-t model with fat-tailed t errors can also be represented as the product of *iid* outlier states $o_{j,t}$ and normally distributed errors. In both cases, $o_{j,t}$ allows each model to pick up on temporary increases in volatility that would be ill-represented by the more persistent variations modeled via the conventional SV processes for $\log \lambda_t$.[22]

Here we provide a closer comparison of the outlier estimates obtained from SVO and SV-t. For ease of comparison, we focus on three regions for possible realizations of the outlier states $o_{j,t}$: below 2, between 2 and 5, and above 5, corresponding to the cases of no (or small) outliers, moderate, and large outliers, respectively.[23] Focusing on a few selected variables in the interest of chart readability, Figures 4–7 display posterior probabilities of $o_{j,t}$ to have fallen in one of the three regions at a particular point in time. The three regions cover the possible support of $o_{j,t}$, and posterior probabilities of a realized outlier to fall in the three regions sum to one. Each figure directly reports probabilities of outliers having fallen between values of 2 and 5 or having been larger than 5, with the complement being the region of values below 2. Each figure compares results from SVO and SV-t for a different variable that result in a common message: Echoing our discussion of each model's properties in Section 2, SV-t sees outliers as being more moderately sized but occurring also more regularly than SVO, which tends to see outlier states to be larger than 5 (when they occur).

[22] In our application, $\log \lambda_t$ follows a multivariate random walk. Similar concerns about leakage from short-lived volatility spikes into estimates of $\log \lambda_t$ apply in the case of highly persistent, but stationary, processes for $\log \lambda_t$ as used elsewhere as well; see, among others, Clark and Ravazzolo (2015).

[23] As described in Section 2, the support of $o_{j,t}$ in the SVO model is between 1 and 20, and in the SV-t case the support is given by the positive portion of the real line. In both cases, the priors place most of their mass on realizations of $o_{j,t}$ around 1 as shown in Figure 2. In the SV-t case, the remaining mass of the prior is largely assigned to values below 5, whereas SVO places equal mass on values between 2 and 20.

For selected variables, Figure 8 also reports prior and posterior probability densities of the SVO parameter p_j, which describes the model's unconditional probability of seeing an outlier state value of 2 or more (at any given point in time). For some variables, like real income in Panel a of the figure, the posterior is shifted somewhat to the right of the prior, reflecting the relatively more frequent occurrence of outliers in this series discussed before.[24] For payroll gains, the posterior is more concentrated around the prior mean, as seen in Panel b, whereas the posteriors of p_j for other variables are shifted more to the left. Overall, and as it should be, the estimated probability of an outlier is quite low, with only negligible mass on values for p_j larger than 5 percent, even in the case of real income, and often below 2 percent for other variables.

Time variation in Σ_t affects our forecasts through two channels: first, the estimation of VAR coefficients Π as discussed in Section 2; 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), to construct predictive densities. The forecast results we have seen so far, for 1985 to 2017, seem to suggest that the latter channel is more relevant than the former, as the RMSE differences between SV and SVO (or SV-t) are very small, while those in CRPS are sometimes larger. The outlier states in SVO and SV-t allow for spikes in volatility 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 time-varying volatility in the residuals of different VAR equations as estimated in quasi-real time over the course of 2020.[25]

For each variable we report estimates generated by SV, SVO, and SV-t, as well as the persistent components of Σ_t imputed from SVO and SV-t when the effects of the outlier states are ignored. In this last case, we compute $\Sigma_t = A^{-1}\,\Lambda_t\,(A^{-1})'$ based on the SVO/SV-t

[24] As in Stock and Watson (2016), the prior for p_j is a beta distribution, centered around a mean of about 2 percent, consistent with having observed an outlier once every 4 years in 10 years' worth of monthly data.

[25] 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).

estimates for Λ_t and A^{-1}. In addition, we consider the corresponding measures of residual volatility obtained from the SV-OutMiss model, described in Section 2, that treats pre-specified outliers as missing data. Figure 9 displays estimates for payroll growth; further results are shown in our online appendix. Over the COVID-19 period, the SVO and SV-t models clearly differentiate between increases in uncertainty that are short- and longer-lived, which the SV model cannot do. SV estimates, shown in Panel (a) of the figure, reflect the impact of COVID-19 in the spring with a strong increase, which leveled off somewhat over the summer, but remained substantially elevated in the fall.

In contrast, SVO proves more nimble in accounting for the extreme data seen in the spring with a big, more than 30-fold, jump in volatility in April as shown in Panel (b) of the figure. However, as revealed by comparison with Panel (e), this jump is largely seen as transitory (both as it occurred in the spring and with the hindsight of estimates constructed based on data for the fall). The same is true, albeit to a somewhat more moderate extent, for the corresponding SV-t estimates shown in Panels (c) and (f). In contrast, the persistent components of volatility in the cases of SVO and SV-t are seen to have risen no more than 8-fold over the course of the year, while estimates from SV-OutMiss have risen by less than 5 times their level at the beginning of the year.

The more moderate rise in estimates of the persistent volatility components obtained with the SVO and SV-t specifications 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.[26] As discussed next, the aforementioned pattern in volatility estimates shown in Figure 9 is mirrored in out-of-sample forecast densities generated over the course of 2020.

[26] A similar conclusion emerges from an approach that adds dummies for each month since March 2020 to every VAR equation that is discussed in Section 6.

6 Forecasts made in 2020

As a reference for the pre-COVID situation, Figure 10 reports forecasts generated by the CONST and SV models in January 2020.[27] In January 2020, prior to the onset of COVID-19's economic effects, predictive densities generated from the CONST and SV models differ a little, but not markedly so. For example, as shown in Panel d of the figure, the CONST model saw the unemployment rate rise back above 4 percent over the course of 2021, consistent with a higher longer-run level of the unemployment rate, rather than the shallower path predicted by SV, which predicted the unemployment rate would remain mostly below 4 percent over the forecast horizon.

Things change dramatically over the course of March and April. The COVID-19 pandemic began to affect the US economy most visibly with the introduction of lockdown measures in the second half of March 2020, resulting in strong swings, particularly among measures of real activity, in subsequent months. Figures 11–14 display the evolution of forecasts for real income, payroll growth, stock market returns, and the unemployment rate over the months of March, April, and August generated from our alternative BVAR models.[28] As noted by Lenza and Primiceri (2020) and Schorfheide and Song (2020), forecasts generated by homoskedastic BVARs, like our CONST specification, display implausible behavior since the spring of 2020.[29] For example, CONST forecasts for the unemployment rate made in April run toward 80 percent with a 68 percent uncertainty band extending beyond rates of 0 percent and 100 percent by 2021; see Panel b of Figure 14.

In contrast, the reaction of point and density forecasts generated by the heteroskedastic VARs (SV, SVO, SV-t) to the incoming data in spring 2020 is much better behaved. Consid-

[27] Forecasts from the other alternatives, notably SVO and SV-t, are similar to those generated by the SV model in January 2020.

[28] 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.

[29] Lenza and Primiceri (2020) consider a slightly smaller VAR 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.

ering again the unemployment-rate forecasts shown in Figure 14, point forecasts from the SV model rise to about 18 percent by the end of 2020 and then fall back to about 14 percent by the end of 2021.[30] Generally, across variables and forecast origins, point forecasts generated by SVO and SV-t are fairly close, as seen also in our comparison of forecast performance pre-COVID-19 in Table 2.[31]

However, among our heteroskedastic VARs, stark differences emerge considering the uncertainty around forecasts made with and without outlier adjustments. As discussed in Section 5, the SV model sees only persistent changes in uncertainty, whereas SVO and SV-t distinguish between persistent and transitory shocks to volatility. Not surprisingly, in response to the incoming COVID-19 data in the spring of 2020, forecast densities from SV widen considerably. In the case of the unemployment rate, the 68 percent bands generated by SV expect outcomes to fall between 7 and 29 percent by late 2020 and between -10 and -40 percent by the end of 2022. In contrast, the corresponding SVO bands widened more moderately, though still noticeably, expecting the unemployment rate to fall between 14.5 and 21 percent by the end of 2020, and between 5 and 20 percent by late 2021. By comparison, in January, the SVO bands — much narrower before the pandemic — put the unemployment rate at the end of 2020 between 2.7 and 4.0 percent and at the end of 2021 between 2.2 and 5.0 percent.[32]

The middle-row panels of Figures 11–14 include a comparison between SVO and SV-t forecasts. As noted in Section 2, the SV-t model can be represented in a form that is very similar to the SVO model, differing only in the density of the *iid* outlier states. Overall, both forms of outlier-adjusted SV generate broadly similar forecast densities over the course of 2020, with a few instances of sizable differences, such as with forecasts of real income made

[30] These forecasts made in April jump off a reading for the unemployment rate of just under 15 percent.

[31] For better readability, forecasts generated by SV are displayed on different scales in the top and bottom rows of panels shown in Figures 11–14. Similarly, the SVO forecast densities shown in the top-row panels of these figures are also shown in the middle-row panels of each figure.

[32] The SVO forecasts made in January 2020 (not shown) are almost identical to those from the SV model displayed in Figure 10.

in April 2020 (see Panel e of Figure 11).[33]

Critically, both SVO and SV-t incorporate adjustments to *random* outliers that occur at unknown times. We also consider two procedures that condition on knowledge of when and which outliers occurred in the data. One criterion for the ex-ante identification of outliers is based on the distance from a data point to its sample median; the other reflects the timing of the COVID-19 pandemic.

When outliers are identified ex-ante, they could be treated as missing data, as we do with the SV-OutMiss approach in an otherwise standard VAR-SV model. In our application, observations that are more than 5 times the inter-quartile range away from their sample median are considered outliers.[34] The resulting forecast densities with jump-off points in 2020 are shown in the middle-row panels of Figures 11–14. In comparison to forecasts based on SVO (or SV-t), the forecast densities from SV-OutMiss tend to be narrower, in particular later in 2020, and dependent on the variables considered. For example, as indicated by the circled data points in Panel f of Figure 12, payroll growth data for the months of April, May, and June are treated as outliers in our application. The resulting 68 percent bands generated in August for annualized payroll growth in late 2021 range from -14 to -18 percent, whereas the corresponding SVO band is about twice as wide (ranging from about -30 to -33 percent). SV-OutMiss does not merely omit outlier data from the estimation of parameters and volatility states; the outliers are also ignored in the data vectors used to simulate predictive densities at every forecast origin. Ignoring the massive drop in payrolls recorded for April, by about 15 percent, also leads to some differences between near-term forecasts obtained from SV-OutMiss and SVO.[35] While SV-OutMiss sees positive payroll growth resuming by the third quarter of 2020 and returning to its longer-run average growth

[33]If anything, the 68 percent bands generated from SV-t tend to be a little wider than those from SVO in the month of April, when the most extreme observations occurred, while the bands are a little narrower in March and August. While small, these differences are consistent with the outlier specification employed by the SVO model being more geared toward filtering out large outliers.

[34]We obtain similar results with a threshold of 10 times the inter-quartile range.

[35]The drop in monthly payrolls of about 15 percent corresponds to an annualized rate of decline of about 85 percent, or an annualized log-change of -180 percent, which is the number shown in Figure 12.

rate of about 4 percent (annualized) by early 2021, SVO predicts a protracted slump until the second half of 2021 and a marked fall through the end of 2022.

As an alternative approach to handling outliers at known dates, we consider a further VAR specification, where each equation in (1) is augmented with dummies for every month since March 2020.[36] In light of the wild swings in at least some of the data, and for the purpose of soaking up potential outliers (rather than measuring average effects during COVID-19), *separate* dummies are added for each month since March, and wide priors are assigned to each dummy coefficient.[37] These dummies are applied to our BVAR model with SV, since the SV version of the model displayed generally beneficial qualities prior to the onset of the extreme observations of the COVID period. The bottom-row panels of Figures 11–14 compare the resulting forecasts for the months of March, April, and August. Strikingly, introduction of the outlier dummies to the BVAR with SV leads to point forecasts that are nearly identical to those obtained from SV and SVO without dummies. However, as the COVID-19 dummies soak up the residuals from every month since March, the width of the uncertainty bands remains stuck at levels estimated for the months prior to the economic onset of COVID-19, which appears to convey an unrealistically tight picture of forecast uncertainty since March.

In the supplementary online appendix, we report results of forecast comparisons limited to an evaluation covering only 2020; see Tables S.1– S.3 there. While the evaluation window is short, and realized values are, as of yet, scarce, some of the findings are interesting.[38] In particular, SV is generally better than CONST, with lower RMSE and CRPS for most

[36] Our dummy specification matters only for forecasts made in or after March 2020. In March 2020, one dummy is added to the VAR, two dummies are added in April, and so on.

[37] Denote the dummy coefficient for each month $t \geq$ 2020:03 by δ_t. The prior for each δ_t is a mean-zero normal distribution, with a large variance set equal to $1/\varepsilon$, where ε is a small number chosen as a function of machine precision (identical to the output of the `eps` function in MATLAB). For $t \geq$ 2020:03, only the sum of δ_t and the residual v_t are identified. In an OLS estimation, designed to minimize squared residuals, the dummy setup would result in $v_t = 0$ (for $t \geq$ 2020:03), whereas our Bayesian estimation will form predictions for these v_t identical to the posterior of the February residual, i.e., the last residual before the first non-zero dummy enters the system.

[38] Incoming data for the current year may, of course, also still be revised. With only six months of COVID data in hand, from March through September 2020, the evaluation is limited to one- and three-months-ahead forecasts, and no formal significance tests have been computed.

variables; similarly, CONST is outperformed by the other heteroskedastic model variants considered. Not surprisingly, given the strong similarity of predictive densities generated from SVO and SV-t discussed above, both forms of outlier-adjusted SV have performed comparably well so far. Since point forecasts obtained from SVO and SV-t in 2020 have (so far) also been similar to SV, so is their performance in terms of RMSE. The narrower forecast densities obtained from SVO and SV-t, compared to SV, also lead to slightly better CRPS at $h = 3$. Compared to these three models, SV-OutMiss has fared even somewhat better in terms of both RMSE and CRPS, suggesting that ignoring large outliers as start-off values for the forecasts may be particularly beneficial.[39]

7 Conclusion

We study the use of an outlier-augmented stochastic volatility specification for Bayesian VARs. This SVO approach extends 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). Our work is prompted by the enormous realizations of many macroeconomic time series witnessed over the course of 2020 as COVID-19 started to impact many economies across the world. As recognized by other recent studies such as Lenza and Primiceri (2020) and Schorfheide and Song (2020), these outliers have strong, and implausibly outsized, effects on forecasts made with standard constant-variance VARs. Instead, as VARs with time-varying volatility tend to down-weight high-volatility observations in the construction of parameter estimates, the resulting forecasts can be better insulated from outliers. As shown in Section 6, different variants of BVARs with time-varying volatility generate point forecasts that are less distorted than in the constant-variance case.

[39] In the supplementary online appendix, we also consider forecast performance during the Great Recession of 2007-09 and its aftermath. As shown in Tables S.5– S.7 there, SV has done generally better than CONST in terms of both RMSE and, in particular, CRPS during that period. SVO performed comparably, the differences are very small, as well as those between SVO and SV-OutMiss. A likely reason for this pattern is that few outliers are detected in this period, after properly accounting for volatility spikes.

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.[40] In contrast, SVO allows the model to fit sharp spikes in current volatility while adapting its uncertainty forecasts more moderately. The SVO model is related to an SV model with t-distributed errors, with SVO placing more prior mass on the occurrence of huge outliers. In our data, there are not many instances of such dramatic changes, as indicated by the frequency of observations far in the tails of the empirical density of the various data series considered in Figure 3. Although we prefer the SVO specification for being nimble in adapting to extreme and rare jumps in volatility like those seen since the onset of the recent pandemic, the SV-t model yields some of the same benefits to forecast accuracy. Pre-COVID, SVO performed quite similarly to SV-t in point and density forecasting, and densities generated with SVO and SV-t since the onset of COVID are broadly similar (though SVO seems to have filtered out more strongly the particular outliers seen in April 2020).

Of course, future data will be needed to assess which of the forecasts made in 2020 will end up being closer to the eventually realized data; and even then, the evaluation of density forecasts made this year will remain restricted to a limited sample of realized values. Nevertheless, we can take signal from an evaluation of simulated out-of-sample forecasts over a longer sample of post-1985 US data, described in Section 4. We find that SVO mildly outperforms standard SV, in particular in terms of density forecasts, while both display benefits over a constant-variance BVAR. In 2020, point forecasts generated from SV and SVO are very similar. But as SVO projections filter out the effects from short-lived outliers on forecast uncertainty, predictive densities constructed with SVO in 2020 widen by much less than those from SV. The ability of SVO to capture these extreme events, while otherwise retaining the beneficial performance of SV, is particularly appealing, and encouraging also for its use in current circumstances. Critically, SVO (and SV-t) treat the occurrence of outliers

[40] Typical implementations of SV differ, at times, in whether log-variances are modeled as random walks, or highly persistent though stationary processes. Concerns about undue extrapolation from a short-lived spike in volatility further into the future arise, however, in either case.

as stochastic events, with unknown timing. As a result, forecast uncertainty generated from both approaches is less compressed than what is obtained from approaches that treat outliers as known. More broadly, treating outliers as random events makes SVO and SV-t attractive for continued use over the yet-unknown course of economic developments related to the COVID-19 pandemic.

References

[1] Antolín-Díaz, Juan, Thomas Drechsel, and Ivan Petrella (2020), "Advances in nowcasting economic activity: Secular trends, large shocks and new data," (*mimeo)*, August.

[2] Artis, Michael J., Anindya Banerjee, and Massimiliano Marcellino (2005), "Factor forecasts for the UK," *Journal of Forecasting*, 24, 279–298, https://doi.org/10.1002/for.957.

[3] Aruoba, S. Boragan, Francis X. Diebold, and Chiara Scotti (2009), "Real-time measurement of business conditions," *Journal of Business & Economic Statistics*, 27, 417–427, https://doi.org/10.1198/jbes.2009.07205.

[4] Banbura, Marta, Domenico Giannone, and Lucrezia Reichlin (2010), "Large Bayesian vector auto regressions," *Journal of Applied Econometrics*, 25, 71–92, https://doi.org/10.1002/jae.1137.

[5] 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, https://doi.org/10.1016/j.jeconom.2019.04.024.

[6] Carriero, Andrea, Todd E. Clark, Massimiliano Marcellino, and Elmar Mertens (2020), "Measuring uncertainty and its effects in the COVID-19 era," Working Paper 20-32, Federal Reserve Bank of Cleveland, https://doi.org/10.26509/frbc-wp-202032.

[7] ——— (2021), "Forecasting with shadow-rate VARs," *mimeo*.

[8] Chan, Joshua C.C. (2019), "Large hybrid time-varying parameter VARs," *mimeo*.

[9] Chib, Siddhartha, Minchul Shin, and Fei Tan (2020), "DSGE-SVt: An econometric toolkit for high-dimensional DSGE models with SV and t Errors," *mimeo*.

[10] 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, https://doi.org/10.1016/j.ijforecast.2017.03.001.

[11] Clark, Todd E. (2011), "Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility," *Journal of Business and Economic Statistics*, 29, 327–341, https://doi.org/10.1198/jbes.2010.09248.

[12] Clark, Todd E., Michael W. McCracken, and Elmar Mertens (2020), "Modeling time-varying uncertainty of multiple-horizon forecast errors," *The Review of Economics and Statistics*, 102, 17–33, https://doi.org/10.1162/rest_a_00809.

[13] Clark, Todd E. and Francesco Ravazzolo (2015), "Macroeconomic forecasting performance under alternative specifications of time-varying volatility," *Journal of Applied Econometrics*, 30, 551–575, https://doi.org/10.1002/jae.2379.

[14] Cogley, Timothy and Thomas J. Sargent (2005), "Drift and volatilities: Monetary policies and outcomes in the post WWII U.S.," *Review of Economic Dynamics*, 8, 262–302, https://doi.org/10.1016/j.red.2004.10.009.

[15] Cúrdia, Vasco, Marco Del Negro, and Daniel L. Greenwald (2014), "Rare shocks, great recessions," *Journal of Applied Econometrics*, 29, 1031–1052, https://doi.org/10.1002/jae.2395.

[16] D'Agostino, Antonello, Luca Gambetti, and Domenico Giannone (2013), "Macroeconomic forecasting and structural change," *Journal of Applied Econometrics*, 28, 82–101, https://doi.org/10.1002/jae.1257.

[17] Diebold, Francis X. (2020), "Real-time real economic activity entering the Pandemic Recession," *Covid Economics*, 62, 1–19, CEPR.

[18] Diebold, Francis X. and Roberto S. Mariano (1995), "Comparing predictive accuracy," *Journal of Business and Economic Statistics*, 13, 253–63, https://doi.org/10.2307/1392185.

[19] Durbin, J. and S. J. Koopman (2002), "A simple and efficient simulation smoother for state space time series analysis," *Biometrika*, 89, 603–615, https://doi.org/10.1093/biomet/89.3.603.

[20] Guerrón-Quintana, Pablo and Molin Zhong (2017), "Macroeconomic forecasting in times of crises," Finance and Economics Discussion Series 2017-018, Board of Governors of the Federal Reserve System (U.S.).

[21] Holston, Kathryn, Thomas Laubach, and John C. Williams (2020), "Adapting the Laubach and Williams and Holston, Laubach, and Williams Models to the COVID-19 Pandemic," *mimeo*, Federal Reserve Bank of New York, May.

[22] Huber, Florian, Gary Koop, Luca Onorante, Michael Pfarrhofer, and Josef Schreiner (forthcoming), "Nowcasting in a pandemic using non-parametric mixed frequency VARs," *Journal of Econometrics*, https://doi.org/10.1016/j.jeconom.2020.11.006.

[23] 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, https://doi.org/10.1016/j.jeconom.2003.09.001.

[24] Johannsen, Benjamin K. and Elmar Mertens (forthcoming), "A time series model of interest rates with the effective lower bound," *Journal of Money, Credit and Banking*.

[25] Jurado, Kyle, Sydney C. Ludvigson, and Serena Ng (2015), "Measuring uncertainty," *American Economic Review*, 105, 1177–1216, https://doi.org/10.1257/aer.20131193.

[26] Justiniano, Alejandro and Giorgio E. Primiceri (2008), "The time-varying volatility of macroeconomic fluctuations," *The American Economic Review*, 98, 604–641, https://doi.org/10.1257/aer.98.3.604.

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

[28] 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, https://doi.org/10.1111/1467-937X.00050.

[29] Koop, Gary (2003), *Bayesian Econometrics*: Wiley-Interscience.

[30] Koop, Gary M. (2013), "Forecasting with medium and large Bayesian VARS," *Journal of Applied Econometrics*, 28, 177–203, https://doi.org/10.1002/jae.1270.

[31] Krüger, Fabian, Sebastian Lerch, Thordis L. Thorarinsdottir, and Tilmann Gneiting (2020), "Predictive Inference Based on Markov Chain Monte Carlo Output," arXiv:1608.06802.

[32] Lenza, Michele and Giorgio E. Primiceri (2020), "How to estimate a VAR after March 2020," NBER Working Paper 27771, National Bureau of Economic Research, Inc, https://doi.org/10.3386/w27771.

[33] McCracken, Michael W. and Serena Ng (2016), "FRED-MD: A monthly database for macroeconomic research," *Journal of Business & Economic Statistics*, 34, 574–589, https://doi.org/10.1080/07350015.2015.1086655.

[34] Perez-Quiros, Gabriel and Margaret M. McConnell (2000), "Output fluctuations in the United States: What has changed since the early 1980's?" *American Economic Review*, 90, 1464–1476, https://doi.org/10.1257/aer.90.5.1464.

[35] Schorfheide, Frank and Dongho Song (2020), "Real-time forecasting with a (standard) mixed-frequency VAR during a pandemic," Working Paper 20-26, Federal Reserve Bank of Philadelphia, https://doi.org/10.21799/frbp.wp.2020.26.

[36] Stock, James H. and Mark W. Watson (2002), "Macroeconomic forecasting using diffusion indexes," *Journal of Business & Economic Statistics*, 20, 147–162, https://doi.org/10.1198/073500102317351921.

[37] ——— (2007), "Why has U.S. inflation become harder to forecast?" *Journal of Money, Credit and Banking*, 39, 3–33, https://doi.org/10.1111/j.1538-4616.2007.00014.x.

[38] ——— (2016), "Core inflation and trend inflation," *The Review of Economics and Statistics*, 98, 770–784, https://doi.org/10.1162/REST_a_00608.

[39] West, Kenneth D. (1996), "Asymptotic inference about predictive ability," *Econometrica*, 64, 1067–1084, https://doi.org/10.2307/2171956.

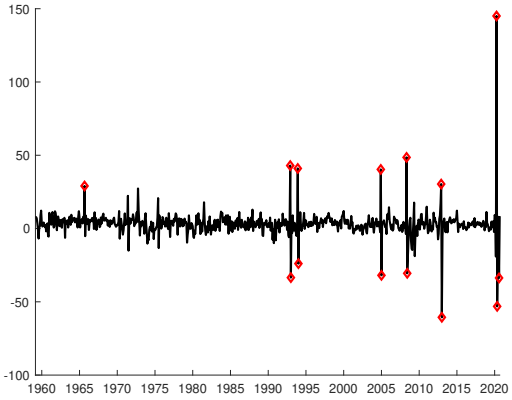
Table 1: List of variables

Variable	FRED-MD code	tcode
Real Income	RPI	5
Real Consumption Exp.	DPCERA3M086SBEA	5
IP	INDPRO	5
Capacity Utilization	CUMFNS	1
Unemployment Rate	UNRATE	1
Nonfarm payrolls	PAYEMS	5
Hours	CES0600000007	1
Hourly Earnings	CES0600000008	5
PPI: Finished Goods	WPSFD49207	5
PCE prices	PCEPI	5
Housing Starts	HOUST	4
S&P 500	SP 500	5
USD / GBP FX rate	EXUSUKx	5
5-Year yield	GS5	1
10-Year yield	GS10	1
Baa spread	BAAFFM	1

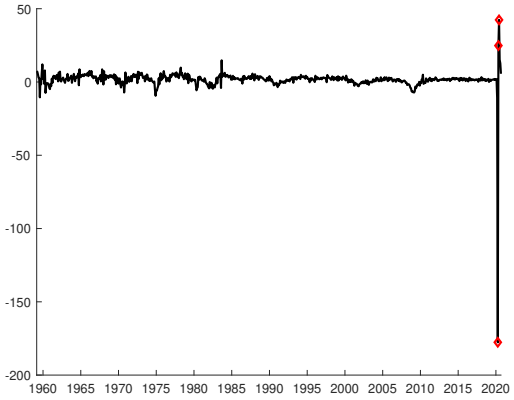
Note: Data obtained from the 2020-10 vintage of FRED-MD. Monthly observations from 1959:M03 to 2020:M09. The column tcode denotes the following data transformation for a series x: (1) no transformation; (2) Δx_t; (3) $\Delta^2 x_t$; (4) $\log(x_t)$; (5) $\Delta \log(x_t) \cdot 1200$; (6) $\Delta^2 \log(x_t)$; (7) $\Delta(x_t/x_{t-1} - 1.0)$.

Figure 1: Some selected data series

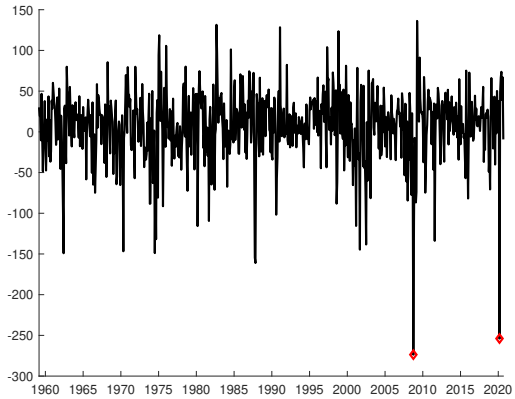
(a) Real personal income



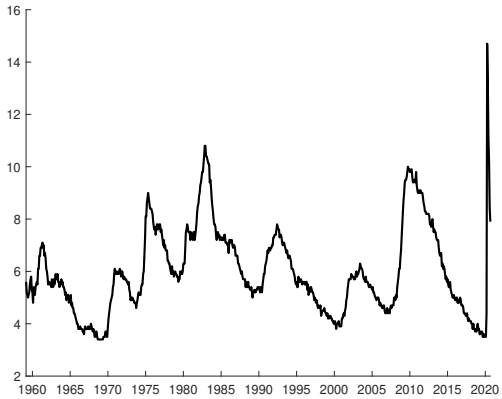
(b) Payroll growth



(c) S&P 500 returns



(d) Unemployment rate

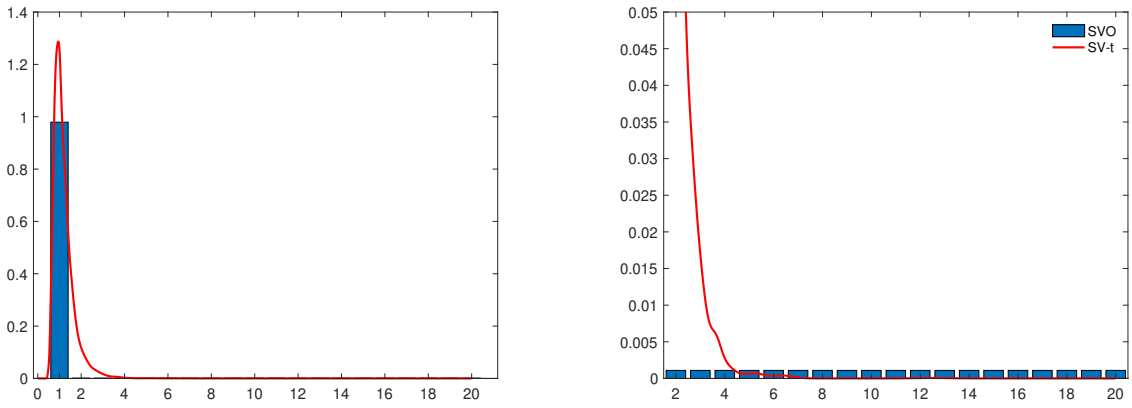


Note: Data for selected time series, with data transformations as listed in Table 1. Red dots denote observations that are more than five times the inter-quartile range away from the series median.

Figure 2: Densities of outlier states: SVO vs. SV-t

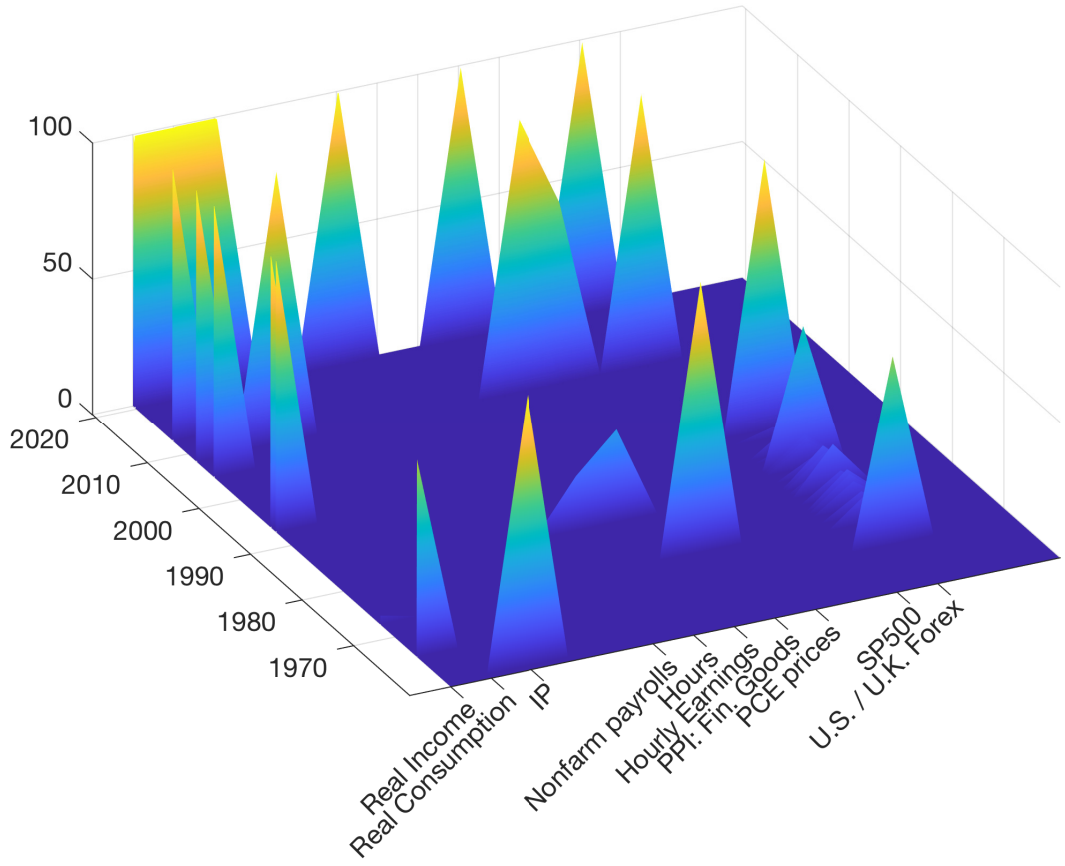
(a) Full figure

(b) Zoomed into right tail



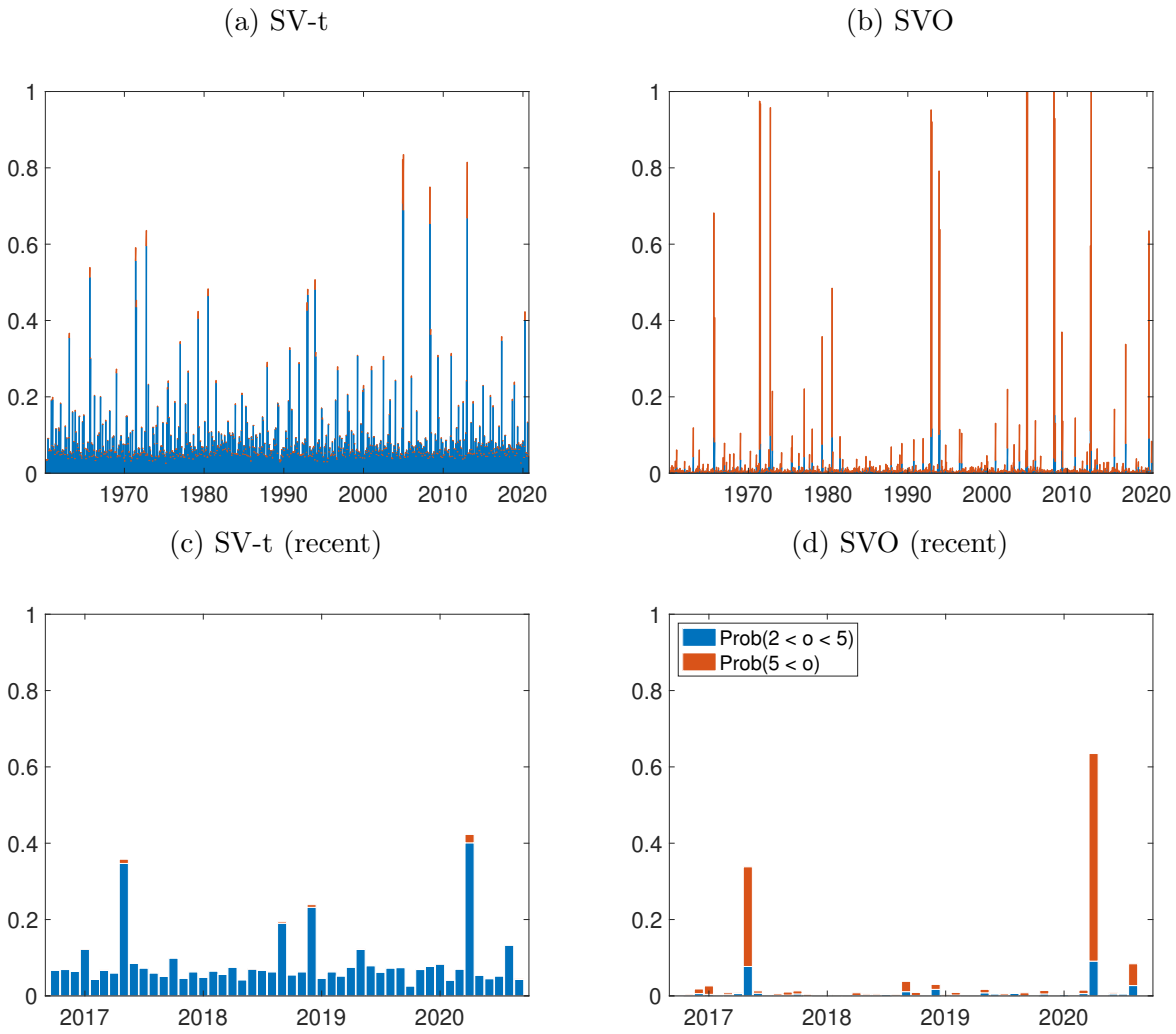
Note: Densities for the outlier state $o_{j,t}$ in the SVO model (blue bars) and the SV-t model of (23). The degrees of freedom for the SV-t model have been set equal to five, which generates roughly the same variance for $o_{j,t}$ as in the calibration of the SVO model, where the outlier probability p has been set to correspond to one outlier every four years in monthly data, $p = 1/(4 \cdot 12)$. (As discussed in the text, the variance would be equalized with a choice for the degrees of freedom equal to 5.70, generating a visually identical picture.)

Figure 3: Potential outliers in the data



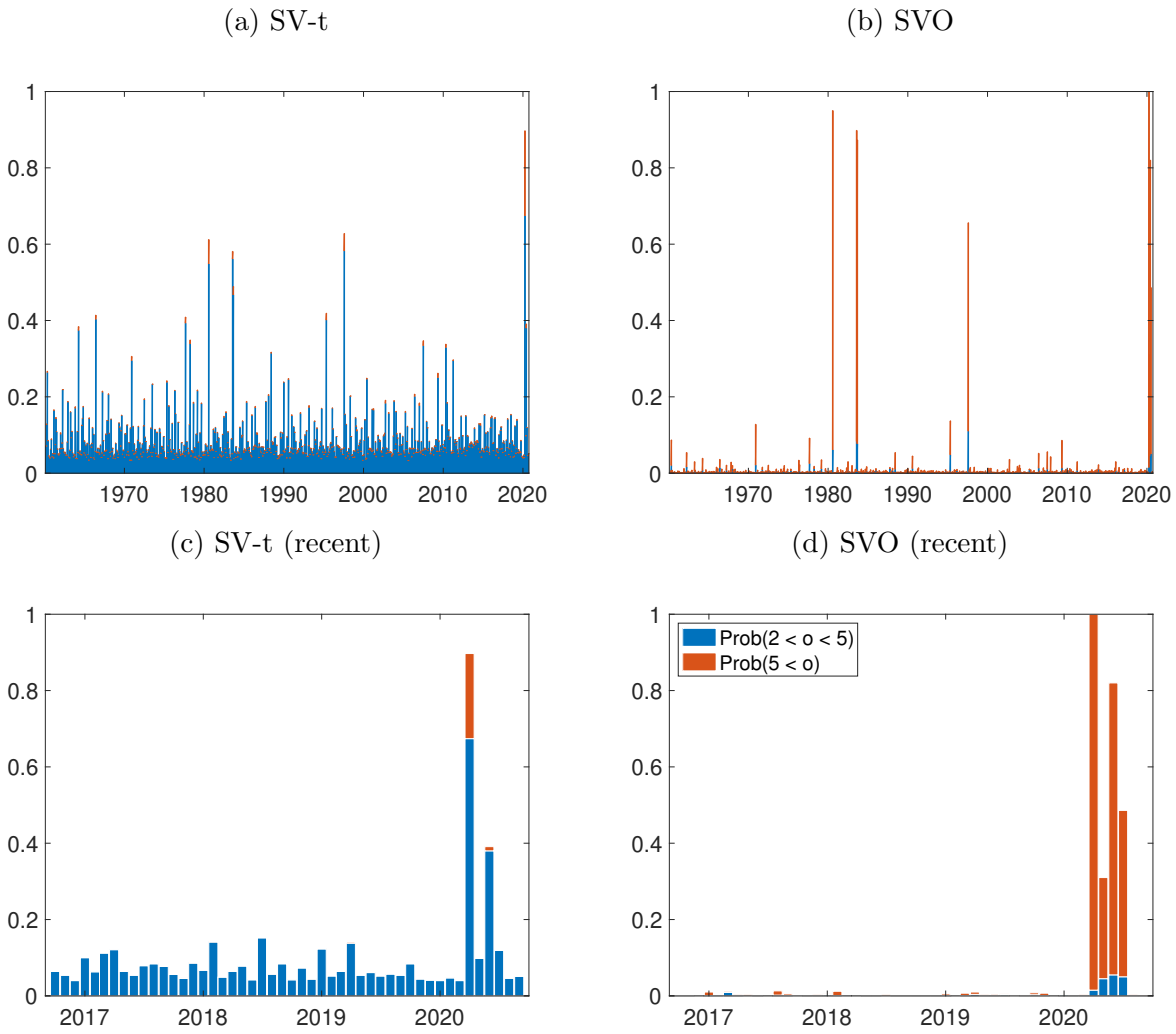
Note: Occurrence of potential outliers in our 16-variable data set (as described in Table 1). Potential outliers are identified as observations that are more than five times the interquartile range away from the series median in a given sample. In quasi-real time, the assessment may change, and the graph above indicates the average occurrence (in percentage points) of an observations being designated as outlier over all quasi-real-time samples that include a given observation. We consider growing quasi-real-time samples, all starting in 1959:M3 with the first sample ending in 1985:M1.

Figure 4: Posteriors of outlier states for real income



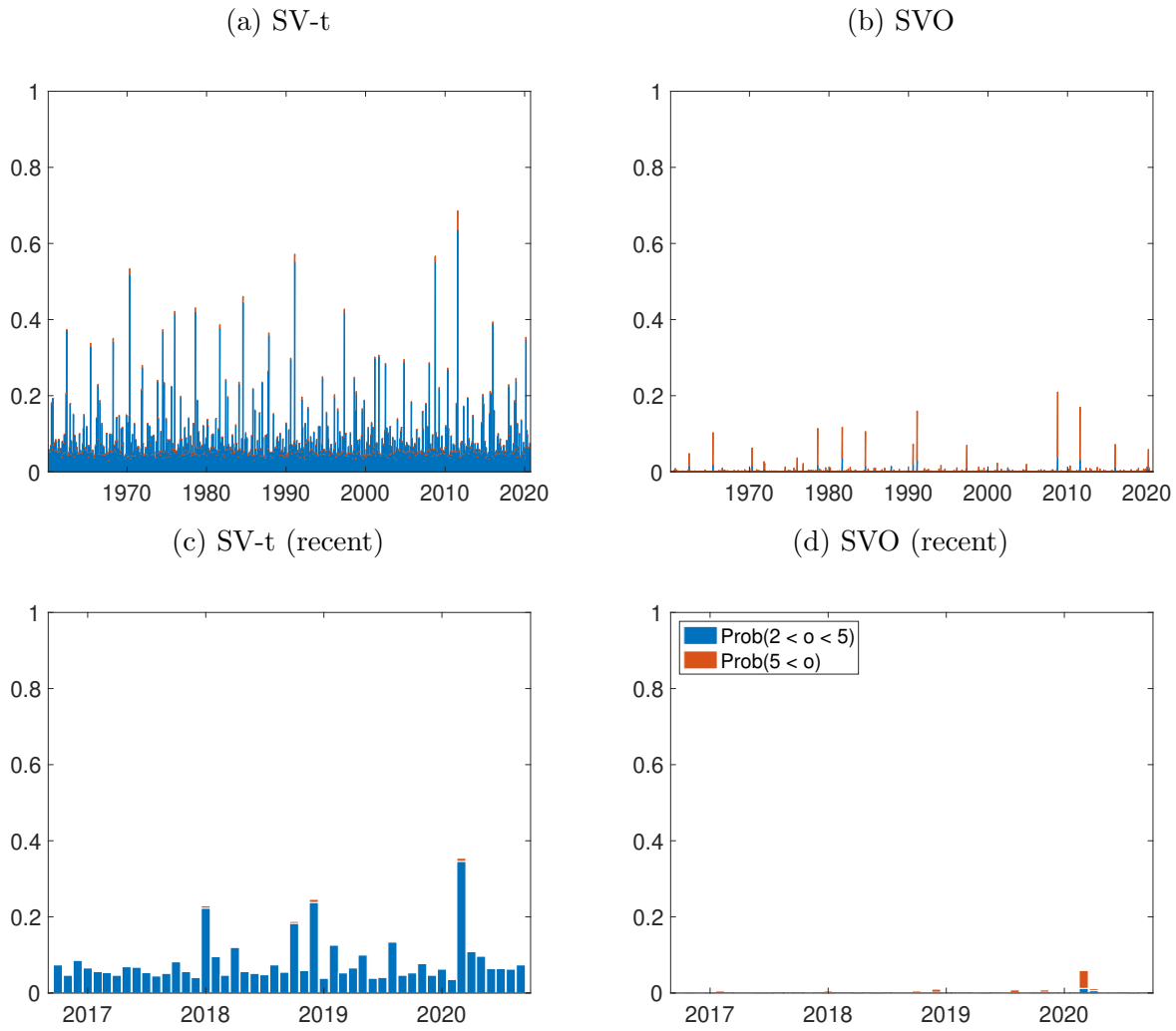
Note: Full-sample estimates per September 2020 of posterior probabilities for realizations of $o_{j,t}$ in SV-t and SVO models. Each panel shows posterior probabilities for $o_{j,t}$ to fall into a range between two and five (blue bars) or to be larger than five (orange bars) in a given month of the sample. The lower row of panels zooms in on results for the last few years (numbers are identical to the corresponding results in the upper-row panels). The SV-t model is estimated with five degrees of freedom.

Figure 5: Posteriors of outlier states for payroll growth



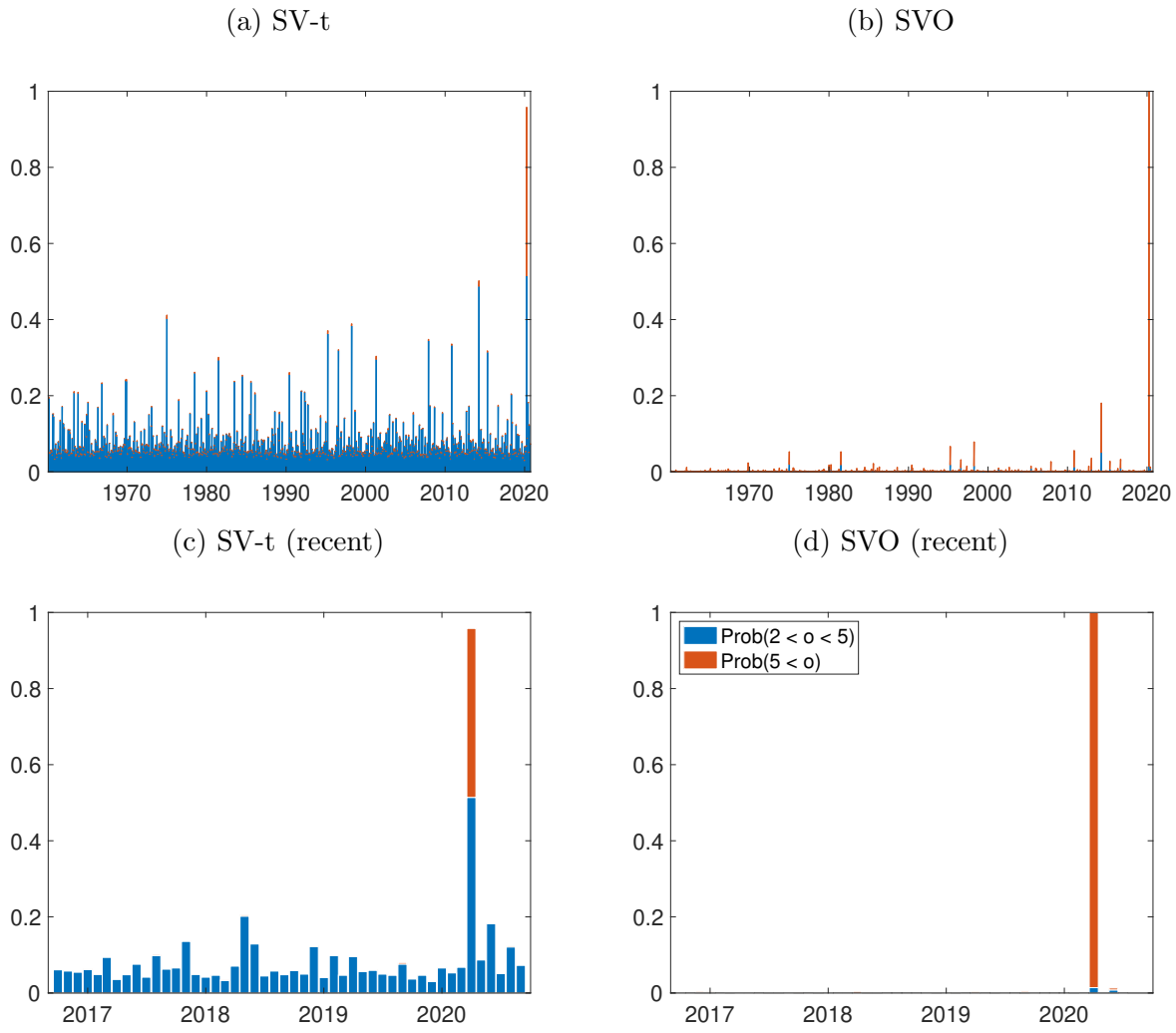
Note: Full-sample estimates per September 2020 of posterior probabilities for realizations of $o_{j,t}$ in SV-t and SVO models. Each panel shows posterior probabilities for $o_{j,t}$ to fall into a range between two and five (blue bars) or to be larger than five (orange bars) in a given month of the sample. The lower row of panels zooms in on results for the last few years (numbers are identical to the corresponding results in the upper-row panels). The SV-t model is estimated with five degrees of freedom.

Figure 6: Posteriors of outlier states for S&P 500 returns



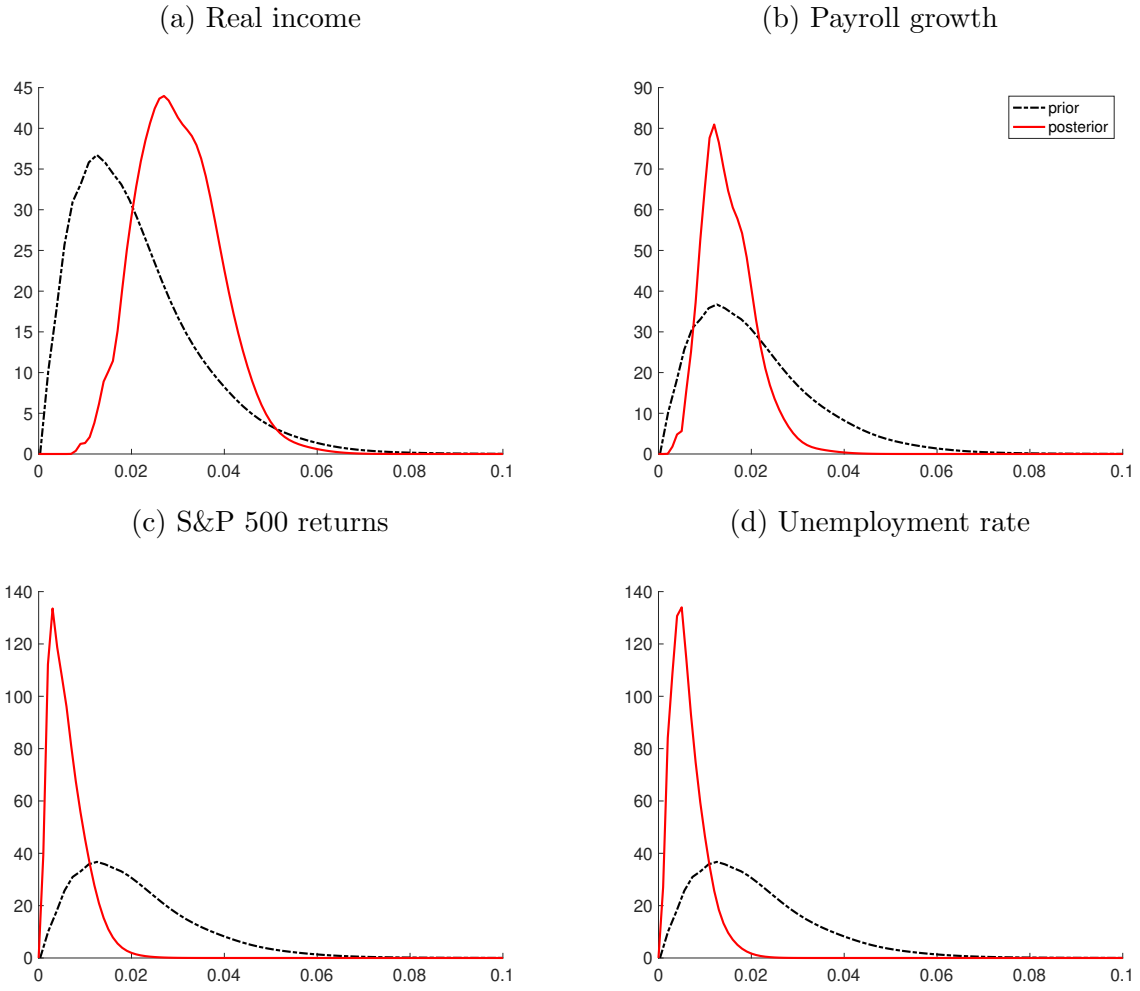
Note: Full-sample estimates per September 2020 of posterior probabilities for realizations of $o_{j,t}$ in SV-t and SVO models. Each panel shows posterior probabilities for $o_{j,t}$ to fall into a range between two and five (blue bars) or to be larger than five (orange bars) in a given month of the sample. The lower row of panels zooms in on results for the last few years (numbers are identical to the corresponding results in the upper-row panels). The SV-t model is estimated with five degrees of freedom.

Figure 7: Posteriors of outlier states for the unemployment rate



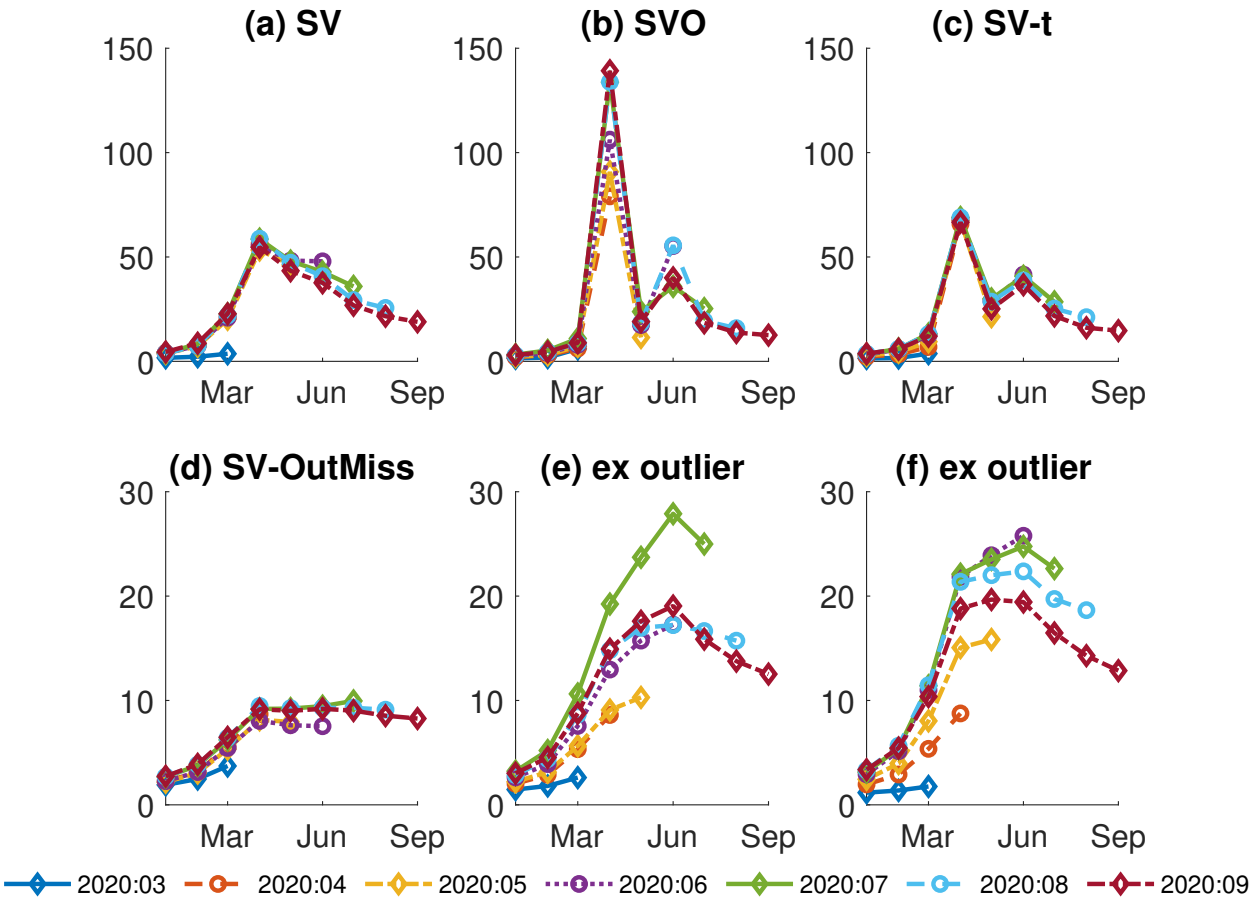
Note: Full-sample estimates per September 2020 of posterior probabilities for realizations of $o_{j,t}$ in SV-t and SVO models. Each panel shows posterior probabilities for $o_{j,t}$ to fall into a range between two and five (blue bars) or to be larger than five (orange bars) in a given month of the sample. The lower row of panels zooms in on results for the last few years (numbers are identical to the corresponding results in the upper-row panels). The SV-t model is estimated with five degrees of freedom.

Figure 8: SVO outlier probabilities



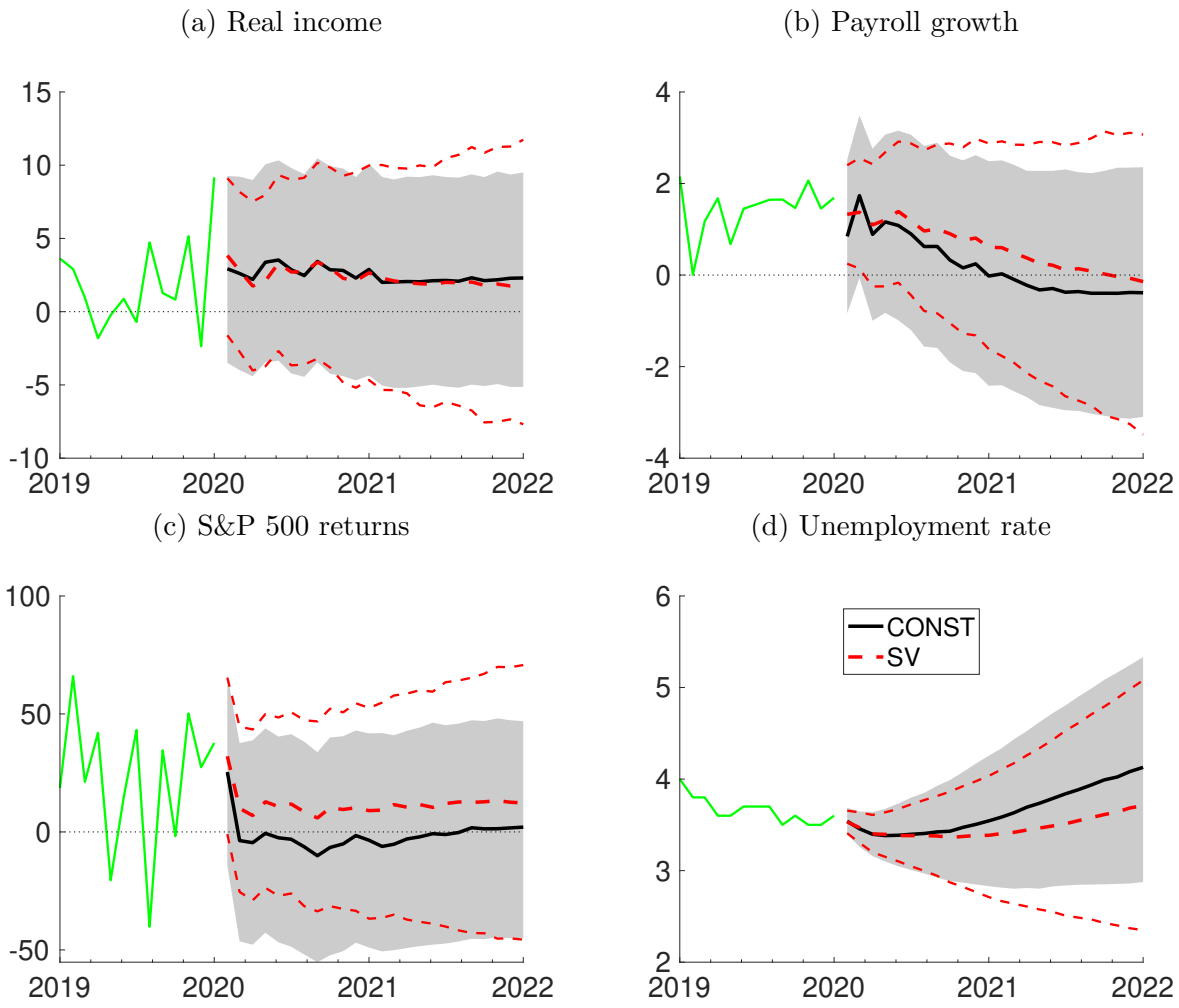
Note: Prior and posterior distribution of the outlier probability p_j in the SVO model for selected variables, estimated from the full sample of data available from March 1959 through September 2020.

Figure 9: Time-varying volatilities in 2020 of payroll growth



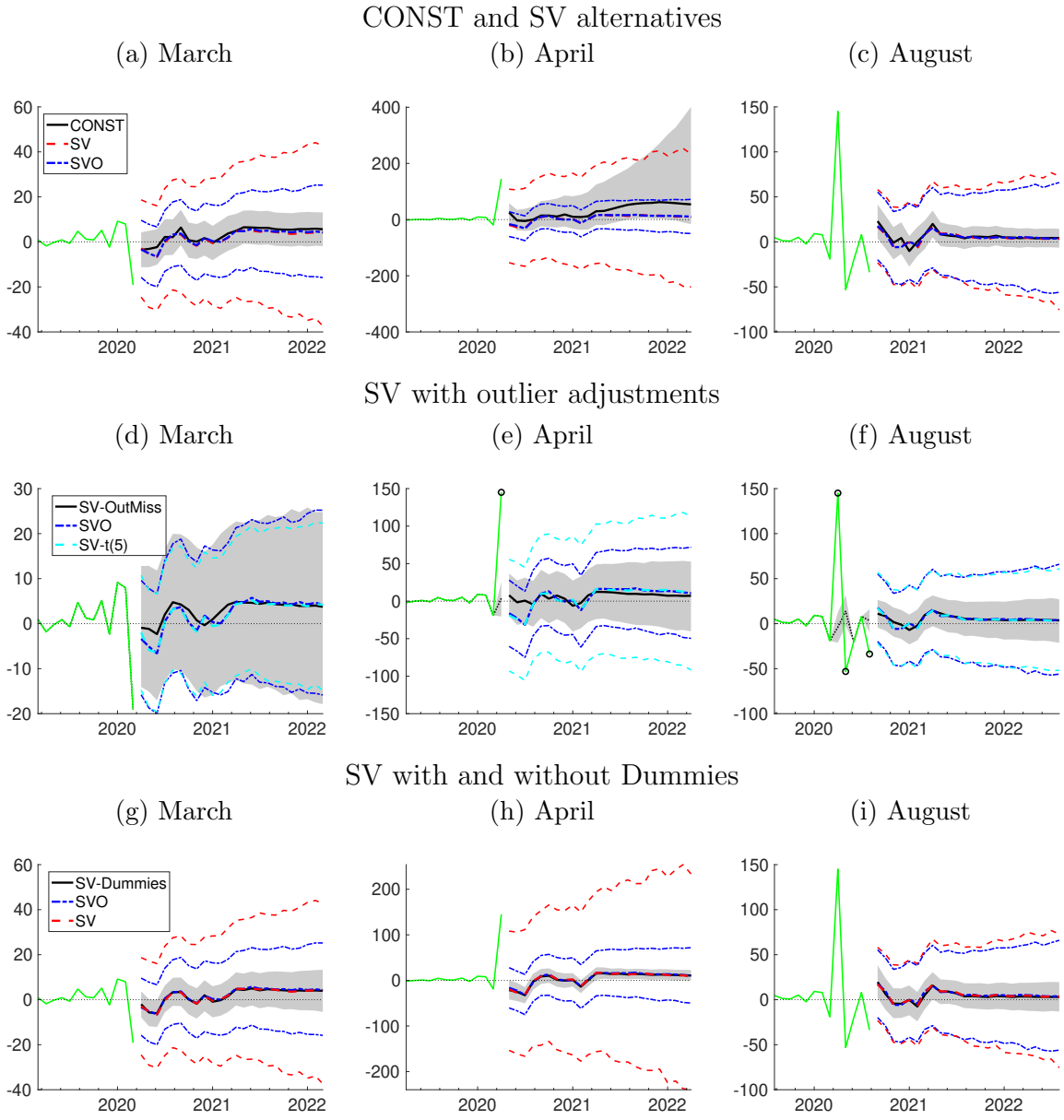
Note: Quasi-real-time trajectories of time-varying volatility in VAR residuals, measured by the diagonal elements of Σ_t. Medians of (smoothed) posterior obtained from different data samples ending at forecast origins as indicated in the figure legend. Panels (e) and (f) display estimates of stochastic volatility for SVO and SV-t, respectively, that ignore the contributions from outliers computed from $\Sigma_t = A^{-1}\,\Lambda_t\,A^{-T}$ (i.e., neglecting the O_t components in the computation of the uncertainty measures shown here, while including 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 10: Predictive densities in January 2020 for selected variables



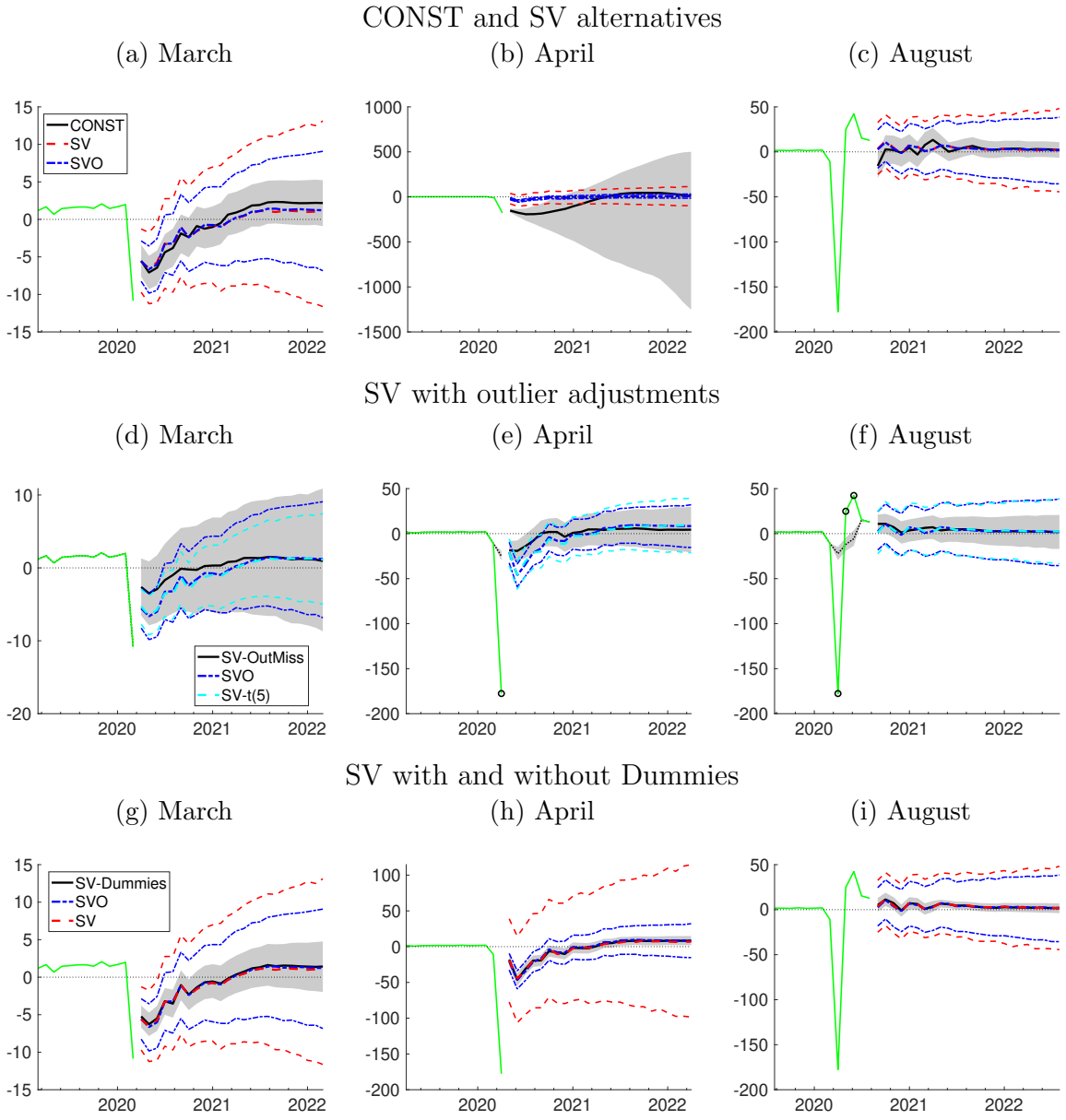
Note: Medians and 68% uncertainty bands of predictive densities, simulated out-of-sample at various forecast origins as indicated in each panel. The solid green line denotes realized data prior to the forecast origin.

Figure 11: Predictive densities since March 2020 for real income



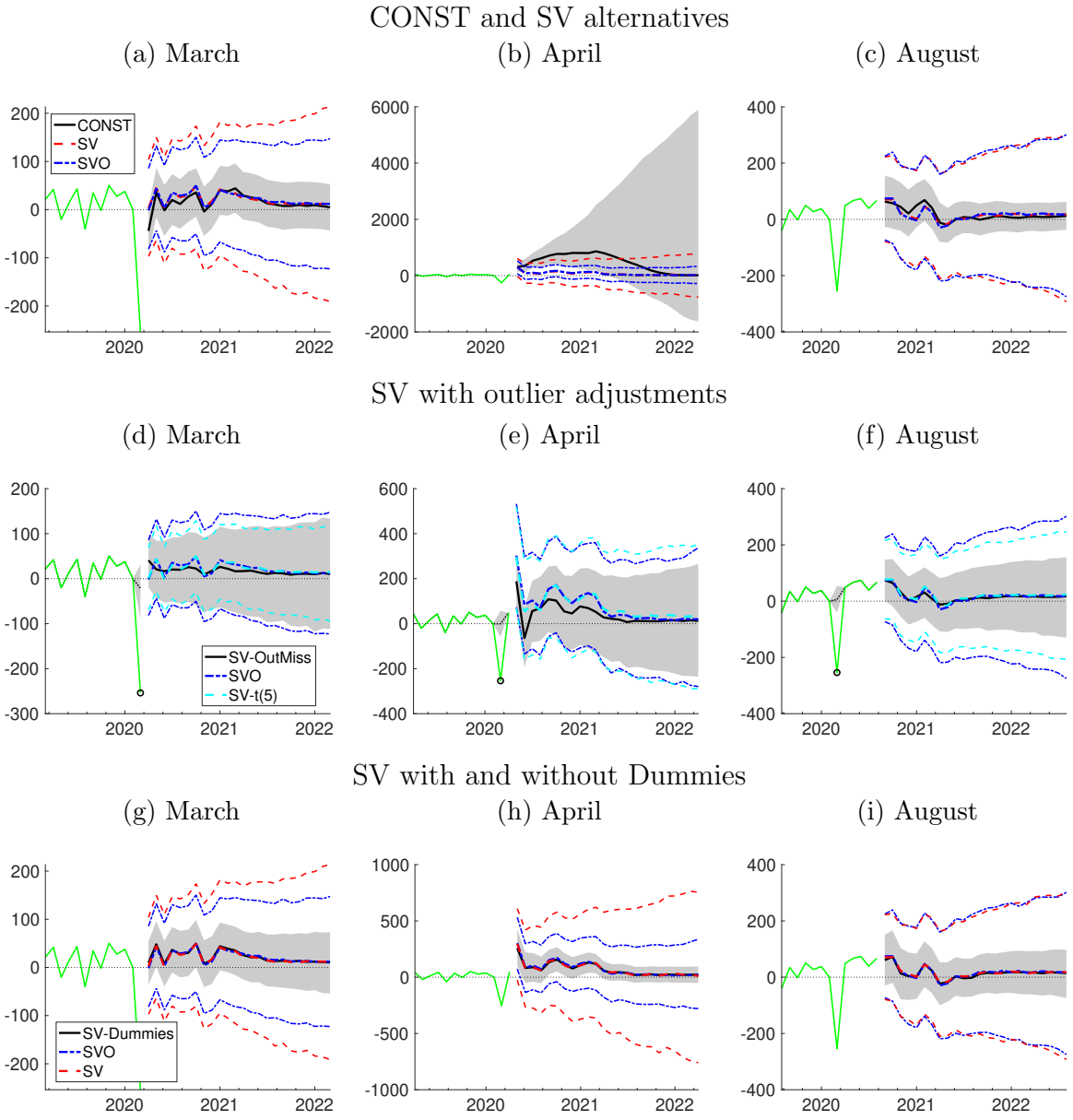
Note: Medians and 68% uncertainty bands of predictive densities, simulated out-of-sample at various forecast origins as indicated in each panel. The solid green line denotes realized data prior to the forecast origin. In panels (d) – (f), observations identified ex-ante as outliers, based on being more than 10 times the inter-quartile range away from the median, are indicated with a circle, and the corresponding backcast densities from the SV-OutMiss model are superimposed.

Figure 12: Predictive densities since March 2020 for payroll growth



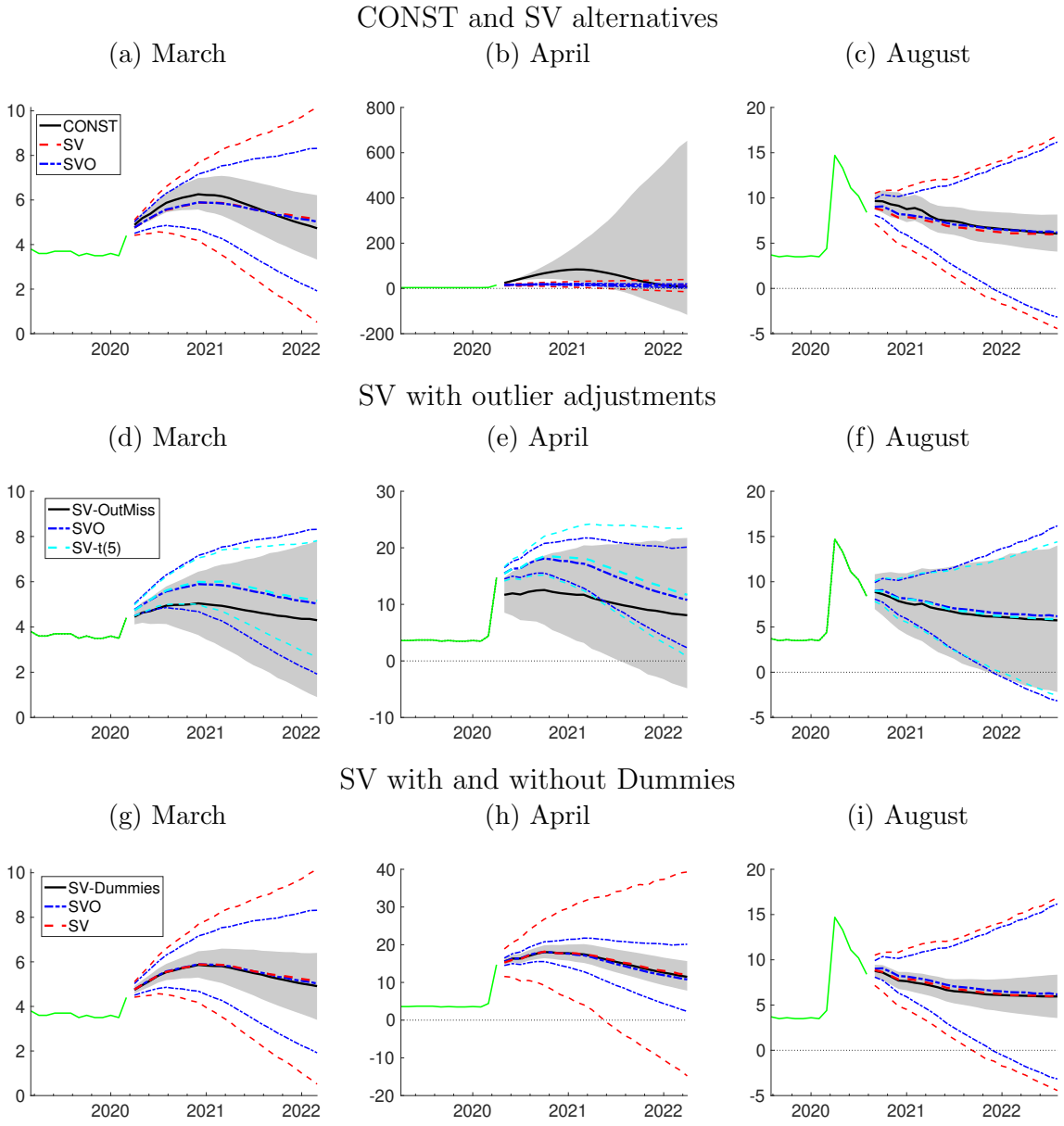
Note: Medians and 68% uncertainty bands of predictive densities, simulated out-of-sample at various forecast origins as indicated in each panel. The solid green line denotes realized data prior to the forecast origin. In panels (d) – (f), observations identified ex-ante as outliers, based on being more than 10 times the inter-quartile range away from the median, are indicated with a circle, and the corresponding backcast densities from the SV-OutMiss model are superimposed.

Figure 13: Predictive densities since March 2020 for S&P 500 returns



Note: Medians and 68% uncertainty bands of predictive densities, simulated out-of-sample at various forecast origins as indicated in each panel. The solid green line denotes realized data prior to the forecast origin. In panels (d) – (f), observations identified ex-ante as outliers, based on being more than 10 times the inter-quartile range away from the median, are indicated with a circle, and the corresponding backcast densities from the SV-OutMiss model are superimposed.

Figure 14: Predictive densities since March 2020 for the unemployment rate



Note: Medians and 68% uncertainty bands of predictive densities, simulated out-of-sample at various forecast origins as indicated in each panel. The solid green line denotes realized data prior to the forecast origin. In panels (d) – (f), observations identified ex-ante as outliers, based on being more than 10 times the inter-quartile range away from the median, are indicated with a circle, and the corresponding backcast densities from the SV-OutMiss model are superimposed.

Table 2: Relative RMSE

					Relative to CONST ...							
	CONST				SV				SVO			
Variable / Horizons	1	3	12	24	1	3	12	24	1	3	12	24
Real Income	7.63	7.69	7.75	7.82	1.00	1.00	1.00	1.19	1.00	0.99	1.00	1.01
Real Consumption	5.38	5.61	5.37	5.08	1.00	0.99	0.99	1.03*	1.00	0.99	0.99	0.99
IP	6.90	7.18	7.76	8.27	0.97**	1.00	1.01	1.01	0.97	0.99	1.01	0.98**
Capacity Utilization	0.47	0.87	2.79	4.04	0.96***	0.98	1.00	1.03	0.97	0.98	0.99	1.01
Unemployment	0.14	0.22	0.74	1.40	1.01	1.02	1.00	0.94	1.01	1.02	0.99*	0.94
Nonfarm payrolls	1.28	1.36	1.92	2.14	0.95***	0.98	0.97	0.96	0.95	0.97	0.97	0.95
Hours	0.20	0.23	0.42	0.46	0.96***	1.00	0.96	0.94	0.97	1.01	0.95	0.92
Hourly Earnings	2.56	2.49	2.62	2.95	0.96**	0.99	0.97	0.94	0.97	0.98	0.96**	0.92
PPI (fin. goods)	7.19	7.48	8.09	8.13	1.00	0.99	0.94*	0.93*	1.00	0.98	0.94	0.92
PCE prices	2.13	2.52	3.05	3.48	0.99	0.96	0.87**	0.83**	0.99	0.96	0.87	0.82
Housing Starts	0.07	0.10	0.24	0.38	1.01	1.01	0.96	0.91**	1.01	1.00	0.95	0.92
S&P 500	43.48	44.57	44.15	43.36	1.01	0.99	1.00	1.02	1.01	0.99	1.00	1.00
USD / GBP FX rate	29.12	30.41	28.77	28.55	0.97**	0.98	0.99	1.20	0.97	0.97	0.99	0.98
5-Year yield	0.27	0.61	1.22	1.33	0.94***	0.93***	0.96	1.13**	0.94	0.92	0.97	1.11
10-Year yield	0.25	0.56	1.20	1.27	0.97**	0.94***	0.94	1.01	0.97	0.93	0.93	1.00
Baa spread	0.32	0.76	1.51	1.73	0.82***	0.79**	0.89*	0.87	0.82	0.78**	0.87*	0.87

Note: Comparison of "CONST" (baseline, in denominator of relative comparisons) against "SV" and "SVO." Values below one indicate improvement over baseline. Evaluation window from 1985:M01 through 2017:M12. Significance assessed by Diebold-Mariano test using Newey-West standard errors with $h + 1$ lags.

Table 3: Relative Avg CRPS

					Relative to CONST ...							
	CONST				SV				SVO			
Variable / Horizons	1	3	12	24	1	3	12	24	1	3	12	24
Real Income	3.44	3.31	3.43	3.48	0.93***	1.01	1.07**	1.23***	0.92	0.97**	1.04*	1.13***
Real Consumption	2.88	3.01	2.99	2.93	0.96**	0.95***	1.02	1.22***	0.96*	0.96**	1.04***	1.21
IP	3.81	3.95	4.30	4.55	0.96***	0.98	1.06***	1.21***	0.96*	0.99**	1.08***	1.21
Capacity Utilization	0.27	0.50	1.53	2.29	0.96***	0.96**	1.01	1.11**	0.97**	0.97***	1.03**	1.13
Unemployment	0.08	0.12	0.39	0.75	1.01	1.00	0.97	0.96	1.01*	1.01**	0.99**	0.98**
Nonfarm payrolls	0.77	0.83	1.12	1.25	0.90***	0.93***	1.01	1.19***	0.91	0.94	1.04***	1.21*
Hours	0.11	0.13	0.24	0.27	0.93***	0.98	0.98	1.10*	0.93	0.99*	0.98	1.08
Hourly Earnings	1.48	1.47	1.56	1.73	0.93***	0.95***	1.01	1.17***	0.94*	0.96**	1.04***	1.18**
PPI (fin. goods)	3.84	3.97	4.30	4.32	0.98	0.98	0.97	1.07*	0.98	0.98	0.98**	1.08
PCE prices	1.14	1.32	1.60	1.80	0.97**	0.96**	0.92*	1.01	0.97	0.97**	0.94***	1.04***
Housing Starts	0.04	0.05	0.12	0.19	1.00	1.00	0.94	0.96	1.01*	1.01***	0.96***	0.99**
S&P 500	23.01	23.71	23.57	23.18	1.00	0.99	1.06***	1.23***	1.00**	0.99	1.07*	1.21**
USD / GBP FX rate	16.06	16.47	15.83	15.77	0.97**	0.98	1.05**	1.17***	0.96**	0.97	1.03***	1.10***
5-Year yield	0.15	0.34	0.69	0.79	0.93***	0.92***	0.94	1.06	0.93	0.92	0.96**	1.09***
10-Year yield	0.14	0.31	0.67	0.76	0.96***	0.93***	0.93	1.02	0.96	0.94	0.95**	1.06***
Baa spread	0.18	0.42	0.85	1.00	0.74***	0.76***	0.90*	1.02	0.74	0.76	0.91	1.06***

Note: Comparison of "CONST" (baseline, in denominator of relative comparisons) against "SV" and "SVO." Values below one indicate improvement over baseline. Evaluation window from 1985:M01 through 2017:M12. Significance assessed by Diebold-Mariano test using Newey-West standard errors with $h+1$ lags. Due to the close behavior of some of the models compared, and rounding of the report values, one of the comparisons shows a significant relative CRPS of 1.00. This case arises from persistent differences in performance that are, however, too small to be relevant after rounding.

Table 4: Relative RMSE (alternative models)

					Relative to SVO ...							
	SVO				SV-t(5)				SV-OutMiss			
Variable / Horizons	1	3	12	24	1	3	12	24	1	3	12	24
Real Income	7.63	7.64	7.75	7.89	1.00	1.00***	1.00	1.00	1.00	1.00	1.00	0.98
Real Consumption	5.37	5.55	5.32	5.04	1.01	1.01**	1.01*	1.00	0.99	0.99	1.00*	1.01
IP	6.72	7.12	7.81	8.08	1.01	1.00	1.00	1.00	1.00	1.01	1.00*	1.02
Capacity Utilization	0.46	0.85	2.76	4.10	1.00	1.00	1.02*	1.02	1.00	1.02	1.00	1.00
Unemployment	0.14	0.22	0.73	1.32	1.00	1.00	1.01	1.01	0.99	1.00	1.00	1.00
Nonfarm payrolls	1.22	1.32	1.86	2.03	1.00	1.00	1.02**	1.03**	0.99	1.00	1.00**	1.01
Hours	0.20	0.23	0.40	0.43	1.00	0.99**	1.01	1.01	1.00	1.00	1.01	1.01
Hourly Earnings	2.48	2.45	2.53	2.73	1.00	1.00	1.02***	1.02	1.00	1.00	1.01*	1.02
PPI (fin. goods)	7.16	7.36	7.59	7.45	1.00	1.00	1.00	1.00	0.99*	0.99	1.00	1.01*
PCE prices	2.11	2.43	2.64	2.86	1.00	1.00	1.01	1.01	1.00	0.99**	1.01	1.01
Housing Starts	0.07	0.10	0.22	0.35	1.01***	1.01**	1.03*	1.04	1.00	1.00	1.00**	1.00
S&P 500	43.71	44.15	44.14	43.57	1.00	1.00	1.00	1.00	1.01	1.00	1.00**	1.00
USD / GBP FX rate	28.14	29.61	28.34	28.05	1.01	1.01	1.00	1.03*	1.02	1.00	1.00	1.01**
5-Year yield	0.25	0.56	1.18	1.48	1.00	1.01**	1.03***	0.99	0.99	0.99**	0.99*	0.97*
10-Year yield	0.24	0.53	1.12	1.26	1.00	1.01**	1.03***	1.02	1.00	1.00	1.00*	0.99**
Baa spread	0.26	0.60	1.30	1.50	1.01	1.01	1.02	0.97	1.02	1.01	1.01	1.02**

Note: Comparison of "SVO" (baseline, in denominator of relative comparisons) against "SV-t(5)" and "SV-OutMiss." Values below one indicate improvement over baseline. Evaluation window from 1985:M01 through 2017:M12. Significance assessed by Diebold-Mariano test using Newey-West standard errors with $h + 1$ lags. Due to the close behavior of some of the models compared, and rounding of the report values, a few comparisons show a significant relative RMSE (alternative models) of 1.00. These cases arise from persistent differences in performance that are, however, too small to be relevant after rounding.

Table 5: Relative Avg CRPS (alternative models)

					Relative to SVO ...							
	SVO				SV-t(5)				SV-OutMiss			
Variable / Horizons	1	3	12	24	1	3	12	24	1	3	12	24
Real Income	3.15	3.23	3.56	3.93	1.00	1.00	0.97***	0.94***	0.99	0.97**	0.96	0.94
Real Consumption	2.75	2.88	3.10	3.55	1.00	0.99**	0.94***	0.91***	0.98	0.98	0.96***	0.95***
IP	3.68	3.90	4.63	5.51	1.00	0.99***	0.95***	0.90***	0.99	1.00	0.97**	0.96***
Capacity Utilization	0.26	0.49	1.58	2.58	1.00	0.98***	0.98	0.95**	0.99	0.99	0.97	0.95
Unemployment	0.08	0.12	0.39	0.74	0.99**	0.99*	0.99	0.98	0.99	0.98	0.97	0.97
Nonfarm payrolls	0.70	0.78	1.16	1.52	0.99**	0.98***	0.95***	0.91***	0.98	0.98	0.96	0.94***
Hours	0.11	0.13	0.23	0.29	0.99***	0.98***	0.97***	0.93***	1.00	1.00***	1.00***	0.99***
Hourly Earnings	1.39	1.41	1.62	2.05	0.99	0.98***	0.94***	0.90***	0.99	0.99	0.96***	0.95***
PPI (fin. goods)	3.74	3.91	4.20	4.65	1.00	0.99*	0.97***	0.93***	0.99	0.99	0.98**	0.96***
PCE prices	1.10	1.27	1.50	1.87	1.00	1.00	0.97***	0.93***	0.98**	0.98**	0.97	0.94*
Housing Starts	0.04	0.05	0.12	0.19	1.00	1.01	1.01	0.99	0.99	0.99**	0.97***	0.97**
S&P 500	23.08	23.45	25.27	28.16	0.99**	0.99***	0.96***	0.92***	1.00	1.00*	0.97***	0.97***
USD / GBP FX rate	15.40	16.03	16.32	17.39	1.01**	1.00	0.99***	0.98***	1.02	1.00	0.99	0.98
5-Year yield	0.14	0.31	0.66	0.86	1.00	1.01	1.02*	0.98	0.99	0.99	0.99**	0.96***
10-Year yield	0.13	0.29	0.64	0.81	1.00	1.00	1.01	0.98*	0.99	0.99	0.99**	0.95***
Baa spread	0.14	0.32	0.77	1.06	1.00	1.00	0.97*	0.90***	1.01	0.99	0.97	0.94***

Note: Comparison of "SVO" (baseline, in denominator of relative comparisons) against "SV-t(5)" and "SV-OutMiss." Values below one indicate improvement over baseline. Evaluation window from 1985:M01 through 2017:M12. Significance assessed by Diebold-Mariano test using Newey-West standard errors with $h + 1$ lags. Due to the close behavior of some of the models compared, and rounding of the report values, a few comparisons show a significant relative CRPS (alternative models) of 1.00. These cases arise from persistent differences in performance that are, however, too small to be relevant after rounding.