Addressing COVID-19 Outliers in BVARs with Stochastic Volatility*

Andrea Carriero
Queen Mary University of London
a.carriero@qmul.ac.uk

Todd E. Clark†
Federal Reserve Bank of Cleveland
todd.clark@clev.frb.org

Massimiliano Marcellino
Bocconi University, IGIER, and CEPR
massimiliano.marcellino@unibocconi.it

Elmar Mertens
Deutsche Bundesbank
elmar.mertens@bundesbank.de

This draft: May 31, 2021

Abstract

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 outlier-adjusted stochastic volatility (SV) models for VAR residuals that combine transitory and persistent changes in volatility. Evaluating forecast performance over the last few decades in quasi-real time, we find that outlier-augmented SV schemes do at least as well as a conventional SV model and outperform standard homoskedastic VARs. Our best-performing model features stochastic volatility, fat tails, and an occasional outlier state. Point forecasts made in 2020 from heteroskedastic VARs are much less sensitive to outliers in the data, and the outlier-adjusted SV models generate more reasonable gauges of forecast uncertainty than a standard SV model. In addition, we consider estimation and forecasting from a VAR with conventional SV that treats outliers in individual variables as missing data. In historical forecast comparisons, this alternative missing-data approach performs on par with our outlier-adjusted SV specifications. Over the limited sample of available data since the onset of the pandemic, treating individual outliers as missing data also generates particularly well-performing forecasts.

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

*We gratefully acknowledge helpful suggestions and comments received from Frank Schorfheide, and seminar participants at the Federal Reserve Bank of New York, University of Pennsylvania, 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. 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. Additional results are provided in a supplementary online appendix.

†Corresponding author.

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] Since then, real income has continued to fluctuate strongly, recording further record rates of increase and decline in early 2021.[3] 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 can 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),

[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 April 2021 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 then second-highest drop in our data (the largest drop in real income, by about 5 percent, that occurred in January 2013).

[3] In February and March of 2021, real income registered its strongest relative swings recorded so far, falling by about 7.5 percent in February and rising by about 18.5 percent in March, after having risen by about 9.5 percent in January 2021.

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 forecast 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.[4] 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.[5] However, some work in the SV literature has extended the conventional SV model to feature fat-tailed, instead of normal, errors, as in Jacquier, Polson, and Rossi (2004), henceforth denoted "SV-t."[6] As we show, these specifications with fat tails accommodate frequent, transitory changes in volatility. 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 and infrequent increases in volatility, in which case it may be desirable to reduce their influence on model

[4]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.

[5]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.

[6]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.

estimates and forecast distribution. The outlier-augmented SV process used by Stock and Watson (2016) with unobserved component models of inflation, henceforth denoted "SVO," accommodates such transient, sharp, increases in volatility.

In this paper, we introduce a novel combination of (1) an SV model with volatility outliers with (2) a VAR with fat-tailed errors. The outlier-adjusted volatility model, labeled SVO, 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. In its original form, first considered by Stock and Watson (2016), the model has Gaussian errors. We combine the outlier-adjusted SV process with VARs that have Gaussian (SVO) or t-distributed (SVO-t) errors, and also consider the case of t-distributed errors without volatility outliers (SV-t). In our baseline model, outlier states are variable-specific, and we also consider a common-outlier specification.

We demonstrate that SVO, SVO-t, and SV-t share the same latent state representation where residuals are written as the product of a normally distributed shock and a set of outlier states, but differ in the assumed densities for the outlier states. In particular, SVO puts more mass on outliers being large events that increase volatility by more than twofold, SV-t sees outliers as more moderately sized, and SVO-t is a combination of the two. While we are particularly interested in the performance of SVO and SVO-t during the COVID-19 episode, we also study its versatility outside the pandemic.

Our SVO-t model extends and nests the SVO approach of Stock and Watson (2016), developed in the context of an unobserved component model with normal errors, and the fat-tail SV model of Jacquier, Polson, and Rossi (2004). Specifically, we consider BVAR-SV models — without and with fat tails — and show that SVO(-t) effectively filters the outliers associated with the unprecedented, temporary volatility induced by the COVID-19 pandemic. In addition, SVO(-t) also detects pre-COVID outliers in macroeconomic and financial time series, whose existence had been noted by, among others, Stock and Watson (2002). Furthermore, volatility estimates resulting from our SVO(-t) model are generally less persistent than those obtained with the standard SV specification, as part of the volatility is attributed to non-persistent volatility outliers.

Conventional BVAR-SV procedures can easily be extended to include outlier state estimation via the SVO(-t) approach. Specifically, we show that the standard MCMC algorithm used for estimation of BVAR-SV models can still be used, but with the addition of two extra steps. First, realized outlier states need to be drawn from their posterior, conditional on draws for each variable's outlier probability. Second, the outlier probability for each variable is drawn from a (conditional posterior) distribution conditional on the draws of the time series of outlier states.

Empirically, we consider the effects of adding the SVO(-t) 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.[7]

The COVID-19 pandemic visibly affected the US economy starting in March 2020. We confirm the findings of Lenza and Primiceri (2020) and Schorfheide and Song (2020) that forecasts generated since then from homoskedastic BVARs are often distorted; for example, the recent outliers cause the forecast paths of some variables to become extreme by historical standards. 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(-t) model explicitly allows for one-off spikes in volatility, resulting in estimates

[7]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.

of forecast uncertainty that are still elevated but, in our subjective assessment, appear less extreme and more reasonable. So, in our assessment, the SVO and SVO-t specifications offer an effective approach for managing infrequent outliers with BVARs used for forecasting.

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.[8] 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(-t) specifications. But estimates of forecast uncertainty remain unrealistically stuck at pre-COVID levels.

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.[9] 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

[8]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.

[9]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).

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(-t) (or SV-t), though not 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 and SVO-t specifications 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 (e.g., Bańbura, Giannone, and Reichlin (2010), Carriero, Clark, and Marcellino (2019), and Koop (2013)). Considering 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. In historical forecast accuracy, the best-performing model is SVO-t, which features stochastic volatility, fat tails, and the outlier state treatment. However, the alternative approach of treating outliers as missing data in an otherwise conventional VAR with SV performs about as well in historical forecasting and performs best in the short sample following the pandemic's outbreak.

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 SVO-t, or the alternative of treating extreme observations as missing, prevents the width of predictive densities from blowing up as they would in the SV case. Importantly, the added value of SVO and SVO-t also holds up over a longer sample outside the recent COVID-19 episode.[10]

The remainder of this paper proceeds as follows: Section 2 introduces the SVO and SVO-t models 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 and early 2021. 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 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.[11] Heteroskedasticity in the VAR residuals has important effects on the estimation of Π, in

[10] 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.

[11] 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).

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 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 $y_{t+h|t}$.[12] 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

[12] 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.

distance between a given data point and the time-series median.[13]

2.1 Model specification

We consider the following seven variants of the VAR model (1). The first five differ in the specified process for the residuals v_t, whereas the last two variants either treat pre-specified outliers as missing data or add dummy variables for the pandemic period.

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:

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

3) SVO-t: The SVO-t model is intended to capture two kinds of outliers that are both modeled as transitory changes in volatility: The first kind captures rare jumps in volatility. The second kind occurs more often, but is less extreme in impact (consistent with draws from the tails of a fat-tailed distribution). Each kind of outliers enters the model in form of a diagonal matrix of scale factors, denoted O_t and Q_t, with diagonal elements $o_{j,t}$ and $q_{j,t}$, respectively, that are mutually *iid* over all j and t.

[13] 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.

The first kind of outliers, $o_{j,t}$, has a two-part distribution that distinguishes between regular observations and outliers. When variable j has a regular observation in period t we have $o_{j,t} = 1$, while for an outlier it is $o_{j,t} > 1$. Outliers in variable j occur with probability p_j and the distribution for $o_{j,t}$ is:

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

for $j = 1, \ldots, N$ and where $U(2, 20)$ denotes a uniform distribution with support between 2 and 20.

The second, less extreme, type of outliers in the SVO-t model is equivalent to having t-distributed VAR residuals (conditional on Λ_t and O_t). Following Jacquier, Polson, and Rossi (2004), we let the squares of the diagonal elements of Q_t, $q_{j,t}$, have inverse-gamma distributions:

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

The vector of VAR residuals in the SVO-t model is written as

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

with A^{-1} and $\Lambda_t^{0.5}$ specified as before. The j^{th} residual $q_{j,t} \cdot \varepsilon_{j,t}$ (adjusted for the rotation by A^{-1} and scaling by $\Lambda_t^{0.5}\,O_t$), has a student-t distribution with d_j degrees of freedom, since $\varepsilon_{j,t} \sim N(0, 1)$ and $d_j/q_{j,t} \sim \chi^2_{d_j}$. Since O_t, Q_t, and Λ_t are diagonal, they commute, and the time-varying variance-covariance matrix of the VAR residuals can conveniently be expressed as $\Sigma_t = A^{-1}\,O_t\,Q_t\;\Lambda_t\,Q_t'\,O_t'\;(A^{-1})'$.

We place a beta prior on the outlier probability p that corresponds to 10 years' worth of prior data. For the t-component of the SVO-t model, we follow Jacquier, Polson, and Rossi (2004) and estimate the degrees of freedom d_j for each variable using a uniform discrete prior with a range of 3 to 40.

4) SVO: We also consider a simplified version of the SVO-t model, denoted SVO, where $Q_t = I$ so that the VAR residuals are Gaussian (conditional on O_t and Λ_t). In this case, the time-varying variance-covariance matrix of the VAR residuals is given by $\Sigma_t = A^{-1} O_t \Lambda_t O_t' (A^{-1})'$. The SVO model is similar to the treatment of volatility outliers by Stock and Watson (2016) in an unobserved component model of inflation.[14] As in Stock and Watson (2016), we place a beta prior on the outlier probability p_j so that the prior mean implies an outlier frequency of once every 4 years in monthly data (and precision consistent with 10 years' worth of prior observations). As discussed further below, the prior mean of $p_j = 1/(4 \cdot 12)$ implies about the same variance of $o_{j,t}$ in the SVO model as do our settings for p_j and d_j in the SVO-t model for the combined outlier states $o_{j,t} \cdot q_{j,t}$.

With the COVID-19 pandemic inducing extreme volatility in a number of variables, some may view it as plausible that the outlier is common to all variables, rather than independent across variables as in the SVO specification. Some other work, such as Lenza and Primiceri (2020), has developed models in which the pandemic induces a common shift in volatility in an otherwise homoskedastic VAR. Accordingly, in a robustness check, we also consider a specification in which the outlier state is common to all variables, in which case the time-varying variance-covariance matrix of the VAR residuals is given by $\Sigma_t = \bar{o}_t^2 A^{-1} \Lambda_t (A^{-1})'$, where $\bar{o}_t$ denotes a scalar outlier state.

5) SV-t: The SV model with t-distributed errors, SV-t, is a simplified version of the SVO-t model where $O_t = I$, so that the time-varying variance-covariance matrix of the VAR residuals is given by $\Sigma_t = A^{-1} Q_t \Lambda_t Q_t' (A^{-1})'$. 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 estimated as in Jacquier, Polson, and Rossi

[14] In an application to inflation data, Stock and Watson (2016) use a $U(2, 10)$ distribution for $o_{j,t} > 1$.

(2004), using a uniform discrete prior with a range of 3 to 40.

Figure 2 illustrates the differences in densities implied for $o_{j,t}$, $q_{j,t}$, and $o_{j,t} \cdot q_{j,t}$ in, respectively, the SVO, SV-t and SVO-t models. The densities for the outlier states, $o_{j,t}$ and $q_{j,t}$, depend on the outlier probability p_j and the t-error degrees of freedom d_j. To create the figure, we set p_j equal to our choices of prior mean, and d_j equal to our estimated degrees of freedom in each model, which imply roughly equal variances $\mathrm{Var}\,(o_{j,t})$, $\mathrm{Var}\,(q_{j,t})$, and $\mathrm{Var}\,(o_{j,t} \cdot q_{j,t})$ in the SVO, SV-t and SVO model, respectively. [15] The density for the outlier states peaks at (SVO) or near (SVO-t, 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. SVO-t is in between, with more probability to values higher than 6 than SV-t but less than SVO. Also, while the outlier states in the SVO case cannot take values below 1, the SV-t and SVO-t cases assign some mass also to values below 1. Overall, SVO is more geared than SV-t toward generating sizable outliers at a variable-specific rate of occurrence p_j that is directly governed by an explicit prior, and SVO-t adds to that the flexibility of a fat-tailed error distribution.

6) 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

[15] For the SVO model, with $p_j = 1/(4 \cdot 12)$ we have $\mathrm{Var}\,(o_{j,t}) \approx 1.54$. In the SV-t case with $d_j = 5$ we get $\mathrm{Var}\,(q_{j,t}) \approx 1.67$ and in the SVO-t case, with $p_j = 1/120$ and $d_j = 9$ we obtain $\mathrm{Var}\,(o_{j,t} \cdot q_{j,t}) \approx 1.56$. The variances can be computed from

$$\mathrm{Var}\,(o_{j,t}) = (1 - p_j) + p_j \cdot \frac{(20-2)^2}{12}\,, \qquad \mathrm{Var}\,(q_{j,t}) = \frac{d_j}{d_j - 2}\,,$$

$$\text{and} \quad \mathrm{Var}\,(o_{j,t} \cdot q_{j,t}) = (1 - p_j) + p_j \cdot \frac{(20-2)^2}{12} \cdot \frac{d_j}{d_j - 2}\,.$$

For the SV-t case, we obtain similar estimates for our paper with $d_j = 6$, which implies $\mathrm{Var}\,(q_{j,t}) = 1.5$.

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 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.

7) SV-Dummy: As a final alternative, we also consider a simple BVAR-SV model with dummy variables included for each month of the pandemic period.

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, but as corrected in Carriero, et al. (2021a). 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 (corrected as in Carriero, et al. (2021a)), 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.[16]

For the SVO-t variant, the following extra steps are added to the original BVAR-SV setup: Realized outlier states $o_{j,t}$ and $q_{j,t}$ need to be drawn from their posteriors. The step for $o_{j,t}$ conditions on draws for the outlier probability p_j and proceeds analogously to the sampling of the mixture states needed with the Kim, Shephard, and Chib (1998) approach to the stochastic volatility states $\log \lambda_t$. The step for $q_{j,t}$ takes a draw from an inverse Gamma distribution. A further additional step draws the outlier probability p_j for each variable from a (conditional posterior) beta distribution conditional on the draws of the time series of outlier states. The algorithms for SVO and SV-t are simplified versions of that for SVO-t.[17]

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-t (or SVO or SV-t) over standard SV is small, since this model adds 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

[16] 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.

[17] The ordering of steps in our MCMC sampler reflects the recommendations of Del Negro and Primiceri (2015) as implemented also by Cúrdia, Del Negro, and Greenwald (2014) (for SV-t) and Stock and Watson (2016) (for SVO). Specifically, the t-error states, $q_{j,t}$, are sampled before the SV mixture states of Kim, Shephard, and Chib (1998), while draws from $o_{j,t}$ condition on those mixture states so that $o_{j,t}$ and p_j are sampled after the SV steps known from Kim, Shephard, and Chib.

[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.

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 2021:M3, taken from the April 2021 vintage of the FRED-MD database maintained by the Federal Reserve Bank of St. Louis. The variables and their transformation to logs or log-differences are listed in Table 1. To avoid issues related to the effective lower bound (ELB) on nominal interest rates, the data set includes only longer-term interest rates and omits a policy rate measure, like the federal funds rate, which was constrained by the ELB from late 2008 to 2016, and then again starting in March 2020.[19]

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 since the pandemic's outbreak, but without registering explicit outliers according to this metric. In some cases, outliers may be attributed to unusual

[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 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. (2021b)), 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 (2021).

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 and BVAR-SVO-t models 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, 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 April 2021 vintage of FRED-MD; we abstract from issues related to real-time data collection. The forecast horizons considered extend from 1 to 24 months. We evaluate point and density forecasts based on root-mean-squared errors (RMSE) and continuous ranked probability scores (CRPS), respectively, as described in, among others, Clark and Ravazzolo (2015) and Krüger, et al. (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 a homoskedastic BVAR

[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.

[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.

and BVARs with SV and SVO-t specifications, taking the SV model as the benchmark. 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 1 to 4 percent in the case of employment growth and as much as 25 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-t 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-t typically performs as well as, and at times even better than, SV. Point forecasts generated by the SVO-t model over the post-war period (and pre-COVID) are generally on par with those from the SV model, with RMSE ratios in some cases a little below or above 1 but often very close to 1. With density forecast accuracy as gauged by the CRPS, at shorter horizons the SVO-t specification performs very similarly to the SV baseline, with CRPS ratios very close to 1. At longer horizons, particularly at 24 months, the SVO-t model often yields modestly most accurate density forecasts, for variables including real income, consumption, industrial production, employment, hours, and stock returns. The SVO-t gains are largest for real personal income, the variable most prone to outliers. Overall, the evidence suggests that consistent use of SVO-t 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-t against versions of the model that strip out the t-distributed component (SVO) or the Stock-Watson outlier state (SV-t) as well as the SV-OutMiss approach, which treats pre-specified outliers as missing data as described in Section 2. Note that these comparisons take SVO-t as the baseline, so that a ratio less (more) than 1 means the alternative model is more (less) accurate than the baseline. By and large, point forecasts from these alternatives are quite similar in accuracy to those from the SVO-t specification. Differences in relative RMSE are never more than 4 percent and typically just 1 percent or less. In density accuracy, the SV-t and SV-OutMiss models are similar to our preferred SVO-t specification, with SV-t particularly so. At horizons of 12 months or less, the SV-OutMiss specification yields density accuracy very similar to the SVO-t baseline; at 24 months, SV-OutMiss forecasts are sometimes modestly less accurate (e.g., industrial production and hours) and sometimes a little more accurate (e.g., bond yields). The most noticeable differences in CRPS accuracy occur with the SVO model. Although at shorter horizons SVO accuracy is quite similar to SVO-t accuracy, at longer horizons SVO-t provides the more accurate forecasts, often by a statistically significant margin, reaching 9 to 10 percent for consumption, industrial production, employment, and hourly earnings.

The supplementary online appendix provides results of another robustness check, of treating the outlier state as common in the SVO specification. Making the outlier common seems to have no advantages and in some cases makes forecasts slightly less accurate compared to the SVO specification that models outliers as independent across variables. In point forecast accuracy, the common outlier specification generally matches the SVO results, with slightly reduced accuracy for some variables at some horizons. As measured by the CRPS, density forecasts from the common outlier specification are essentially the same in accuracy as those from the SVO model at horizons of 12 months or less but consistently less accurate at the 24 months horizon. The common-outlier specification registers virtually no outliers prior to the COVID-19 pandemic. Instead, the common-outlier specification sees outliers only in the early stages of the pandemic period, from March through June 2020, when a good number of variables experienced enormous realizations at the same time, but none in late 2020 or early 2021. Imposing the same outlier on all variables during COVID-19 leads to some marked differences in width of predictive densities compared to the SVO(-t) models that feature variable-specific outliers, but fairly identical performance in point forecasts over the pandemic period.

5 Outlier estimates in 2020 and before

As described in Section 2, the SVO-t approach extends the baseline SV model by adding latent outlier states $o_{j,t}$ and $q_{j,t}$ for each variable $j = 1, \ldots, N$, with the former uniformly distributed and squares of the latter having an inverse Gamma distribution. The outlier states enrich the dynamics of the time-varying variance-covariance matrix, Σ_t, so that volatility can change due to transitory changes in $o_{j,t}$ and $q_{j,t}$, as well as the persistent variations induced through the log-SV terms $\log \lambda_t$. The SVO model adds just the state $o_{j,t}$ to an SV model, whereas the SV-t specification adds just the state $q_{j,t}$. In each case, the additional latent states serve 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-t, SVO, and SV-t. In each case, the model estimates permit the computation of the posterior probabilities of an outlier of a selected magnitude. For the SVO-t model, we examine the probability of a given value of an outlier in the reduced form innovation that reflects both O_t and Q_t and the connections across variables captured in A^{-1}. Specifically, we compute reduced-form outlier scales from the ratios between diagonal elements of $\Sigma_t = A^{-1} O_t Q_t \Lambda_t Q_t' O_t A^{-T}$ and $\tilde{\Sigma}_t = A^{-1} \Lambda_t A^{-T}$; denoting the j-th diagonal elements by $\tilde{\sigma}_t(j)^2$ and $\sigma_t(j)^2$, the outlier scale for the j-th variable is $\sigma_t(j)/\tilde{\sigma}_t(j)$. To simulate the posterior distribution for the reduced-form outlier states, these computations are performed for every MCMC draw. (For SVO and SV-t, corresponding computations are performed using only O_t and Q_t, respectively.) For each variable, we then examine the probability of a given value of the combined outlier term for the SVO-t model. For ease of comparison, we focus on three regions for possible

[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).

realizations of the outlier states: below 2, between 2 and 5, and above 5, corresponding to the cases of small (or no) 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 the outlier estimate to have fallen in one of the three regions at a particular point in time. The three regions cover the possible support of the outlier, and posterior probabilities of a realized outlier to fall in the three regions sum to one.[24] 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. To help distinguish estimates for recent years, the top panel of each figure provides estimates for the full sample, and the bottom panel reports probabilities for just the 2017:M1-2021:M3 period.

Echoing our discussion of each model's properties in Section 2, these results show that, when SVO and SV-t are compared, 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). For example, in a few months of 2020, both SV-t and SVO yield high probabilities of outliers in payroll growth, but SVO indicates the outlier to be larger in magnitude than does SV-t. As may be expected, our preferred SVO-t specification captures aspects of both SV-t-type outliers and SVO-type outliers. With SVO-t, the probability of a small outlier is a little lower than in the SV-t case, with some probability mass shifted to a large outlier. Similarly, with SVO-t, the probability of a large outlier is modestly lower than in the SVO results, with some probability mass shifted to a small outlier. The results for payroll growth in 2020 provide an example. The SV-t estimates yield a high probability of small outliers in 2020, whereas the SVO estimates put a high probability on large outliers.

[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 $o_{j,t}$ and $q_{j,t}$ 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. In the SVO-t case, for the combined outlier state, $o_{j,t} \cdot q_{j,t}$, the prior mass blends features of both SVO and SV-t. In particular, in the upper tail, for realizations of the combined outlier state above 5, the SVO-t prior has more mass than SV-t but less than SVO.

[24] Possible realizations of $q_{j,t}$ beyond the maximal value of $o_{j,t}$ attract only negligible density as shown in Figure 2.

The SVO-t estimates yield a mixture of small and large outliers last year.

For selected variables, Figure 8 also reports prior and posterior probability densities of the outlier probability parameter p_j in the SVO-t and SVO models, which describes the model's unconditional probability of seeing an outlier state $o_{j,t}$ value of 2 or more (at any given point in time). In the SVO case, 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.[25] For payroll growth, 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. A comparison between the SVO and SVO-t estimates (shown in the bottom half of the figure) reveals that, with the additional outlier state added in the SVO-t model, the posterior distributions of the outlier probability parameter p_j tend to shift a little to the left and become less dispersed — more evidence that, with the small outlier state also in the model, the probability of a large outlier is reduced. Of course, as the figure indicates, in the SVO-t specification, the prior distributions are also shifted to the left and less dispersed than in the SVO model; as noted above, in SVO-t, we deliberately use a prior that implies large outliers to be less frequent.

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-t are very small, while those in CRPS are sometimes larger. The outlier states in SVO-t (as well as SVO and SV-t) allow for spikes in volatility

[25] 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.

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.[26]

For each variable we report estimates of time variation in the volatility of forecast errors generated by SV and SVO-t, as well as the persistent components of Σ_t imputed from SVO-t when the effects of the outlier states $o_{j,t}$ and $q_{j,t}$ are ignored. For this counterfactual, we compute $\tilde{\Sigma}_t = A^{-1}\,\Lambda_t\,(A^{-1})'$ based on the SVO-t estimates for Λ_t and A^{-1}.[27] 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-t model clearly differentiates between increases in uncertainty that are short- and longer-lived, which the SV model cannot do. Volatility estimates from the SV model, 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-t proves more nimble in accounting for the extreme data seen in the spring with a big jump in volatility in April as shown in Panel (b) of the figure. However, as revealed by comparison with Panel (d), 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). In contrast, the persistent component of volatility in the case of SVO-t is seen to have risen no more than 8-fold over the course of the year. That is, the SVO-t estimates yield a much smaller rise in the persistent component of volatility than do the estimates

[26] 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).

[27] The full variance-covariance matrix of forecast errors in the SVO-t model is instead given by

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

from the SV model. The SV-OutMiss model yields an even smaller increase in the persistent component of volatility (the only component of volatility in that model); the estimates from SV-OutMiss shown in Panel (c) have risen by less than 5 times their level at the beginning of the year.

The more moderate rise in estimates of the persistent volatility component obtained with the SVO-t specification yields noticeably narrower (and arguably less extreme) uncertainty bands around forecasts compared to the SV model. In contrast, forecasts that condition on knowledge of when outliers occurred, but otherwise ignore any further information from their realization (as in the SV-OutMiss case), lead to particularly narrow uncertainty bands.[28] 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.

6 Forecasts made in 2020-2021

As a reference for the pre-COVID situation, Figure 10 reports forecasts generated by the CONST and SV models in January 2020.[29] 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 in most cases. For example, as shown in Panel (d) of the figure, the CONST model saw the unemployment rate rise back to 4 percent over the course of 2021, consistent with a higher longer-run level of the unemployment rate, whereas the SV model predicted a modestly shallower path, with the unemployment rate 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 mea-

[28] 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.

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

sures 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 June generated from our alternative BVAR models.[30] As noted by Lenza and Primiceri (2020) and Schorfheide and Song (2020), forecasts generated by homoskedastic BVARs, like our CONST specification, can display extreme behavior since the spring of 2020.[31] For example, CONST forecasts for the unemployment rate made in April run to nearly 35 percent, with a 68 percent uncertainty band of -17 percent to 22 percent in April 2022; see Panel (b) of Figure 14.

In contrast, the reaction of point and density forecasts generated by the SV and SVO-t specifications to the incoming data in spring 2020 is much better behaved. Considering again the unemployment rate forecasts shown in Figure 14, point forecasts from the SV model rise to about 17 percent by the end of 2020 and then fall back to about 13 percent by the end of 2021.[32] Generally, across variables and forecast origins, point forecasts generated by SV and SVO-t are fairly close, as seen also in our comparison of forecast performance pre-COVID-19 in Table 2.[33]

However, stark differences emerge considering the uncertainty around forecasts made with and without outlier adjustments. In keeping with the volatility comparisons provided above, while the observations of 2020 widen the predictive densities of both SV and SVO-t forecasts, their impact is much greater for the former than the latter. As indicated in the top row of each of the figures, SVO-t generates much narrower bands than SV in response to the particularly large swings in incoming data seen in March and April 2020. Moreover, the

[30]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.

[31]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.

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

[33]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.

SVO-t bands also remain narrower for forecasts made in subsequent months, such as June 2020.

The middle-row panels of Figures 11–14 compare our preferred SVO-t results to those for the more restrictive SVO and SV-t specifications. As expected, while the point forecasts of these specifications are difficult to distinguish, bigger differences are evident in the widths of the predictive densities. Across the variables and forecast origins shown, the predictive densities are generally the narrowest with the SVO-t forecasts. The SVO model generally yields wider densities, although in most cases the differences are less stark in June compared to March and April. Estimates are more varied with the SV-t model. In some cases (e.g., for payroll growth at the March 2020 forecast origin), the SV-t forecast intervals are very similar to the SVO-t estimates. But, in other cases, the SV-t intervals are wider than the SVO-t estimates; examples include real income and the unemployment rate in the April 2020 forecasts, when incoming data for these variables displayed particularly large jumps.

Critically, SVO-t, 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 bottom-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 (i) of Figure 12, payroll growth data for the months of March, April,

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

May, and June are treated as outliers in our application. The resulting 68 percent bands generated in June for annualized payroll growth in late 2021 range from -16 to 19 percent, whereas the corresponding SVO-t band is much wider, ranging from about -25 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-t during the onset of the COVID-19 recession.[35] While the April forecast generated by SV-OutMiss sees payroll growth turning positive in the fourth quarter of 2020 and hovering around 4 percent (annualized) for the remainder of the forecast horizon, SVO-t predicts a protracted slump until mid-2021, followed by a swifter pace of payrolls than predicted by SV-OutMiss.

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 June. Strikingly, introduction of

[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.

[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.

the outlier dummies to the BVAR with SV leads to point forecasts that are nearly identical to those obtained from SVO-t 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.

Figure 15 provides predictive densities for more recent forecast origins, ranging from September 2020 to March 2021, focusing on the SVO-t specification, the missing data approach, and the pandemic dummies approach. These latest forecasts display a now familiar pattern: Even almost a year after the onset of the COVID-19 pandemic impacted economic data, uncertainty bands from SVO-t remain noticeably wider than before the pandemic (evident in a comparison of Figure 10 to Figure 15). In most cases, forecast densities obtained from SV-OutMiss or the dummy approach, which both treat the timing of outliers as known, remain relatively tight. However, exceptions are evident in the unemployment rate forecasts provided in the bottom row, with the SV-OutMiss bands wider than those of SVO-t for forecasts made with data in September and December 2020. Although harder to discern in the wide scales of the charts necessitated by the extreme realizations of actual data, the point forecasts produced by the alternative methods tend to be broadly similar at longer forecast horizons, although more sizable differences can occur at shorter horizons.

To permit some assessment of the recent performance of alternative methods, Tables 6 and 7 provide RMSE and CRPS results for forecast origins from March 2020 through February 2021, taking the SV model as the baseline.[38] In light of the very small sample, we focus on short forecast horizons, of 1, 3, and 6 months, and we don't assess the statistical significance of any differences.[39] In these 2020-2021 results, consistent with historical results, the

[38] In the supplementary online appendix, we also consider forecast performance during the Great Recession of 2007-09 and its aftermath. As shown 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, SV-t and SV-OutMiss. A likely reason for this pattern is that few outliers are detected in this period, after properly accounting for volatility spikes.

[39] The April 2021 vintage of FRED-MD provides realized values through March 2021. Thus, to evaluate one-step ahead forecasts, the forecast evaluation ends with the forecast origin of February 2021. For longer forecast horizons, the evaluation window is shortened as needed.

accuracy of point forecasts are very similar for SV and SVO-t, except for the Baa spread and longer-term yields (for which SVO-t does better). Point forecasts from SV-dummy also perform typically similarly to those from SV except for longer-term yields and the Baa spread, while point forecasts from SV-OutMiss proved more accurate than SV for many variables, with gains in one-step ahead forecasts reaching up to 50 percent for the 5-year yield (and about 40 percent for hourly earnings and PCE inflation). By the CRPS measure of density accuracy, SVO-t is similar to SV for 1-month-ahead forecasts but becomes better for most variables for the 3- and 6-month horizons. SV-Dummy is again overall comparable to SVO-t, while SV-OutMiss is best, with quite some gains over SV for most variables at the 1-month horizon, and for all variables at the 3- and 6-month horizons.[40] On balance, the approach of treating pre-screened outliers as missing data in a BVAR with SV has worked relatively well for (near-term) forecast accuracy since early 2020. However, given the yet scarce number of realized data points since March 2020, the comparison is based only on relatively few non-overlapping forecast windows (even for near-term forecasts).

To visualize the sensitivity of these results to individual observations, Figure 16 shows absolute losses of one-step ahead forecasts generated from March 2020 onward for selected variables by the models covered by Tables 6 and 7.[41] For some variables, like capacity utilization or hourly earnings, the performance of SV-OutMiss accrues mainly in the early stage of the pandemic period by avoiding singularly large errors. In a few cases, like the 10-year yield or the Baa spread, SV-OutMiss gains have been a little more consistent, whereas in other cases, like PCE inflation or housing starts, relative losses have fluctuated quite unevenly over this short pandemic evaluation window.

[40]From unreported results, over the 2020-21 period SVO-t is similar to SVO and SV-t in terms of point forecasts, while for density forecasts SVO-t yields small improvements in accuracy.

[41]We plot absolute instead of squared errors to provide better visibility of large and small forecast misses. Corresponding plots for additional variables can be found in the supplementary online appendix.

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 sometimes 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.

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.[42] In contrast, SVO allows the model to fit sharp spikes in current volatility while adapting its uncertainty forecasts more moderately. Alternatively, dummy variables could be added to the standard VAR-SV model for every month of the pandemic. By soaking up all information contained in data since the onset of the pandemic, the dummy approach generates point forecasts comparable to our outlier adjusted SV models. As the dummy approach is conditioned on ex-ante knowledge that all COVID-19 related data points are highly unusual, its forecast densities are much tighter than those derived from our more agnostic outlier-adjusted SV models. 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

[42] 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.

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. Our preferred model, SVO-t, puts together the features of SVO and SV-t.

Of course, future data will be needed to assess which of the forecasts made in 2020 and 2021 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-t outperforms standard SV, in particular in terms of density forecasts and at longer horizons, while both display benefits over a constant-variance BVAR. In 2020-2021, point forecasts generated from SV and SVO-t are very similar. But as SVO-t projections filter out the effects from short-lived outliers on forecast uncertainty, predictive densities constructed with SVO-t 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, the SVO-t model (as well as SVO and SV-t) treats the occurrence of outliers as stochastic events, with unknown timing. As a result, forecast uncertainty generated from these outlier-adjusted SV approaches is less compressed than what is obtained from approaches that treat the occurrence of outliers as known. More broadly, treating outliers as random events makes SVO-t, SVO, and SV-t attractive for continued use over the yet-unknown course of economic developments related to the COVID-19 pandemic. That being said, an alternative approach would be to pre-screen the data to identify outliers in individual variables based on a simple measure of historical norms, and then treat these variable-specific outliers as missing observations in an otherwise conventional VAR with SV. This alternative missing-data approach performs about as well in historical forecasting and does particularly well over the short sample of available data over which near-term forecasts for 2020-2021 can so far be evaluated.

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://onlinelibrary.wiley.com/doi/abs/10.1002/for.957, 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] Bańbura, 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, Joshua Chan, Todd E. Clark, and Massimiliano Marcellino (2021a), "Corrigendum to: Large Bayesian Vector Autoregressions with Stochastic Volatility and Non-Conjugate Priors," *manuscript*.

[6] Carriero, Andrea, Todd E. Clark, and Massimiliano Marcellino (2019), "Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors," *Journal of Econometrics*, 212, 137–154, https://doi.org/10.1016/j.jeconom.2019.04.024.

[7] 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.

[8] ——— (2021b), "Forecasting with shadow-rate VARs," *mimeo*.

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

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

[11] 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.

[12] 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.

[13] 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.

[14] 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`.

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

[16] 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`.

[17] 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`.

[18] Del Negro, Marco, and Giorgio E. Primiceri (2015), "Time varying structural vector autoregressions and monetary policy: A corrigendum," *Review of Economic Studies*, 82, 1342–1345, `https://doi.org/10.1093/restud/rdv024`.

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

[20] 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`.

[21] 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`.

[22] 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.).

[23] 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.

[24] 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`.

[25] 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`.

[26] Johannsen, Benjamin K., and Elmar Mertens (2021), "A time series model of interest rates with the effective lower bound," *Journal of Money, Credit and Banking*, `https://doi.org/10.1111/jmcb.12771`.

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

[28] 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`.

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

[30] 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`.

[31] Koop, Gary (2003), *Bayesian Econometrics*: Wiley-Interscience, `https://pureportal.strath.ac.uk/en/publications/bayesian-econometrics`.

[32] 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`.

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

[34] 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`.

[35] 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`.

[36] 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`.

[37] 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`.

[38] 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`.

[39] ——— (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`.

[40] ——— (2016), "Core inflation and trend inflation," *The Review of Economics and Statistics*, 98, 770–784, `https://doi.org/10.1162/REST_a_00608`.

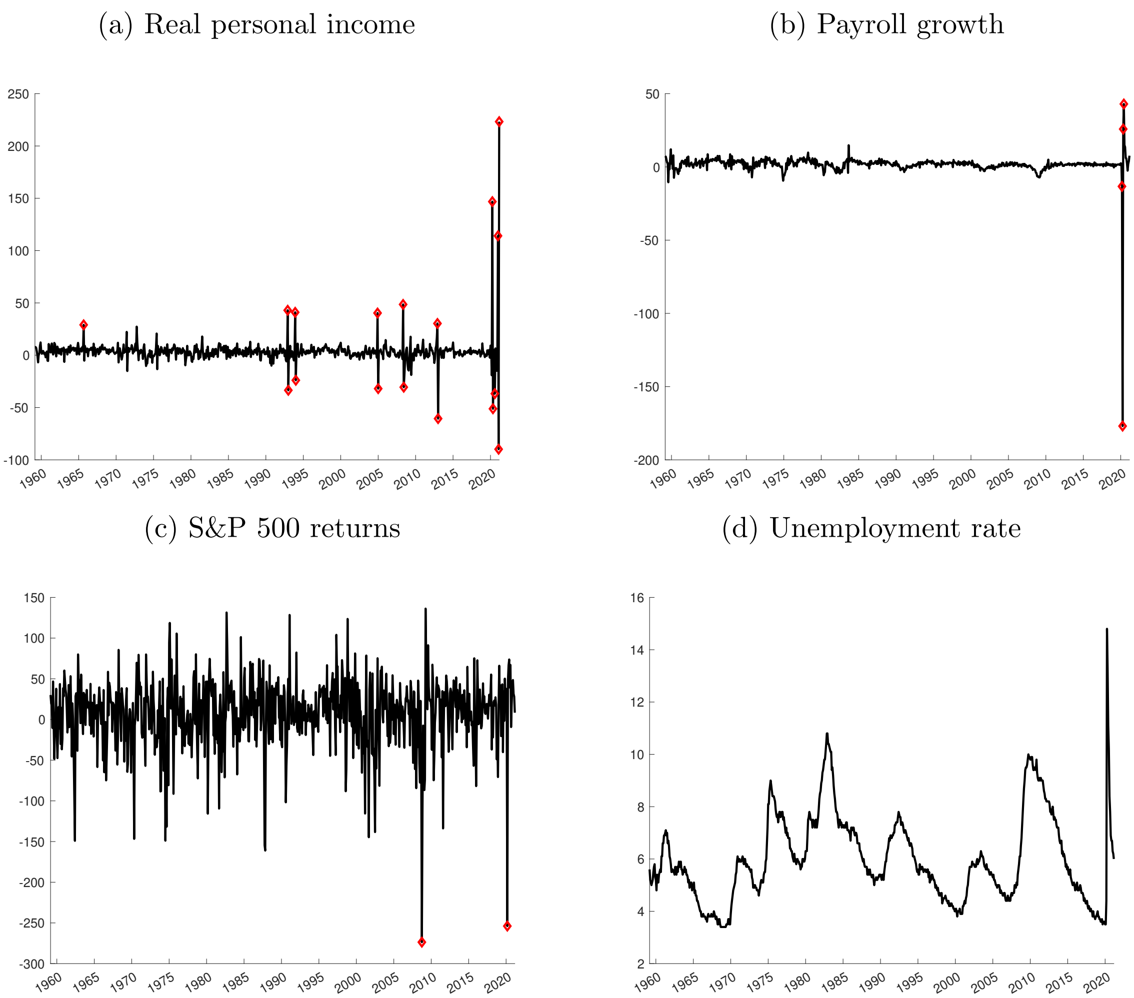
[41] West, Kenneth D. (1996), "Asymptotic inference about predictive ability," *Econometrica*, 64, 1067–1084, `https://doi.org/10.2307/2171956`.

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 2021-04 vintage of FRED-MD. Monthly observations from 1959:M03 to 2021:M03. 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

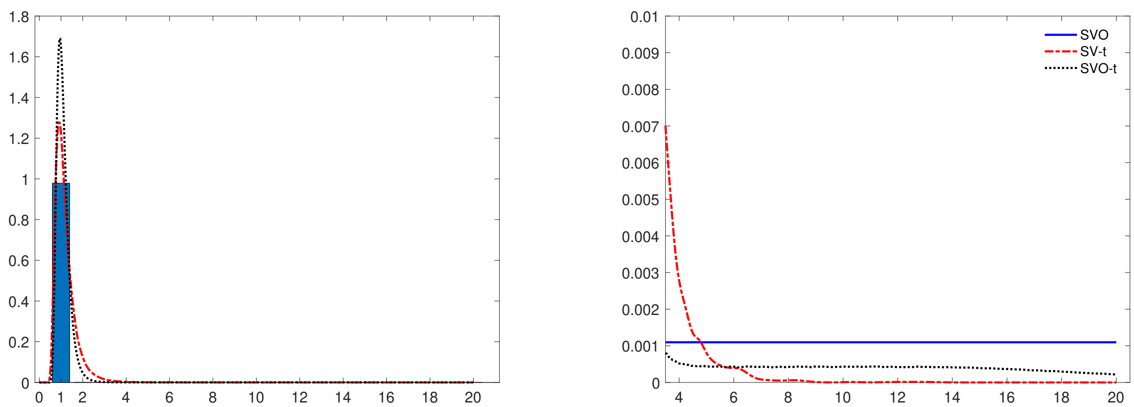
(d) Unemployment rate

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: Prior densities of outlier states in different models

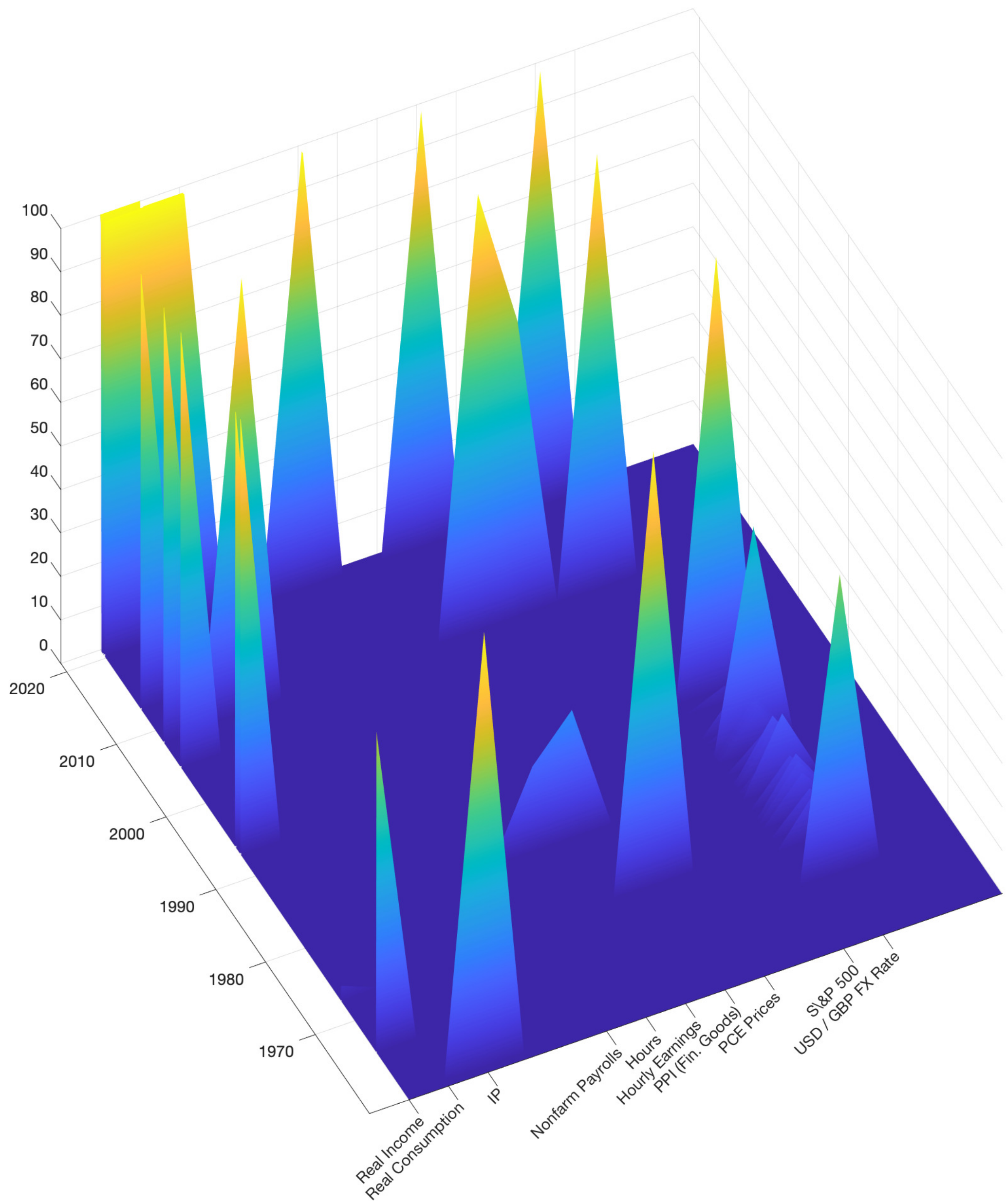
(a) Full figure

(b) Zoomed into right tail



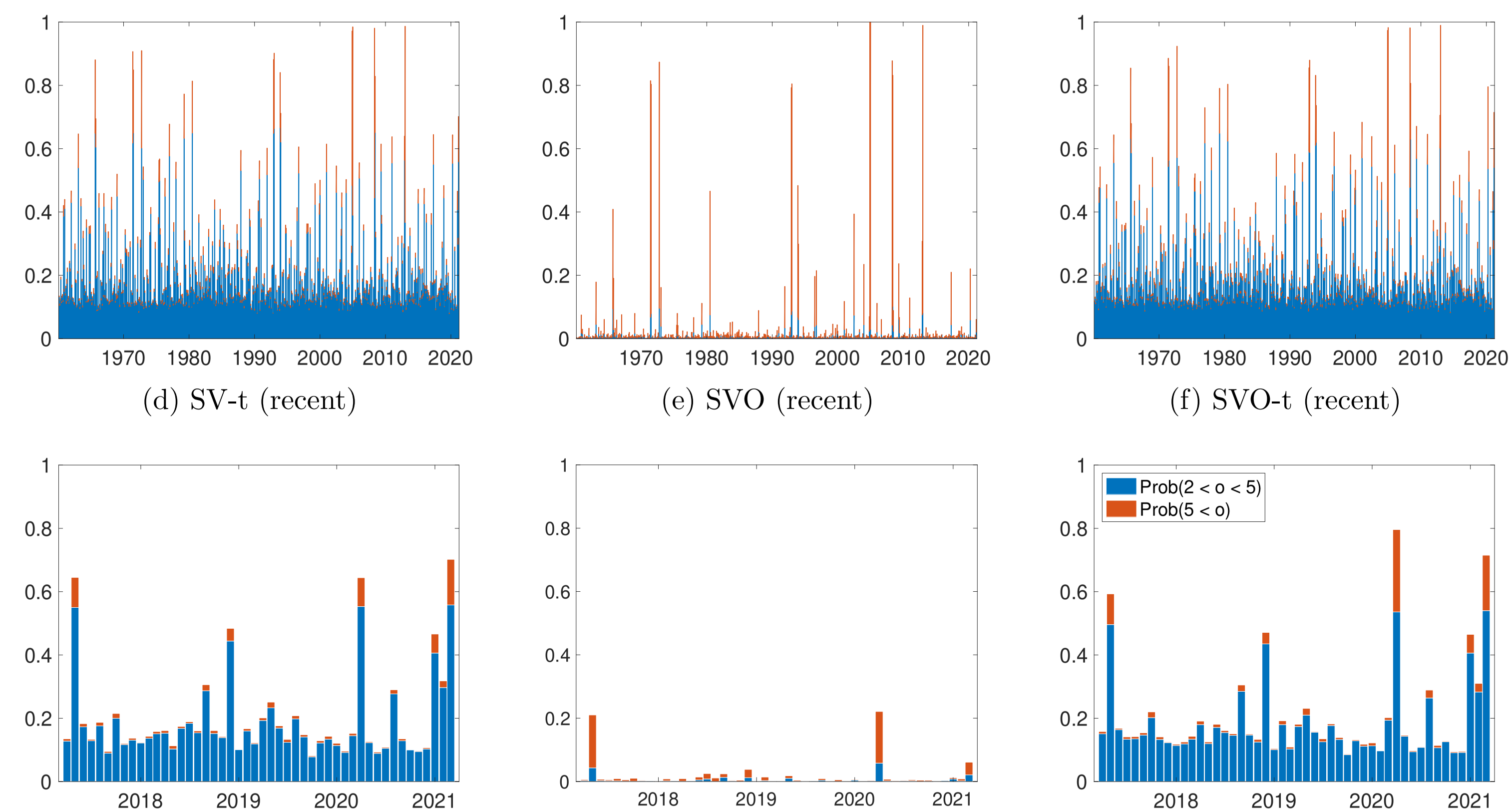
Note: Densities for the outlier state $o_{j,t}$ in the SVO model, $q_{j,t}$ in the SV-t model of Jacquier, et al. (2004), and $o_{j,t} \cdot q_{j,t}$ in the combined SVO-t model. The densities are calibrated to generate roughly the same variance of the outlier states. For the SVO model, outlier probability p_j has been set to correspond to one outlier every four years in monthly data, $p_j = 1/(4 \cdot 12)$. The degrees of freedom for the SV-t model have been set equal to five. For the SVO-t model, the outlier probability has been lowered to correspond to one outlier every ten years, and the degrees of freedom of the t-component have been set to 9.

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 inter-quartile 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

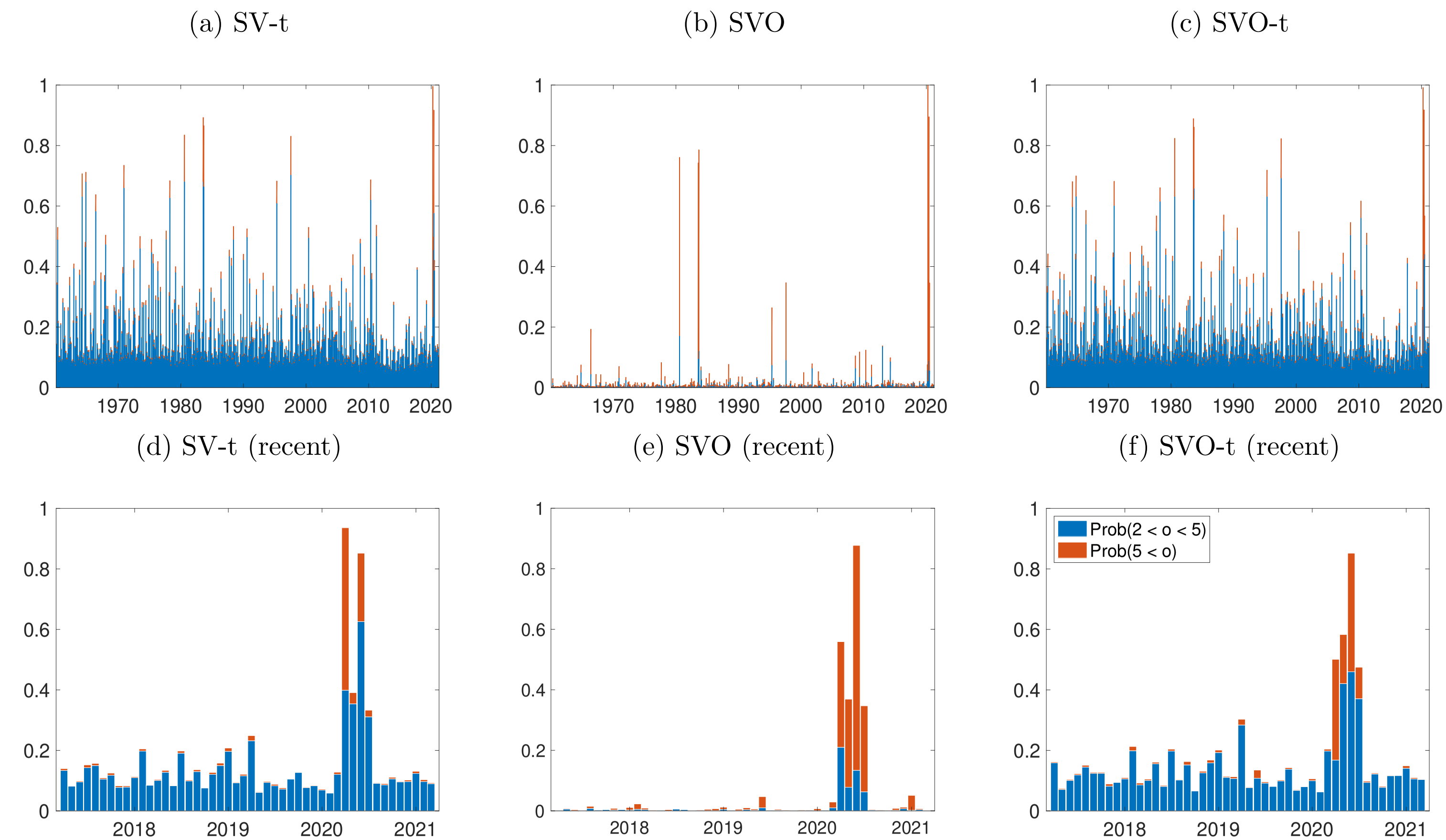


(a) SV-t (b) SVO (c) SVO-t

(d) SV-t (recent) (e) SVO (recent) (f) SVO-t (recent)

Note: Full-sample estimates per March 2021 of posterior probabilities for realizations of the reduced-form outlier states in different models. For the SVO-t model, we compute reduced-form outlier scales from the ratios between diagonal elements of $\Sigma_t = A^{-1} O_t Q_t \Lambda_t Q_t' O_t A^{-T}$ and $\tilde{\Sigma}_t = A^{-1} \Lambda_t A^{-T}$; denoting the j-th diagonal elements by $\tilde{\sigma}_t(j)^2$ and $\sigma_t(j)^2$, the outlier scale for the j-th variable is $\sigma_t(j)/\tilde{\sigma}_t(j)$. For SVO and SV-t, corresponding computations are performed using only O_t and Q_t, respectively. To simulate the posterior distribution for the reduced-form outlier states, these computations are performed for every MCMC draw. Each panel shows posterior probabilities for the outlier states 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).

Figure 5: Posteriors of outlier states for payroll growth

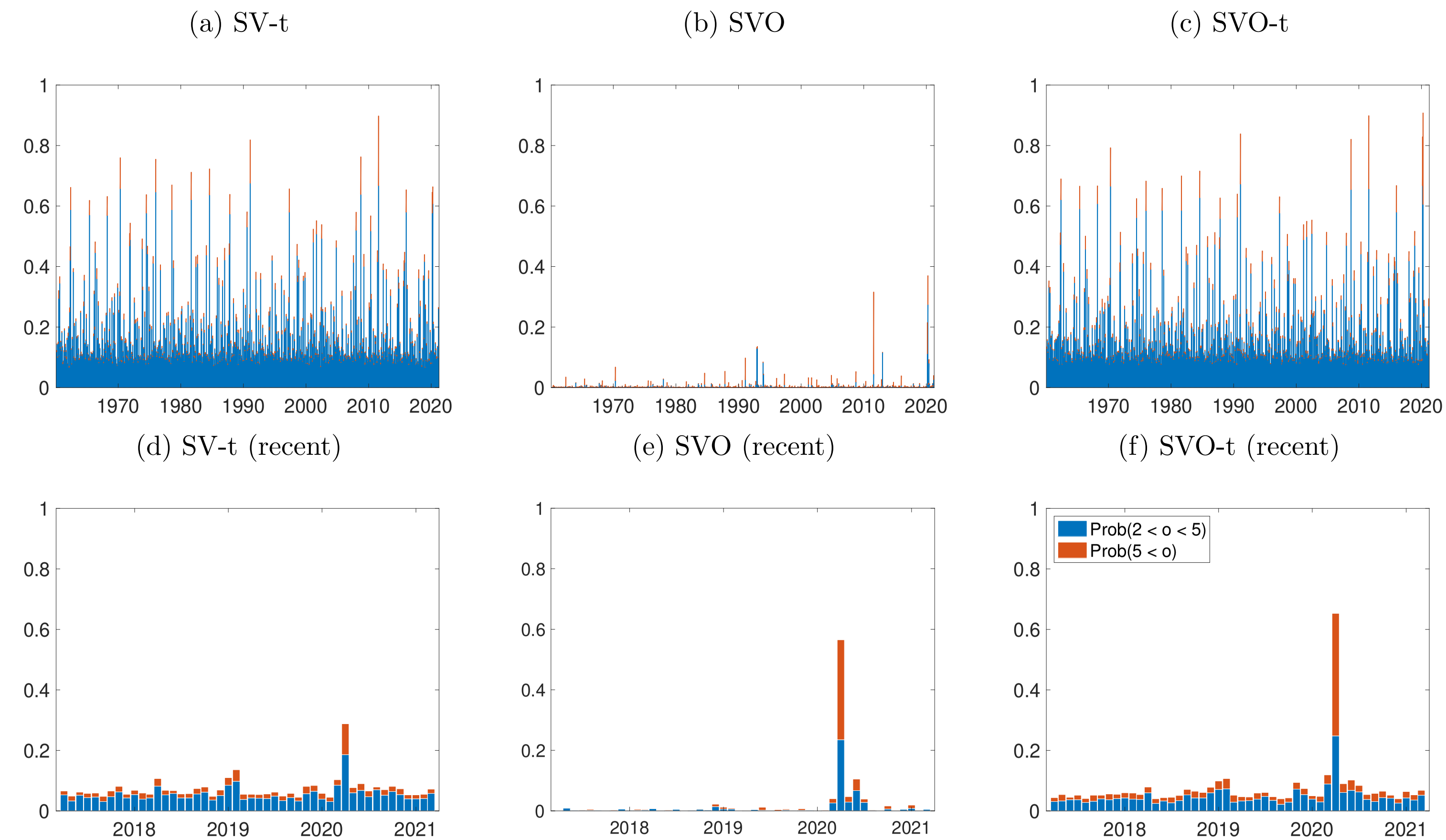


(a) SV-t (b) SVO (c) SVO-t

(d) SV-t (recent) (e) SVO (recent) (f) SVO-t (recent)

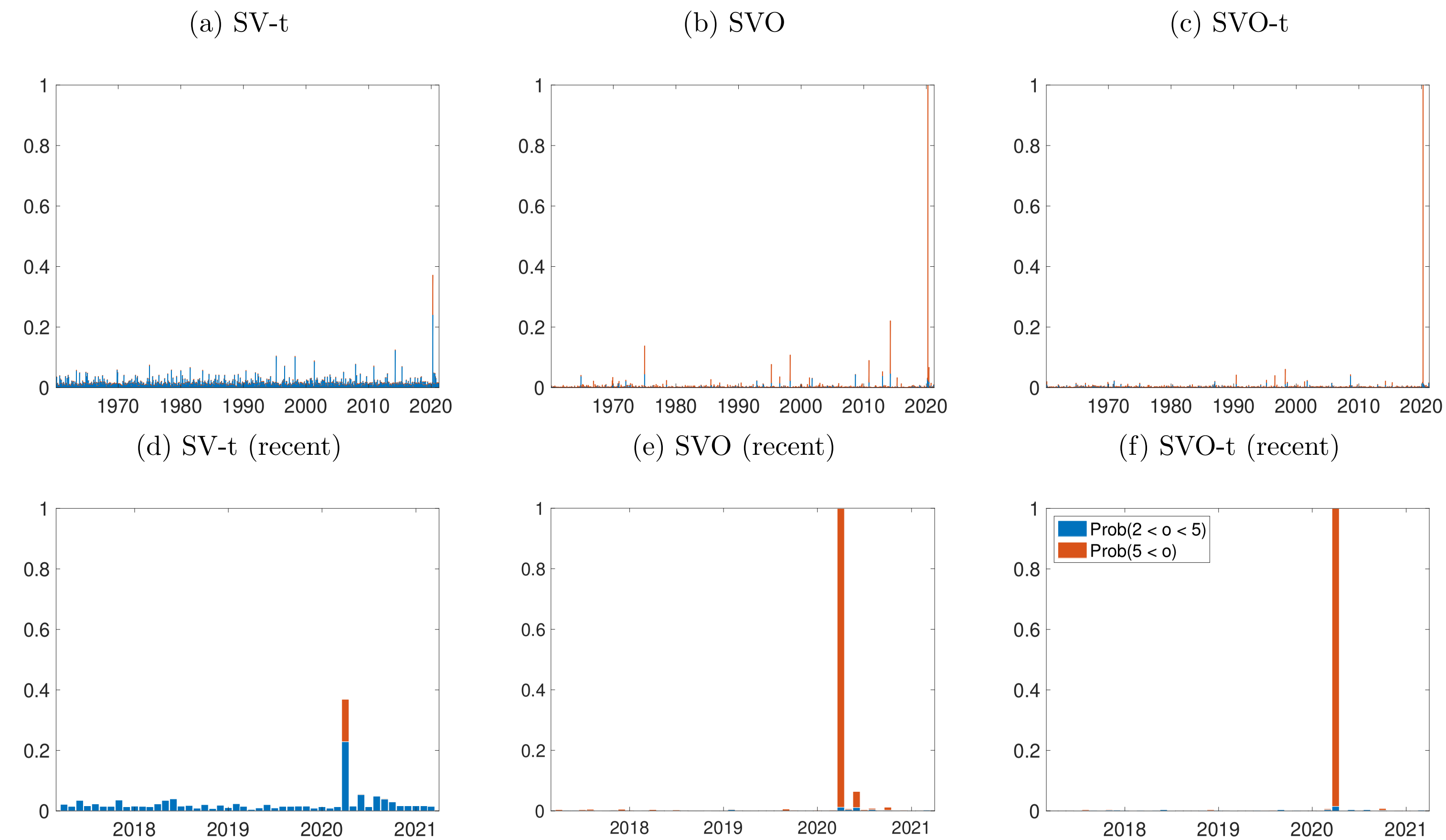
Note: Full-sample estimates per March 2021 of posterior probabilities for realizations of the reduced-form outlier states in different models. For the SVO-t model, we compute reduced-form outlier scales from the ratios between diagonal elements of $\Sigma_t = A^{-1} O_t Q_t \Lambda_t Q_t' O_t A^{-T}$ and $\tilde{\Sigma}_t = A^{-1} \Lambda_t A^{-T}$; denoting the j-th diagonal elements by $\tilde{\sigma}_t(j)^2$ and $\sigma_t(j)^2$, the outlier scale for the j-th variable is $\sigma_t(j)/\tilde{\sigma}_t(j)$. For SVO and SV-t, corresponding computations are performed using only O_t and Q_t, respectively. To simulate the posterior distribution for the reduced-form outlier states, these computations are performed for every MCMC draw. Each panel shows posterior probabilities for the outlier states 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).

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



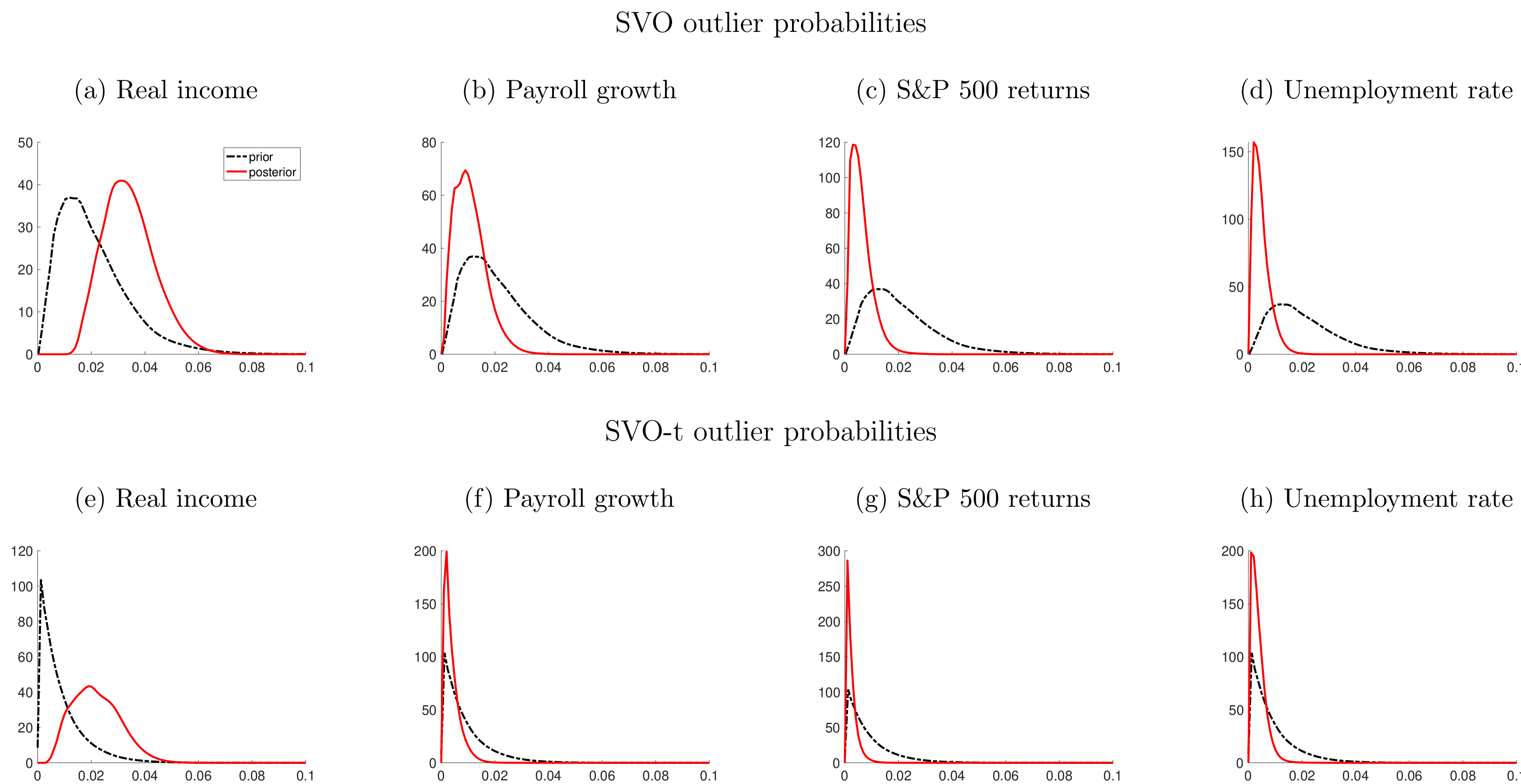
Note: Full-sample estimates per March 2021 of posterior probabilities for realizations of the reduced-form outlier states in different models. For the SVO-t model, we compute reduced-form outlier scales from the ratios between diagonal elements of $\Sigma_t = A^{-1} O_t Q_t \Lambda_t Q_t' O_t A^{-T}$ and $\tilde{\Sigma}_t = A^{-1} \Lambda_t A^{-T}$; denoting the j-th diagonal elements by $\tilde{\sigma}_t(j)^2$ and $\sigma_t(j)^2$, the outlier scale for the j-th variable is $\sigma_t(j)/\tilde{\sigma}_t(j)$. For SVO and SV-t, corresponding computations are performed using only O_t and Q_t, respectively. To simulate the posterior distribution for the reduced-form outlier states, these computations are performed for every MCMC draw. Each panel shows posterior probabilities for the outlier states 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).

Figure 7: Posteriors of outlier states for the unemployment rate



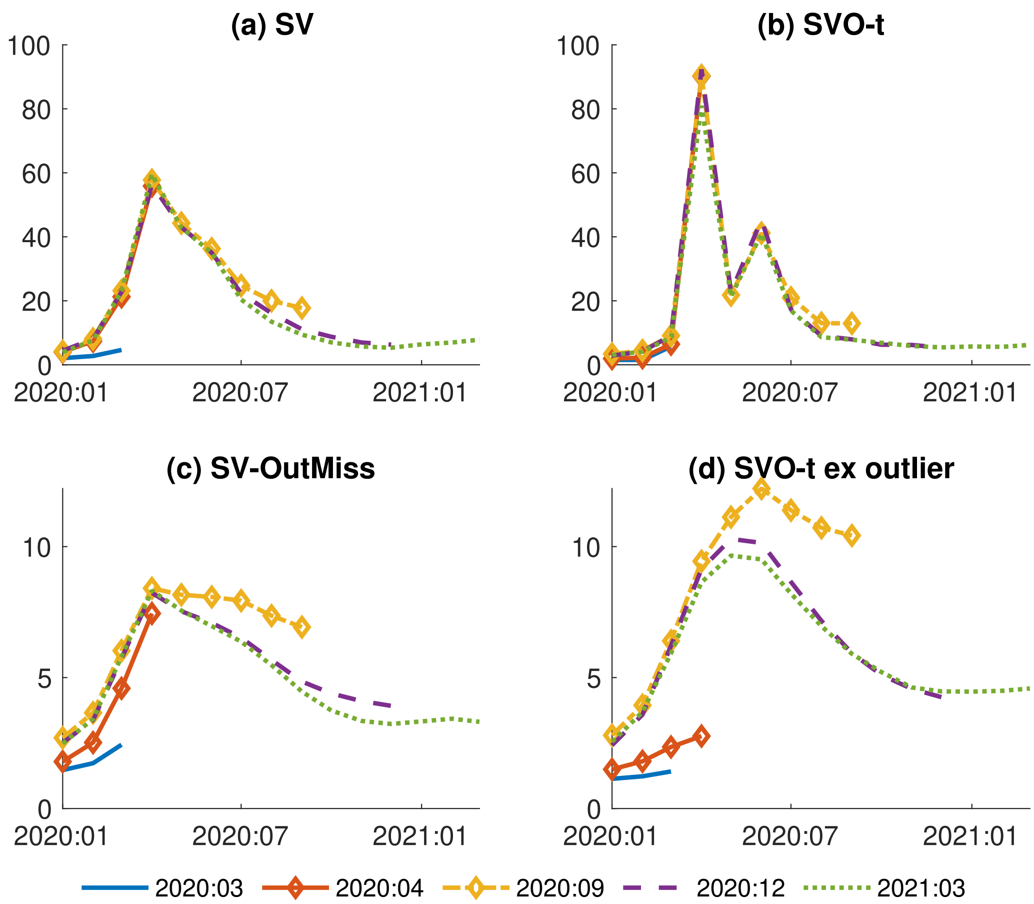
Note: Full-sample estimates per March 2021 of posterior probabilities for realizations of the reduced-form outlier states in different models. For the SVO-t model, we compute reduced-form outlier scales from the ratios between diagonal elements of $\Sigma_t = A^{-1} O_t Q_t \Lambda_t Q_t' O_t A^{-T}$ and $\tilde{\Sigma}_t = A^{-1} \Lambda_t A^{-T}$; denoting the j-th diagonal elements by $\tilde{\sigma}_t(j)^2$ and $\sigma_t(j)^2$, the outlier scale for the j-th variable is $\sigma_t(j)/\tilde{\sigma}_t(j)$. For SVO and SV-t, corresponding computations are performed using only O_t and Q_t, respectively. To simulate the posterior distribution for the reduced-form outlier states, these computations are performed for every MCMC draw. Each panel shows posterior probabilities for the outlier states 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).

Figure 8: Outlier probabilities of SVO and SVO-t



Note: Prior and posterior distribution of the outlier probability p_j in in the orthogonalized VAR residuals of selected variables in the SVO and SVO-t model for selected variables, estimated from the full sample of data available from March 1959 through March 2021.

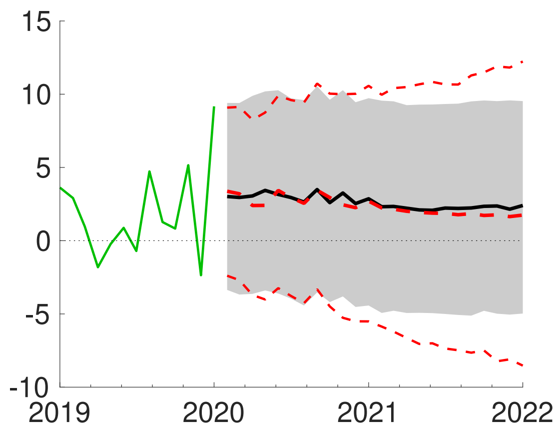
Figure 9: Time-varying volatilities since 2020 of payroll growth



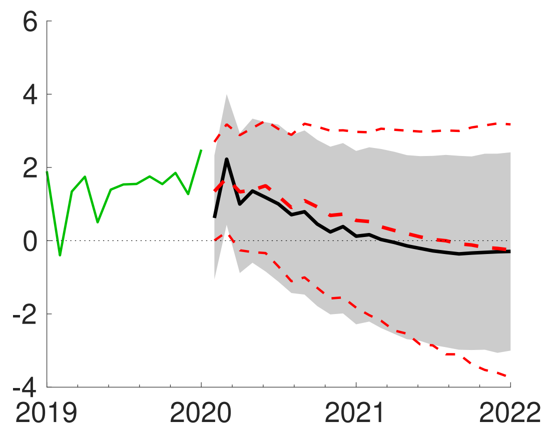
Note: Quasi-real-time trajectories of time-varying volatility in VAR residuals, measured by the diagonal elements of $\mathrm{Var}_t\left(v_t\right) = \Sigma_t$ implied by different models. Medians of (smoothed) posterior obtained from different data samples ending at forecast origins as indicated in the figure legend. Panels (b) and (d) display estimates of stochastic volatility for SVO-t that ignore the contributions from outliers and that are computed from $\tilde{\Sigma}_t = A^{-1}\,\Lambda_t\,A^{-T}$ (i.e., neglecting the O_t and Q_t components in the computation of the uncertainty measures shown here, while including these outliers in estimation of A^{-1}, Λ_t, etc.). Reflecting the sizable differences in the size of estimates resulting with and without outlier treatment, different scales are used in upper- and lower-row panels.

Figure 10: Predictive densities in January 2020 for selected variables

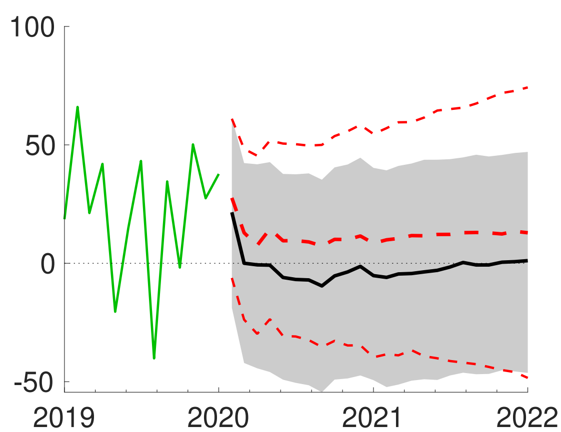
(a) Real income



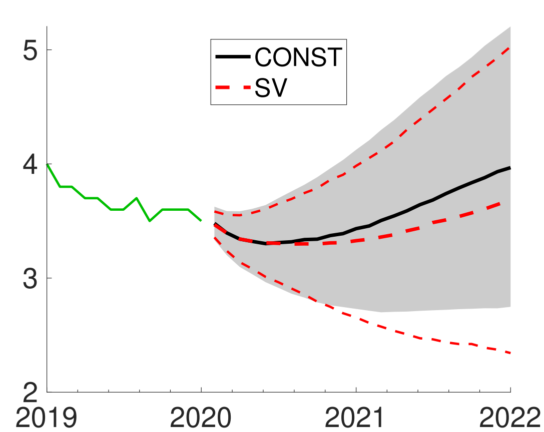
(b) Payroll growth



(c) S&P 500 returns

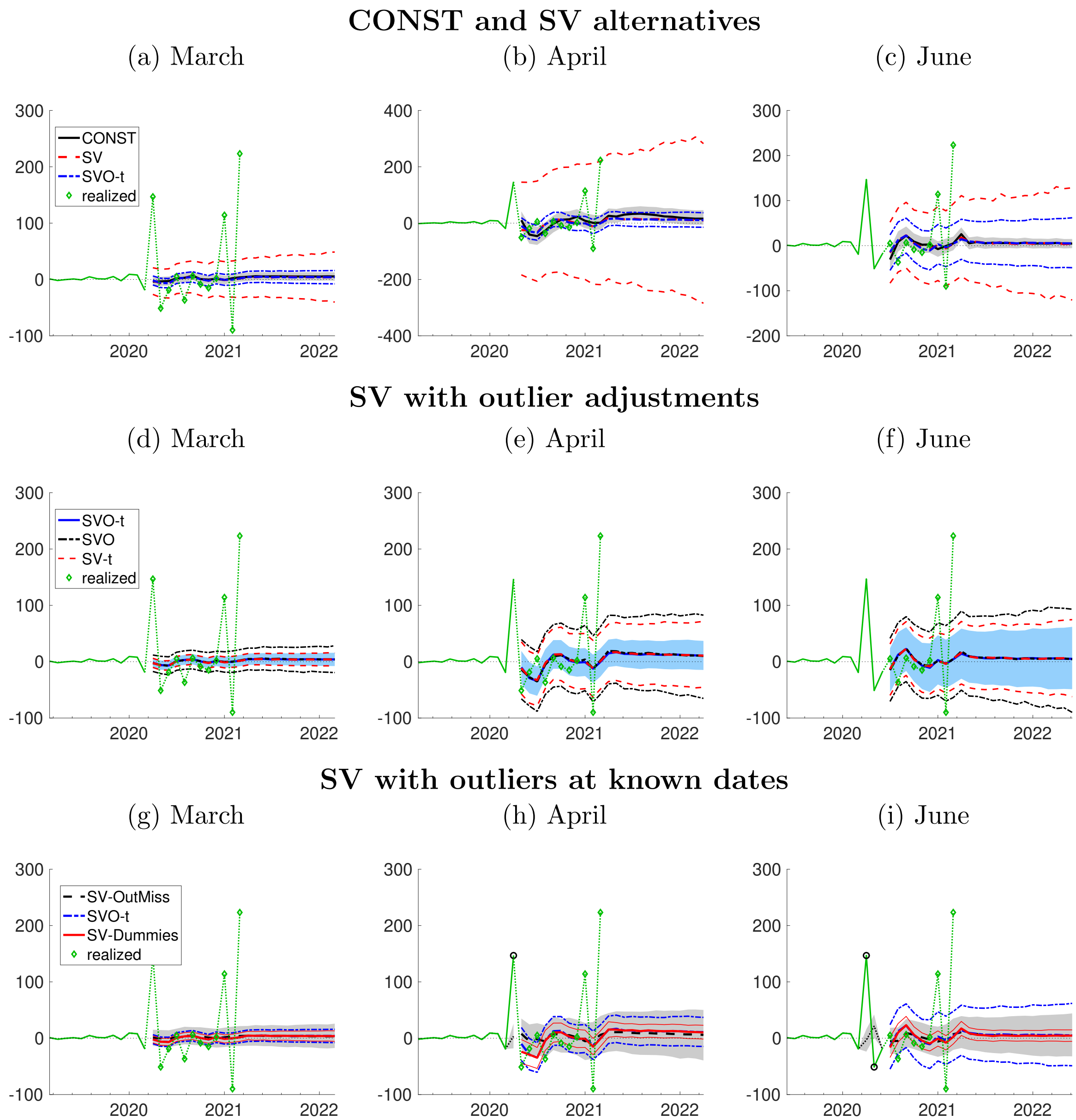


(d) Unemployment rate



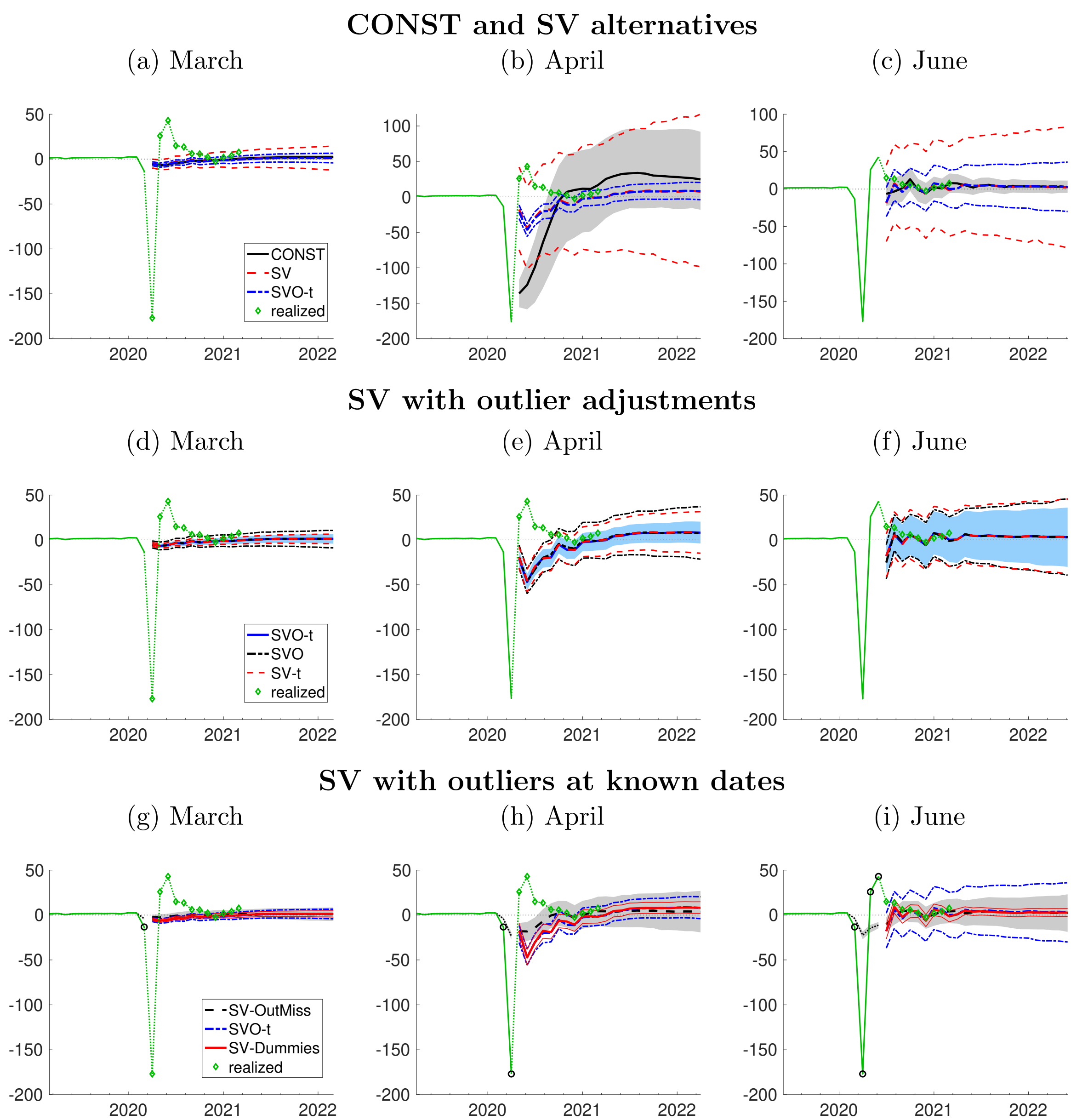
Note: Medians and 68% uncertainty bands of predictive densities, simulated out-of-sample at various forecast origins as indicated in each panel. The solid green line denotes realized data prior to the forecast origin.

Figure 11: Predictive densities since March 2020 for real income



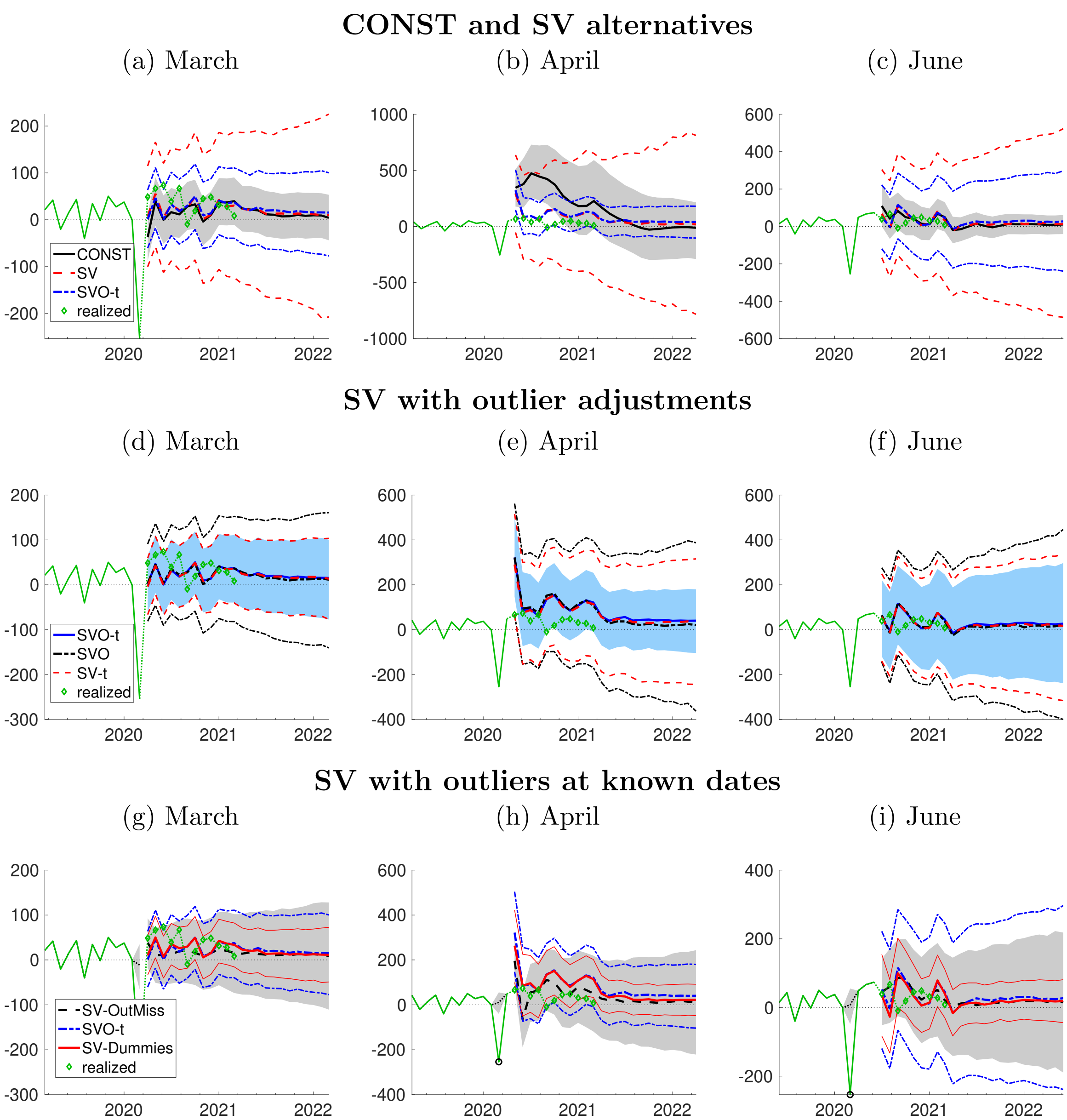
Note: Medians and 68% uncertainty bands of predictive densities, simulated out-of-sample at various forecast origins as indicated in each panel. The solid green line denotes realized data prior to the forecast origin. In panels (g) – (i), observations identified ex-ante as outliers, based on being more than 5 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



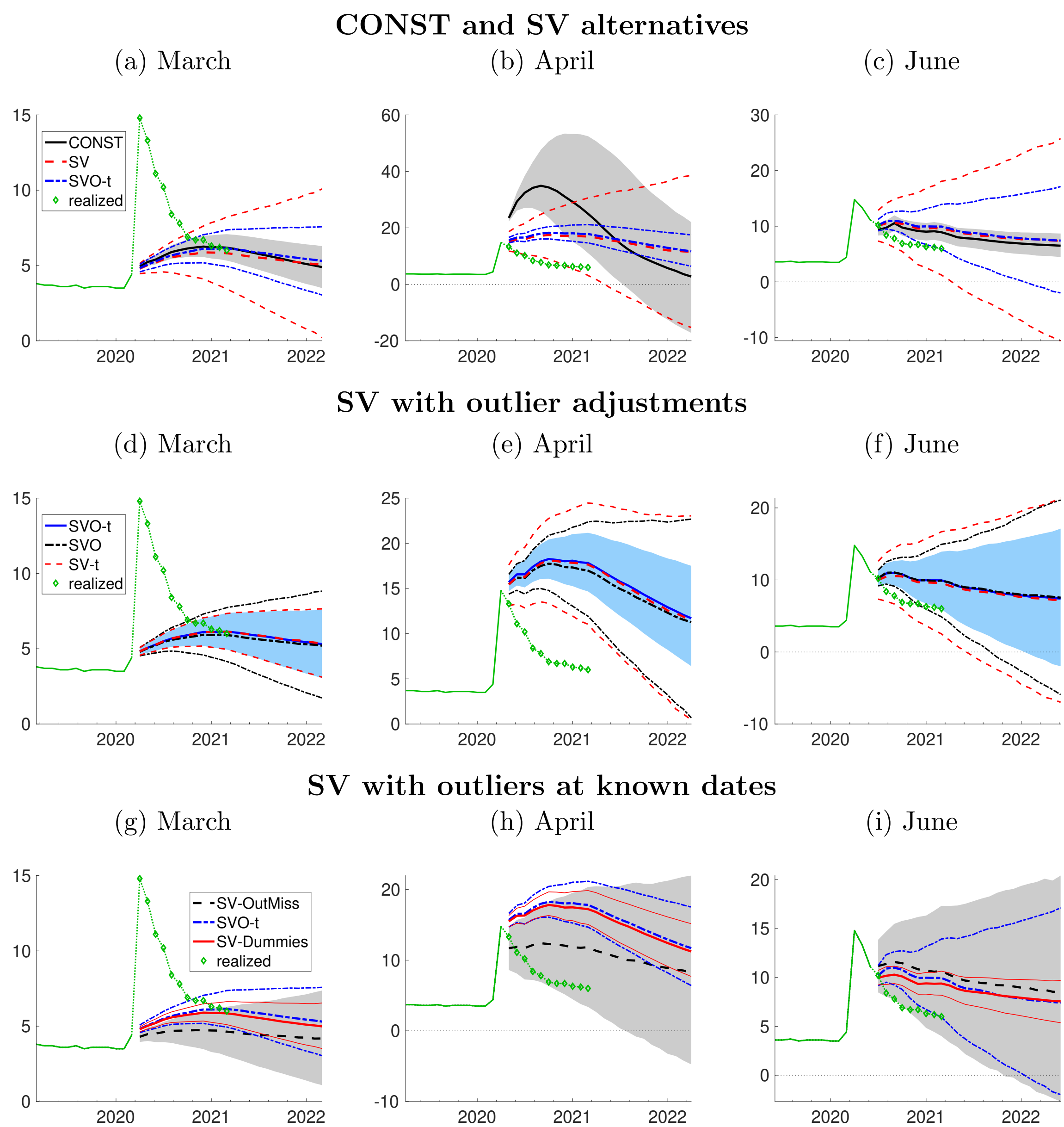
Note: Medians and 68% uncertainty bands of predictive densities, simulated out-of-sample at various forecast origins as indicated in each panel. The solid green line denotes realized data prior to the forecast origin. In panels (g) – (i), observations identified ex-ante as outliers, based on being more than 5 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



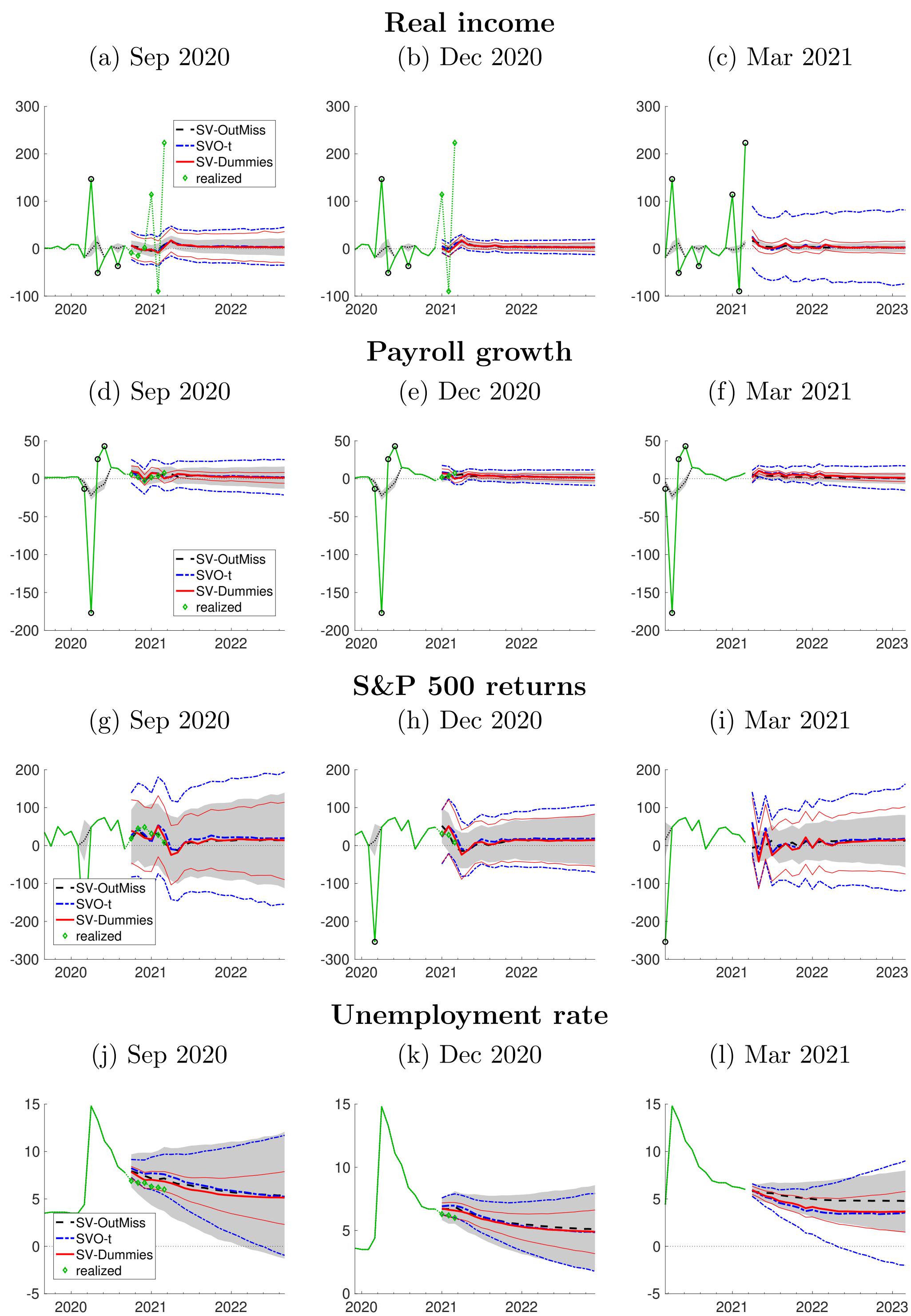
Note: Medians and 68% uncertainty bands of predictive densities, simulated out-of-sample at various forecast origins as indicated in each panel. The solid green line denotes realized data prior to the forecast origin. In panels (g) – (i), observations identified ex-ante as outliers, based on being more than 5 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 (g) – (i), observations identified ex-ante as outliers, based on being more than 5 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.

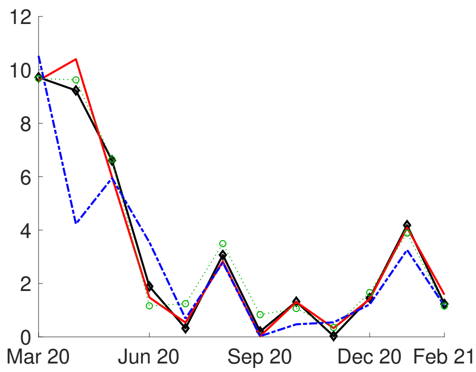
Figure 15: Predictive densities since late 2020



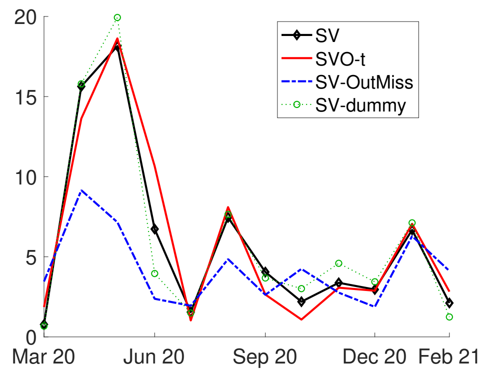
Note: Medians and 68% uncertainty bands of predictive densities, simulated out-of-sample at various forecast origins as indicated in each panel. Forecasts generated from the SV-OutMiss approach idenitfy observations ex-ante as outliers, based on being more than 5 times the inter-quartile range away from the median; these outliers are indicated with a circle, and the corresponding backcast densities from the SV-OutMiss model are superimposed.

Figure 16: Absolute error forecast losses since March 2020

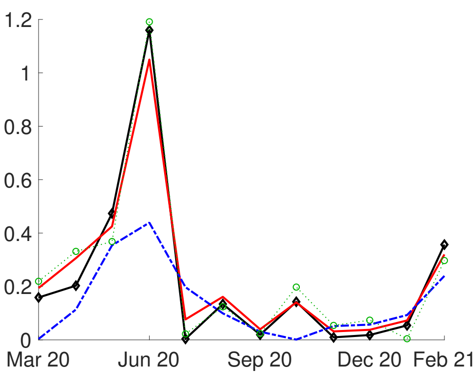
(a) Capacity utilization



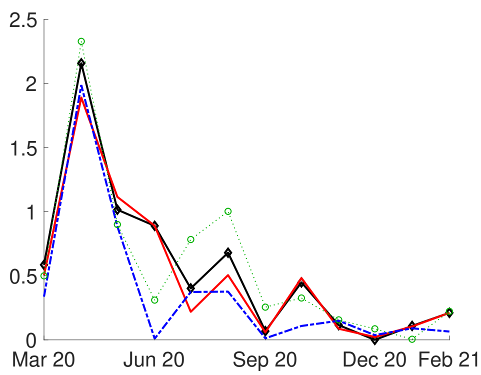
(b) Hourly Earnings



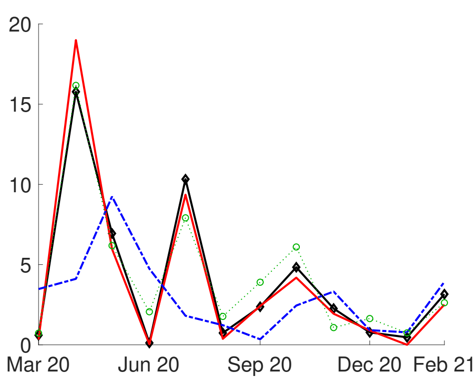
(c) 10-year yield



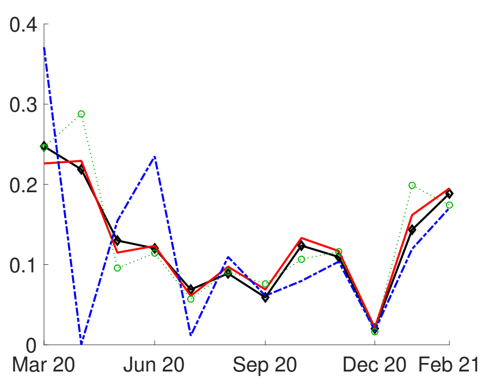
(d) Baa spread



(e) PCE price inflation



(f) Housing starts



Note: Absolute forecast errors for one-steap ahead forecasts, generated out-of-sample by various models since March 2020. Dates on the x-axis indicate the forecast origin.

Table 2: RMSE (baseline comparisons)

					Relative to SV ...							
	SV				CONST				SVO-t			
Variable / Horizons	1	3	12	24	1	3	12	24	1	3	12	24
Real Income	7.60	7.65	7.72	8.44	1.00	1.01	1.00	0.92*	1.01	1.00	1.01**	0.93*
Real Consumption	5.38	5.60	5.36	5.05	1.00	1.00	1.00	1.00	1.01	1.00	1.00	1.01
IP	6.68	7.08	7.80	8.39	1.01	1.00	0.99	0.98	1.00	0.99	1.00	0.96***
Capacity Utilization	0.45	0.85	2.79	4.28	1.02*	1.00	0.99	0.96	1.00	0.99	1.00	0.97
Unemployment	0.14	0.22	0.73	1.33	0.99	0.97	1.00	1.05	1.00	0.99	0.99	0.99
Nonfarm Payrolls	1.21	1.33	1.86	2.11	1.04*	1.02	1.02	1.01	0.99	1.00	1.01	0.98
Hours	0.19	0.23	0.40	0.43	1.04***	1.00	1.04	1.07*	1.00	1.00	0.99	1.00
Hourly Earnings	2.48	2.46	2.54	2.73	1.04**	1.01	1.03	1.06	1.01	1.00	1.01**	1.03*
PPI (Fin. Goods)	7.16	7.35	7.58	7.43	1.00	1.02	1.05	1.07	1.00	0.99	1.00	1.00
PCE Prices	2.11	2.42	2.64	2.81	1.01	1.03	1.13*	1.19**	1.00	1.00	1.01	1.03*
Housing Starts	0.07	0.10	0.23	0.35	0.99	0.99	1.03	1.08**	0.99	0.99	0.99	1.03***
S&P 500	43.71	44.17	44.03	43.28	0.99	1.01	1.00	1.00	1.00	1.00	1.00	1.01**
USD / GBP FX Rate	28.35	29.79	28.26	33.35	1.03**	1.02	1.02	0.86	1.00	1.00	1.00	0.86
5-Year yield	0.25	0.55	1.14	1.43	1.05***	1.06***	1.04	0.92	1.01**	1.00	1.01	0.97
10-Year yield	0.23	0.51	1.08	1.22	1.03	1.06**	1.08*	1.03	1.01**	1.00	1.01	0.98
Baa Spread	0.26	0.61	1.33	1.51	1.20***	1.25**	1.12*	1.15	1.00	0.99	0.99	0.97

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

Table 3: Avg CRPS (baseline comparisons)

					Relative to SV ...							
	SV				CONST				SVO-t			
Variable / Horizons	1	3	12	24	1	3	12	24	1	3	12	24
Real Income	3.17	3.32	3.63	4.19	1.09***	1.00	0.94*	0.82***	0.99	0.96***	0.94***	0.86***
Real Consumption	2.76	2.88	3.05	3.56	1.04**	1.04***	0.98	0.82***	1.00	0.99	0.97***	0.91***
IP	3.65	3.83	4.50	5.37	1.03**	1.02	0.95***	0.84***	1.00	0.99*	0.96***	0.90***
Capacity Utilization	0.26	0.48	1.54	2.57	1.03**	1.04*	0.99	0.90**	1.00	0.99	1.00	0.96
Unemployment	0.08	0.12	0.38	0.72	0.99	1.00	1.03	1.05	1.00	1.00	1.01	1.00
Nonfarm Payrolls	0.69	0.77	1.12	1.48	1.10***	1.07***	0.99	0.84***	0.99	1.00	0.98*	0.93***
Hours	0.10	0.13	0.23	0.29	1.07***	1.02	1.02	0.91**	0.99	0.99	0.98*	0.92***
Hourly Earnings	1.39	1.41	1.58	2.04	1.07***	1.05***	0.98	0.84***	1.00	0.99**	0.98***	0.93***
PPI (Fin. Goods)	3.75	3.90	4.15	4.58	1.02	1.02	1.02	0.93**	1.00	0.99*	0.98***	0.95***
PCE Prices	1.10	1.26	1.47	1.81	1.03**	1.04*	1.07	0.96	1.01**	1.00	1.00	0.98***
Housing Starts	0.04	0.05	0.12	0.19	0.99	1.00	1.05	1.03	1.00	1.00	1.01	1.01*
S&P 500	22.98	23.44	25.00	28.39	1.00	1.01	0.94***	0.82***	1.00	0.99**	0.97***	0.92***
USD / GBP FX Rate	15.59	16.14	16.53	18.39	1.03**	1.02	0.96**	0.86***	0.99	0.99*	0.97***	0.92***
5-Year yield	0.14	0.30	0.63	0.80	1.06***	1.08***	1.06	0.97	1.01***	1.00	1.01*	1.01
10-Year yield	0.13	0.28	0.60	0.75	1.04***	1.06***	1.08*	1.00	1.01***	1.01	1.01	1.01*
Baa Spread	0.14	0.32	0.75	1.01	1.34***	1.30***	1.11*	0.99	1.00	0.99	0.99	0.97**

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

Table 4: Relative RMSE (outlier-adjusted models)

	SVO				SV-t				SV-OutMiss			
Variable / Horizon	1	3	12	24	1	3	12	24	1	3	12	24
Real Income	0.99	1.00	1.00	1.01**	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.02
Real Consumption	0.99**	1.00	0.99	0.99	1.00	1.00**	1.00	1.00	0.98	0.98	1.00	1.00
IP	1.00	1.00	1.00	1.02*	1.00	1.01*	1.00	1.01	1.00	1.01	1.00	1.02**
Capacity Utilization	0.99	1.00	0.99	1.01	1.00	1.01	1.00	0.99	1.00	1.03	0.99	1.00
Unemployment	1.00	1.01	1.00	1.00	1.00	1.00	1.00	0.99**	0.99	1.00	1.00	1.01
Nonfarm Payrolls	1.01	1.00	0.99	1.00	1.01	1.01	0.99**	1.00	1.00	1.00	0.98	1.00
Hours	1.00	1.01	0.99	0.99	1.00	1.00	1.00	1.00	1.00	1.01	1.00	1.00
Hourly Earnings	0.99	1.00	0.98***	0.96***	1.00	1.00	1.00	1.00	0.99	1.00	0.99	0.97**
PPI (Fin. Goods)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.01	1.00	1.01	1.00	1.00
PCE Prices	0.99	1.00	0.99	0.97*	1.00	1.00	1.00	1.00	0.99	1.00	1.01	0.99
Housing Starts	1.00	1.00	1.00	0.98*	1.00	1.00	1.00	1.00	1.00	1.01	1.00	0.98**
S&P 500	1.00	1.00	1.00	1.00	1.00	1.00**	1.00	1.00	1.00	1.00	1.00	0.99
USD / GBP FX Rate	0.99**	1.00	0.99**	0.99	1.00	1.00	1.00	1.01	1.01	0.99	1.00	0.98**
5-Year yield	0.99*	0.99	0.99**	1.02	1.00	1.00	1.00	0.99	0.99	0.99	0.99	0.99
10-Year yield	1.00	0.99	0.99**	1.00	0.99	1.00	1.01	1.00	0.99	0.99	0.99	0.99
Baa Spread	1.00	0.99	0.98	1.02	1.00	1.00	1.01	1.00	1.01	1.00	0.99	1.04

Note: Comparison of "SVO-t" (baseline, in denominator) against "SVO" and "SV-t." Values below 1 indicate improvement over baseline. Evaluation window from 1985:M01 through 2017:M12. Significance assessed by Diebold-Mariano-West 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 significant relative RMSE 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 (outlier-adjusted models)

	SVO				SV-t				SV-OutMiss			
Variable / Horizon	1	3	12	24	1	3	12	24	1	3	12	24
Real Income	1.00	1.01*	1.03***	1.08***	1.00	1.00	1.00**	0.99**	0.99	0.99*	1.00	1.01
Real Consumption	0.99	1.01**	1.05***	1.09***	1.00	1.00	0.99***	0.98***	0.98	0.99	1.01	1.04***
IP	0.99	1.02**	1.05***	1.10***	1.00	1.00	0.99***	0.99***	0.99	1.02*	1.02***	1.07***
Capacity Utilization	1.00	1.01	1.02	1.05*	1.00	1.00	0.99***	0.98***	0.99	1.02	0.99	1.00
Unemployment	1.00	1.01**	1.01	1.02	1.00	1.00	0.99***	0.99***	0.99	0.99	0.98	1.00
Nonfarm Payrolls	1.02**	1.02**	1.04***	1.09***	1.00	1.00	0.98***	0.98***	1.00	1.00	1.00	1.03**
Hours	1.01**	1.02***	1.02**	1.07***	1.00	1.00*	0.99***	0.98***	1.01	1.02*	1.02	1.05***
Hourly Earnings	1.00	1.02***	1.05***	1.09***	1.00	1.00	0.98***	0.98***	0.99	1.01	1.01	1.04***
PPI (Fin. Goods)	1.00	1.01**	1.03***	1.06***	1.00	1.00	0.99**	0.99***	0.99	1.00	1.01**	1.02***
PCE Prices	0.99**	1.01	1.02**	1.05***	1.00**	1.00	0.99***	0.98***	0.98**	0.99**	1.00	0.99
Housing Starts	1.00	1.01**	1.01*	1.01	1.00	1.00	0.99	0.99	0.99	1.00	0.99	0.98**
S&P 500	1.01	1.02**	1.04***	1.07***	1.00	1.00	0.99***	0.99***	1.00	1.01	1.01**	1.05***
USD / GBP FX Rate	0.99*	1.00	1.01***	1.03***	1.00	1.00	1.00	1.00	1.01	0.99	1.00	1.02
5-Year yield	1.00	0.99	1.00	1.03***	0.99*	1.00	1.00	0.98***	0.99	0.99	0.99	0.98***
10-Year yield	1.00	1.00	1.01	1.04***	0.99**	1.00	1.00	0.98***	0.99	0.99	0.99	0.98***
Baa Spread	1.00	1.00	1.02	1.07***	1.00	1.00	1.00	0.98***	1.00	0.99	0.99	1.01

Note: Comparison of "SVO-t" (baseline, in denominator) against "SVO" and "SV-t." Values below 1 indicate improvement over baseline. Evaluation window from 1985:M01 through 2017:M12. Significance assessed by Diebold-Mariano-West 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 significant relative CRPS of 1.00. These cases arise from persistent differences in performance that are, however, too small to be relevant after rounding.

Table 6: Relative RMSE during COVID-19 episode

	SVO-t			SV-OutMiss			SV-dummy		
Variable / Horizon	1	3	6	1	3	6	1	3	6
Real Income	1.01	1.00	1.01	0.98	0.99	1.02	1.01	1.01	1.02
Real Consumption	1.01	1.02	1.04	1.01	0.97	0.95	0.99	0.99	1.01
IP	1.03	1.03	1.04	0.99	0.92	1.03	1.02	1.01	0.99
Capacity Utilization	1.02	1.07	1.12	0.88	0.75	0.77	1.01	1.06	1.11
Unemployment	1.02	1.06	1.06	1.03	0.88	0.74	1.00	0.99	1.01
Nonfarm Payrolls	1.01	1.01	1.12	0.97	0.81	0.40	1.00	0.99	1.05
Hours	1.02	1.03	1.09	0.84	0.70	0.75	1.00	1.04	1.07
Hourly Earnings	1.02	1.13	0.93	0.60	0.78	0.79	1.04	1.05	0.94
PPI (Fin. Goods)	1.10	1.04	0.88	0.83	0.71	0.76	1.02	1.06	0.99
PCE Prices	1.07	1.00	0.99	0.63	0.85	0.75	0.98	1.04	0.94
Housing Starts	1.01	1.06	1.01	1.10	1.06	1.02	1.08	1.09	1.03
S&P 500	1.10	0.98	1.09	0.79	0.75	0.65	0.88	0.90	1.05
USD / GBP FX Rate	1.06	1.05	1.16	1.05	1.22	0.75	1.01	0.92	1.00
5-Year yield	0.89	0.78	0.61	0.60	0.77	0.63	1.05	0.68	0.38
10-Year yield	0.93	0.77	0.59	0.50	0.83	0.66	1.02	0.66	0.38
Baa Spread	0.92	1.01	1.07	0.82	0.85	0.96	1.05	1.21	1.34

Note: Comparison of "SV" (baseline, in denominator) against "SVO-t" and "SV-OutMiss." Values below 1 indicate improvement over baseline. Evaluation window from 2020:M03 through 2021:M03. Due to the low number of observations in the evaluation window, significance tests have not been performed.

Table 7: Relative Avg CRPS during COVID-19 episode

	SVO-t			SV-OutMiss			SV-dummy		
Variable / Horizon	1	3	6	1	3	6	1	3	6
Real Income	1.02	0.92	0.89	1.04	0.92	0.98	1.05	0.96	0.93
Real Consumption	1.07	0.85	0.69	1.11	0.78	0.63	1.01	0.78	0.68
IP	1.10	1.02	0.69	1.01	0.90	0.84	1.12	0.94	0.46
Capacity Utilization	1.10	1.18	1.16	0.89	0.84	0.83	1.16	1.23	1.12
Unemployment	1.10	1.28	1.28	1.01	0.93	0.85	1.01	1.17	1.26
Nonfarm Payrolls	1.09	1.00	0.59	0.99	0.72	0.31	1.09	0.95	0.44
Hours	1.06	1.06	0.96	0.92	0.81	0.70	1.06	1.08	0.88
Hourly Earnings	1.04	0.94	0.73	0.70	0.74	0.66	1.04	0.87	0.65
PPI (Fin. Goods)	1.07	0.89	0.72	0.81	0.74	0.70	1.04	0.81	0.59
PCE Prices	1.03	0.94	0.74	0.68	0.73	0.60	0.99	0.96	0.54
Housing Starts	0.97	0.94	0.88	0.98	0.92	0.85	0.97	0.91	0.88
S&P 500	0.98	0.80	0.71	0.70	0.56	0.48	0.79	0.64	0.51
USD / GBP FX Rate	0.99	0.94	0.87	0.93	0.98	0.69	0.95	0.82	0.75
5-Year yield	0.93	0.81	0.65	0.63	0.88	0.74	1.11	0.78	0.46
10-Year yield	1.00	0.90	0.71	0.61	0.92	0.80	1.06	0.83	0.57
Baa Spread	0.95	1.02	0.94	0.78	0.86	0.88	1.04	1.23	1.08

Note: Comparison of "SV" (baseline, in denominator) against "SVO-t" and "SV-OutMiss." Values below 1 indicate improvement over baseline. Evaluation window from 2020:M03 through 2021:M03. Due to the low number of observations in the evaluation window, significance tests have not been performed.