KOÇ UNIVERSITY-TÜSİAD ECONOMIC RESEARCH FORUM
WORKING PAPER SERIES

USING SURVEY INFORMATION FOR IMPROVING THE DENSITY NOWCASTING OF US GDP WITH A FOCUS ON PREDICTIVE PERFORMANCE DURING COVID-19 PANDEMIC

Cem Çakmaklı
Hamza Demircan

Working Paper No: 2016
October 2020

KOÇ UNIVERSITY-TÜSİAD ECONOMIC RESEARCH FORUM
Rumelifeneri Yolu 34450 Sarıyer/Istanbul

Using Survey Information for Improving the Density Nowcasting of US GDP with a Focus on Predictive Performance during Covid-19 Pandemic

Cem Çakmaklı * [1] and Hamza Demircan, † [2]

[1] *Koç University*
[2] *Central Bank of the Republic of Turkey*

October 6, 2020

Abstract

We provide a methodology that efficiently combines the statistical models of nowcasting with the survey information for improving the (density) nowcasting of US real GDP. Specifically, we use the conventional dynamic factor model together with a stochastic volatility component as the baseline statistical model. We augment the model with information from the survey expectations by aligning the first and second moments of the predictive distribution implied by this baseline model with those extracted from the survey information at various horizons. Results indicate that survey information bears valuable information over the baseline model for nowcasting GDP. While the mean survey predictions deliver valuable information during extreme events such as the Covid-19 pandemic, the variation in the survey participants' predictions, often used as a measure of 'ambiguity', conveys crucial information beyond the mean of those predictions for capturing the tail behavior of the GDP distribution.

Keywords:*Dynamic factor model; Stochastic volatility; Survey of Professional Forecasters; Disagreement; Predictive density evaluation; Bayesian inference*
JEL Classification: C32, C38, C53, E32, E37

*Cem Çakmaklı, Koç University, Istanbul, Turkey, e-mail: ccakmakli@ku.edu.tr.
†Hamza Demircan, Central Bank of the Republic of Turkey, Istanbul, Turkey, e-mail: Hamza.Demircan@tcmb.gov.tr.

1 Introduction

Monitoring economic conditions in a timely and accurate manner is crucial for economic agents. Gross domestic product (GDP), however, as the key indicator of the economic conditions, is not available instantly as it is released with a substantial delay. Two sources of GDP projections are typically available for the decision maker. First, econometric models of nowcasting have been proved to be very useful in providing accurate predictions of the GDP density using a large set of economic and financial indicators that are timely available. Second, surveys, reporting predictions of forecasters, serve as an important guide reflecting prompt reactions of these forecasters to changing economic conditions.

Econometric specifications constructed for nowcasting often relies on dynamic factor models, among others,[1] to extract the common movement in the economy, which in turn, is used for the prediction of the current GDP. Dynamic factor models have the advantage of processing datasets composed of a large number of macroeconomic and financial variables sampled at mixed frequencies in a statistically optimal manner to extract the common behavior in these series or put differently, factors, see for example Giannone et al. (2009); Banbura et al. (2013); Bok et al. (2018). Coupled with the time-varying volatility, factor models constitute the workhorse for predicting the current economic activity and its distribution, see for example Marcellino et al. (2016) for a recent study.

Surveys, on the other hand, reflecting the expectations of the key economic agents, bear specific information that either based on the specific model or the judgment of the survey participants about future economic conditions. Essentially, survey participants form their 'judgemental' expectations based on both public and private information they might have. Ang et al. (2007); Campbell (2007); D'agostino et al. (2012); Faust and Wright (2013), among others, document that surveys indeed convey useful information beyond that is provided by the statistical models. Following the great recession of 2008-9, a recent strand of research focuses on the ambiguity reflected in the predictions of survey participants as a proxy for uncertainty especially during turmoil periods, see, for example, Giordani and Söderlind (2003); Lahiri and Sheng (2010); Bloom (2014); Grischenko et al. (2019). Therefore, we

[1] Other popular methods used for nowcasting the GDP involve mixed frequency VARs, see for example Giannone et al. (2009); Foroni et al. (2013), among others, or MIDAS and bridge equations, see for example Aastveit et al. (2014); Schumacher (2016), among others, which make use of datasets at mixed frequencies.

distinguish two important features of survey-based information. First, survey participants provide useful information that might diverge from the statistical models particularly at times of high uncertainty. Second, disagreement among survey participants serves as an important proxy of ambiguity those agents confront.

Our point of departure in this paper is the construction of an efficient combination of the baseline statistical model with the survey information for exploiting the distinct features of these two approaches. First, we depart from conventional models of nowcasting based on the dynamic factor model and we augment this model with a stochastic volatility component for the GDP. By doing this, we construct a competitive baseline statistical model for nowcasting the density of the GDP using the time variation in the mean and the volatility. Second, our methodology of integrating the survey information into the statistical model relies on the fact that survey information represents the predictions formed by the survey participants. Hence, our combination strategy aligns the predictions implied by the statistical model with the predictions of the survey participants. We do this by matching both the mean and the variance of the predictive distributions obtained from the statistical model together with the mean and the variance of the survey expectations. This corresponds to augmenting the baseline model with new measurement equations representing these alignments.

We evaluate a dynamic factor model (using a dataset comprised of 17 variables) with and without survey information and stochastic volatility components based on a horse-race in terms of various measures of density evaluation. In particular, we employ marginal and predictive likelihood metrics as well as Probability Integral Transforms (PIT) to evaluate the full-sample and out-of-sample performance of competing models with a special focus on the tail behavior. Considering baseline statistical models, we observe that adding stochastic volatility over the conventional baseline model with constant volatility does not improve the model fit but it improves nowcasting and forecasting performances. On the other hand, when we combine statistical models with the survey information, obtained from the Survey of Professional Forecasters (SPF), the model performance improves substantially. While, using only the mean of the survey expectations improves upon the baseline model, incorporating the variance of the survey expectations provides a further improvement. Interestingly, survey expectations at all horizons including the survey nowcasts of the current quarter and forecasts up to 4-quarter ahead bear additional information beyond that is provided by the baseline statistical model. These findings are robust for both nowcasts and short- and medium-horizon forecasts up to two quarters ahead. The results based on the PITs are also in line with these findings in the sense that models, where survey information is integrated into the baseline models, perform better than the plain baseline models. A crucial finding is on the value-added provided by the disagreement of the survey participants captured by the variance of the survey-based expectations. Results show that the models perform sufficiently well at tails of the distribution only when we integrate the second moment of the survey-based expectations to the baseline statistical model. Hence, the measure of disagreement delivers useful information regarding the tails of the GDP distribution.

We further demonstrate the efficacy of the proposed framework by focusing on the month-by-month performance of the selected models during the first six months of 2020, that is, during the exceptional recent periods when the Covid-19 pandemic hit the US economy. In this case, while the baseline statistical model predicts a relatively milder downturn than the realized extreme downturn in the second quarter, the survey forecasters' mean prediction turns out to be quite accurate. Combining these two sources, the proposed model outperforms the baseline statistical model clearly both in terms of point forecast and density forecast. These demonstrate the valuable information provided by the survey participants during turbulent times.

Our approach is closely related to the approach followed by Grischenko et al. (2019) who combines the predictions from an unobserved components model together with the information in survey-based expectations. They do so for measuring the inflation uncertainty and whether inflation expectations are anchored in the US and Euro Area by taking this uncertainty into account, see also Kozicki and Tinsley (2006, 2012) who uses only the first moment of survey expectations. Similarly, Altug and Çakmaklı (2016) utilizes this approach for forecasting inflation and measuring the deviation of inflation from target inflation. Aruoba (2020) derives the term structure of inflation expectations by matching the first moment of the survey-based inflations with the predictions of a dynamic factor model, specifically the Nelson-Siegel model. A related line of research uses survey information to combine it with the forecasts based on Bayesian VectorAutoregressions (BVAR) involving a relatively limited number of variables typically using entropic tilting methods. Krüger et al. (2017), for example, employ entropic tilting using both the first and the second moments for matching the medium-term forecasts from a BVAR with the nowcasts from surveys, see also Tallman and Zaman (2019) for a similar approach. Altavilla et al. (2017) use entropic tilting for combining survey-based expectations with the predictions of the Nelson-Siegel model, for predicting the yield curve. Our approach, on the other hand, focuses on nowcasting and short-term forecasting of the density of the US real GDP by utilizing a dynamic factor model using a large dataset with 17 variables with mixed frequency together with the survey-based information on the first and second moments of the real GDP growth distribution.

The remainder of this paper is as follows. Section 2 presents model specifications. Section 3 discusses the data, details on estimation methodology and model evaluation.[2] Section 4 evaluates the empirical results and provides a detailed comparison between alternative model strategies. Section 5 discusses the performance of competing models during the Covid-19 pandemic recession. Finally, Section 6 concludes.

2 Model Details

In this section, we present the baseline dynamic factor model together with stochastic volatility for nowcasting the density of the GDP. For the baseline model, we closely follow Banbura et al. (2013) with the addition of the stochastic volatility. The key component of the model constitutes the incorporation of the survey expectations about the first and second moments of the predictive distribution of the GDP to the baseline model.

Econometric models of GDP nowcasting with mixed frequency datasets often involve GDP as the only quarterly variable that is complemented with a bulk of monthly or higher frequency variables that are timely available. We first demonstrate the model in terms of monthly variables in the set of high frequency variables and then we incorporate the GDP to the model.[3] Let $y_t^m = [y_{1,t}^m, y_{2,t}^m, \ldots, y_{n_m,t}^m]'$ for $t = 1, 2, .., T$ denote the n_m-dimensional

[2]We provide full details about the simulation-based inference of the competing models in Section A of the supplementary material for the sake of brevity.

[3]Banbura et al. (2013) show that variables that are sampled at higher frequency than monthly provide

vector of variables in period t. The superscript of 'm' denotes the variables at monthly frequency. The variables are transformed to ensure stationarity if necessary and further standardized. We assume that these variables admit a factor structure as

$$y_t^m = \lambda^m f_t^m + \varepsilon_t^m \ , \quad \varepsilon_t^m \sim N(0, \Sigma), \tag{1}$$

where f_t^m is a vector of latent common factors, λ^m is n_m-dimensional vector of factor loadings and the covariance matrix, denoted as Σ, is of diagonal structure with the diagonal elements as $\sigma_1^2, \sigma_2^2, \ldots, \sigma_{n_m}^2$. We assume that idiosyncratic factors, denoted as $\varepsilon_t^m = [\varepsilon_{1,t}^m, \varepsilon_{2,t}^m, .., \varepsilon_{n_m,t}^m]'$, are uncorrelated with f_t^m for all leads and lags.

For the dynamics of the common factor, we proceed with a single factor structure that obeys an AR(1) specification as

$$f_t^m = \phi f_{t-1}^m + u_t^m \ , \quad u_t^m \sim N(0, \sigma_u^2). \tag{2}$$

As noted in Banbura et al. (2013), employing more than one factor together with higher order autoregressive dynamics do not change the results qualitatively. Additionally, such simple specification keeps the model parsimonious and tractable for further extensions.

Since GDP is a quarterly flow variable, a special attention is required on its link to the monthly factor. Let z_t^q and z_t denotes the logarithm of quarterly and monthly (unobserved) GDP, respectively. Further, let y_t^q and y_t be the growth rates of the quarterly and monthly (unobserved) GDP, respectively. Following the approximation $z_t^q \approx \sum_{k=0}^{2} z_{t-k}$ employed by Mariano and Murasawa (2003) and Bańbura and Modugno (2014), among others, y_t^q is linked to y_t as

$$\begin{aligned} y_t^q = (\sum_{k=0}^{2} z_{t-k}) - (\sum_{k=0}^{2} z_{t-3-k}) &= \Delta z_t + 2\Delta z_{t-1} + 3\Delta z_{t-2} + 2\Delta z_{t-3} + \Delta z_{t-4} \\ &= y_t + 2y_{t-1} + 3y_{t-2} + 2y_{t-3} + y_{t-4} \\ &= \sum_{k=0}^{4} w_k y_{t-k} \end{aligned} \tag{3}$$

where w_k denotes the weights for aggregation. Accordingly, we represent the measurement

little or no improvement for prediction of the GDP.

equation for the quarterly GDP growth rate as

$$y_t^q = \sum_{k=0}^{4} w_k y_{t-k} \quad = \quad \sum_{k=0}^{4} w_k(\lambda^m f_{t-k}^m + \varepsilon_{t-k}) = \lambda^q f_t^q + \varepsilon_t^q \tag{4}$$

if observed, otherwise treated as missing observation. (4) implies an aggregation of the monthly common and idiosyncratic factors to the quarterly frequency where the quarterly factor $f_t^q = \sum_{s=0}^{4} w_s f_{t-s}^m$ for $t = 3k$ and $k = 1, 2 \ldots, K$ where K is the number of quarterly observations. We assume that ε_t^q is white noise at quarterly frequency following Banbura et al. (2013). The final model can be cast into a state-space representation as

$$\begin{aligned} y_t^q &= \lambda^q f_t^q + \varepsilon_t^q \qquad \text{for } t = 3k \text{ and } k = 1, 2 \ldots, K \\ y_t^m &= \lambda^m f_t^m + \varepsilon_t^m \quad \text{for } t = 1, 2, \ldots, T \\ \begin{bmatrix} I_2 & W_t \\ 0 & I \end{bmatrix} \begin{bmatrix} \tilde{f}_t^q \\ f_t^m \end{bmatrix} &= \begin{bmatrix} I_2 & 0 \\ 0 & \phi \end{bmatrix} \begin{bmatrix} \tilde{f}_{t-1}^q \\ f_{t-1}^m \end{bmatrix} + \begin{bmatrix} 0 \\ u_t \end{bmatrix} \end{aligned} \tag{5}$$

where ε_t^q and ε_t^m are following Normal distribution with mean 0 and variances as σ_q^2 and σ_m^2, respectively. W_t includes aggregation weights as given in (3) depending on the time period t. $\tilde{f}_t^q = [f_t^q \ \bar{f}_t^q]'$ has an auxiliary aggregator variable, $\bar{f}_t^q$ and $\tilde{\lambda}^q = [\lambda^q \ 0]$. The auxiliary aggregator is required for the state equation evaluated smoothly, but it is not affecting any of the observed variables. We provide the details of the state-space representations in Section A of the supplementary material. As discussed earlier, the model provided in (5) constitutes the main workhorse for nowcasting the GDP using the dynamic factor models. This model serves as the reference model in our performance evaluation.

Next, we incorporate stochastic volatility to the baseline model for nowcasting the density of GDP growth. For doing this, we allow for the variance of the GDP to evolve over time as

$$h_t \quad = \quad h_{t-3} + \eta_t \ , \quad \eta_t \sim N(0, \sigma_\eta^2) \tag{6}$$

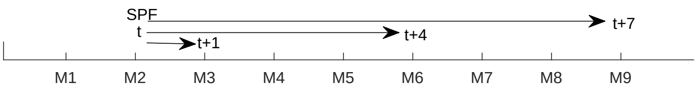
where, $h_t = \log(\sigma_{1,t}^2)$ and $t = 3k,\ k = 1, 2, .., K$ with the $\sigma_{1,t}^2$ being the conditional variance of the GDP in time period t. Alternatively, stochastic volatility could also be specified for the conditional variance of the factors leading to a spanned volatility case, see for example Marcellino et al. (2016). Our motivation for this choice is twofold. First, as we will further exploit the first and second moment of the survey based expectations of the real GDP, we opt for modeling the time variation in the volatility for GDP. Second, modeling the stochastic volatility in the measurement equation rather than the state equation facilitates the inference in such mixed frequency setup with relatively large number of variables.

2.1 Incorporating the density of survey expectations to the baseline statistical model

We extend the baseline statistical models displayed in (5) together with/without stochastic volatility in (6) with the information provided by the survey participants. We do this by using the first and the second moments derived from the expectations of the survey participants. Here, we incorporate the first moment of these expectations to the baseline model without stochastic volatility. In the next part, we proceed to extend the baseline model with stochastic volatility further with the second moment of these expectations as well.

2.1.1 Incorporating the mean

Let $E_t^S[y_{t+h}^q]$ be the expectations of the survey participants of h period ahead GDP growth, where $h = 3k+1$ for $k = 0, 1, 2, 3, 4$. Here S stands for the Survey implying that the expectation is computed using the predictions of the survey participants for the current quarter, $k = 0$, i.e. nowcasts, for the next quarter, $k = 1$, until a year ahead, $k = 4$. We use Survey of Professional Forecasters (SPF) for the survey information about the quarterly GDP, which is released in the second month of the current quarter. That is, $E_t^S[y_{t+h}^q]$ is available for the months $t = 3k - 1$ for $k = 1, 2, \ldots, K$. We illustrate the timeline of the survey data release in the following demonstration.



Here SPF denotes the Survey of Professional Forecasters that is released in the second month of the first quarter of the reference year. The release is constituted by the expectations of the survey participants for the current quarter, which is represented by the short arrow stretched until $t + 1$, expectations for the next quarter, represented by the arrow stretched until $t+4$, and so on. We incorporate the survey based expectations of the GDP to the baseline statistical model by aligning the model based expectations with those derived from the survey. To do that, we first compute the expectations implied by the statistical model. Let $E_t^M[y_{t+h}^q]$ be the $h-$period ahead GDP growth prediction of the model as described in (5), where $h = 3k+1$ for $k = 0,1,2,3,4$. Here M stands for the baseline statistical Model. In line with the timeline of the survey based expectations, we first compute the model based expectations in the second month of the quarter, t, for the current quarter GDP growth, $t+1$, following (4) as

$$\begin{aligned} E_t^M[y_{t+1}^q] &= \lambda^q(E_t^M[f_{t+1}^m] + 2f_t^m + 3f_{t-1}^m + 2f_{t-2}^m + f_{t-3}^m) \\ &= \lambda^q(\phi f_t^m + 2f_t^m + 3f_{t-1}^m + 2f_{t-2}^m + f_{t-3}^m) \end{aligned} \tag{7}$$

where in the second line, the prediction of f_{t+1}^m is computed using the transition equation of the monthly factors that follows an AR(1) process. Similarly, for the forecast of the next quarter which corresponds to the time period $t+4$, the model based expectation can be computed as

$$\begin{aligned} E_t^M[y_{t+4}^q] &= \lambda^q(E_t^M[f_{t+4}^m] + 2E_t^M[f_{t+3}^m] + 3E_t^M[f_{t+2}^m] + 2E_t^M[f_{t+1}^m] + f_t^m) \\ &= \lambda^q(\phi^4 f_t^m + 2\phi^3 f_t^m + 3\phi^2 f_t^m + 2\phi f_t^m + f_t^m). \end{aligned} \tag{8}$$

For the forecast horizons $h = 3k+1$ for $k = 0,1,2,3,4$, i.e. up to four quarter ahead, we can formulate the model based predictions as

$$\begin{aligned} E_t^M[y_{t+h}^q] &= \lambda^q(E_t^M[f_{t+h}^m] + 2E_t^M[f_{t+h-1}^m] + 3E_t^M[f_{t+h-2}^m] + 2E_t^M[f_{t+h-3}^m] + E_t^M[f_{t+h-4}^m]) \\ &= \lambda^q\left(\sum_{k=0}^{4} w_k \phi^{h-k}\right) f_t^m. \end{aligned} \tag{9}$$

The focus of the forecasting exercise is efficiently combining the predictions provided by the statistical model as described in (9) with the predictions provided by the survey participants. We do that by aligning these two sources of predictions by extending the baseline statistical model in (5) with further measurement equations following this aim. Specifically, we extend the model with the following measurements using the dataset provided by the surveys as

$$E_t^S[y_{t+h}^q] = E_t^M[y_{t+h}^q] + \psi_{h,t} \quad \text{with} \quad h = 3k+1 \text{ for } k = 0,1,2,3,4, \tag{10}$$

which leads to the following measurement equations

$$
\begin{aligned}
E_t^S[y_{t+1}^q] &= \lambda^q(\phi f_t^m + 2f_t^m + 3f_{t-1}^m + 2f_{t-2}^m + f_{t-3}^m) + \psi_{1,t} \\
E_t^S[y_{t+h}^q] &= \lambda^q\left(\sum_{k=0}^{4} w_k\phi^{h-k}\right) f_t^m + \psi_{h,t} \quad \text{with} \quad h = 3k+1 \text{ for } k = 1,2,3,4.
\end{aligned}
\tag{11}
$$

(11) implies that the statistical model based expectations should be in line with the expectations obtained from surveys subject to occasional differences or error terms, $\psi_{h,t}$, that follows a normal distribution with variance $\sigma^2_{\psi_h}$. (5) together with the measurement equations (11) constitute the final model where we combine two sources of predictions in a statistically coherent way. For exploring the value-added provided by the survey information depending on the horizon of predictions, we estimate five models where we include, first, only a single measurement equation involving the survey nowcasts, i.e. $k = 0$, and then we add the remaining measurement equations involving the survey forecasts successively for $k = 1, 2, 3, 4$.

Notice that inclusion of the additional measurement equation(s) using survey information potentially changes the estimates of the underlying factors (together with model parameters) and thereby changing the final predictive distribution by changing the mean of the distribution directly. Alternatively, tilting methods that rely on the change of measure could also be conducted. Tilting combines the (moment(s)) of the predictive distribution obtained from the plain statistical model together with the (moment(s)) of the predictive distribution from the additional source of information *ex post*. On the contrary, our approach yields a unique predictive distribution by incorporating the additional source of information into the statistical model structure *ex ante*. We opt for this approach because, first, the state-space structure already allows for a relatively easier extension of the model structure by including additional measurement equations, and second, it leads to considerable computational efficiency.

2.1.2 Incorporating the variance

The variance of the predictions provided by the participants, also called as disagreement, serves as an important proxy of uncertainty, see for example Bomberger (1996) for an earlier analysis. We follow this practice of using disagreement as a measure of uncertainty.

Let $D_{t,h}$ denote the disagreement among the survey participants in period t about the $h-$period ahead GDP growth. This can be computed as

$$D_{t,h} = \frac{1}{N_t}\sum_{i=0}^{N_t}\left(S_{i,t+h} - E_t^S[y_{t+h}^q]\right)^2 \tag{12}$$

where N_t is the number of survey participants in period t performing $h-$step ahead predictions. We align the disagreement computed in each quarter for the $h-$period ahead predictions of the survey participants with the predictions of volatility from the baseline statistical model with stochastic volatility derived using (6). Specifically, assuming a random walk process for the (log-)volatility, $h-$period ahead predictions, $E_t^M[h_{t+h}]$, correspond to the current volatility, h_t. Therefore, we incorporate the disagreement, $D_{t,h}$, computed using the survey information, with the model based predictions as follows

$$\log(D_{t,h}) = h_t + \xi_{t,h} \quad \text{with} \quad h = 3k+1 \text{ for } k = 0, 1, 2, 3, 4. \tag{13}$$

This implies five additional measurement equations that serve for the combination of the (log-) variance obtained from the baseline statistical model together with the (log-) disagreement obtained from surveys. As in the case of the first moment, we estimate five models where we include, first, only a single measurement equation involving the disagreement among survey nowcasts, i.e. $k = 0$, and then we add the remaining measurement equations involving the disagreement among survey forecasts one by one for $k = 1, 2, 3, 4$.

3 Data and Estimation Details

3.1 Data

We consider the US real GDP over the period starting from the last quarter of 1968 until the end of 2019 for the measure of output.[4] This analysis is intended for evaluation of the performance of competing models without including the extreme periods of Covid-19 pandemic in the first two quarters of 2020. For the periods of the so-called 'pandemic recession', we provide a detailed month by month analysis in the next sections. For the

[4]Essentially, the output is measured as Gross National Product (GNP) until 1991 and Gross Domestic Product(GDP) thereafter. Still, the output is denoted as GDP for the whole sample period for the sake of demonstration.

monthly dataset, we construct a broad set of monthly variables involving 16 variables. These variables are employed in Banbura et al. (2013) as well, and thus, it provides ample opportunities to compare the resulting model with the existing popular and successful models which we use as the benchmark.[5] For the survey data, we use the predictions from the Survey of Professional Forecasters (SPF) published by the Federal Reserve Bank of Philadelphia.[6] We provide details on these variables in Table 1.

[Insert Table 1 about here]

In particular, we consider the predictions of the survey participants for the real GDP growth at the individual level. We display the evolution of the mean of these predictions and the disagreement among the forecasters as computed in (12) in Figures 1 and 2.

[Insert Figures 1-2 about here]

Figure 1 shows that the mean of the survey predictions tracks the real GDP growth smoothly with limited variation especially during expansions. This variation further reduces with longer horizon predictions. The decline in the current real GDP growth predictions coincides quite accurately with the actual recession dates, displayed with the grey shaded areas. Note that while these predictions are provided timely, actual real GDP values are released with a lag.

Considering Figure 2, the disagreement before the 1980s lessens considerably after the mid-1980s in line with the notion of great moderation. This refers to the decline of variation in many US macroeconomic series, see McConnell and Perez-Quiros (2000), among others, for details. We observe that the periods before the mid-1980s are comprised of quite erratic swings around high levels of disagreement. On the contrary, the uncertainty is typically quite low after the mid-1980s but it is aggravated almost instantly during turbulent periods which rapidly reverses back afterward. The disagreement in current quarter predictions, i.e. survey-based nowcasts, is typically larger than the corresponding forecasts as can be seen in

[5] The original dataset of Banbura et al. (2013) includes some variables at daily and weekly frequency as well, including return on the market index at the daily frequency and initial jobless claims at the weekly frequency. They note that the daily and weekly variables do not provide further information beyond that of the monthly variables for nowcasting GDP. We follow this practice to exclude the variables at higher than monthly frequency. This also facilitates the computation substantially for the recursive out-of-sample exercise without deteriorating the overall performance of the model.

[6] https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters

many periods of severe recessions of 1982 and 2007. This is because long-horizon predictions tend to follow the long-run level which possesses less variation than the predictions of current values. As the forecasts are provided $h-$periods earlier than the realizations, the nowcasts are provided concurrently, and thus, a shock to the economy has an instant effect on nowcasts which is translated into larger disagreement among these predictions. For a visual representation of survey based nowcasts together with the implied mean and uncertainty, i.e. predictive density for the current quarter GDP growth, we display (kernel approximations of) the distributions over time in the upper panel of Figure 3.

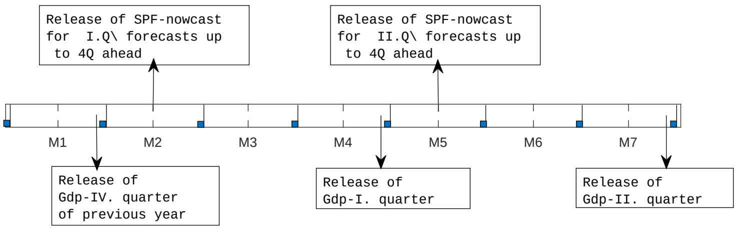
[Insert Figure 3 about here]

Rapid changes in the location of the distributions during recessions together with the changing uncertainty can nicely be traced in the upper panel of Figure 3. The light color indicates the distributions that are thinner representing a limited amount of uncertainty over the periods of the 1990s while the darker colors imply wider distributions indicating the sudden changes in uncertainty during severe recessions as well as during the periods before the mid-1980s.

We perform nowcasts and forecasts at the end of each month recursively. As the GDP is released at the end of the first month following the quarter, this leads us to perform nowcasts rather than backcasts. Moreover, the timing for the release of survey-based predictions was also unclear before 1990. For the periods before the 2000s, survey-based predictions were released towards the end of the second month of the corresponding quarter, while after the 2000s it was released at the end of the second week rather than at the end of the month. Performing the predictions at the end of the month ensures that we do not use information that is not available in real-time.[7] We provide a detailed timeline of release dates with the time of estimation represented in blue dots at the end of months in the following demonstration.[8]

[7]For the periods between the fourth quarter of 1968 and the first quarter of 1990, the survey was conducted by the American Statistical Association (ASA) and the National Bureau of Economic Research (NBER). Following 1990, the Federal Reserve Bank of Philadelphia has taken over the conduct of surveys. For the release dates of the survey information after the second quarter of 1990, see `https://www.philadelphiafed.org/-/media/research-and-data/real-time-center/survey-of-professional-forecasters/spf-release-dates.txt?la=en`

[8]These release dates correspond to the periods after the first quarter of 2005.



We first conduct a full sample estimation for the analysis of the models. Next, we perform a recursive out-of-sample analysis in pseudo-real-time using the ragged edge datasets due to publication delays for evaluation of the model performance with/out survey information.

3.2 Estimation Procedure

We use Bayesian inference using Markov Chain Monte Carlo (MCMC) simulation techniques for estimation and inference in the unobserved component model and its variants. Specifically, we use Metropolis within Gibbs sampling together with data augmentation (see Tanner and Wong, 1987) to obtain posterior results. While the baseline statistical model allows for using plain MCMC involving only Gibbs sampling using standard conditional distributions, the fact that incorporating survey information leads to nonlinear parameter structure as in (11). We employ the Metropolis-Hastings algorithm in such cases. The posterior distribution for the model parameters is proportional to the product of the likelihood function and the prior distributions for model parameters. While for the likelihood specification we make use of the multivariate Normal distribution following the convention, for the prior specifications of the model parameters, we specify uninformative priors. Details on the prior specifications and the resulting posterior distributions along with the simulation scheme are provided in Appendix A.

3.3 Evaluation of Models

We compare the competing models with and without survey information to uncover the value-added provided by the survey information on top of the baseline statistical models. Therefore, we use various measures of model comparison for evaluation of the model performance with a special focus on density prediction. First, we use the marginal likelihood metric to explore the performance of models using the full sample. A drawback of marginal likelihoods is its sensitivity to the specifications of prior distributions. Even though we use uninformative priors,[9] we still want to support our comparison with a more robust metric of model evaluation. An alternative way of model comparison is employing predictive Bayes factors computed using predictive likelihoods. This approach has the advantage that it is not affected by the choice of prior distributions and overfitting, while it is directly related to the posterior model probabilities. For a given model, the predictive likelihood of the observation at $t_0 + 1$, y_{t_0+1}, conditional on previous observations y^{t_0}, is given by

$$f(y_{t_0+1}|y^{t_0}) = \int f(y_{t_0+1}|\theta)f(\theta|y^{t_0})d\theta, \tag{14}$$

where $p(\theta|y^{t_0})$ is the posterior distribution of the model parameters, estimated using the data until t_0, gathered in the parameter set, θ, given the observations until t_0. $p(y_{t_0+1}|\theta)$ is the density of the observation y_{t_0+1}, which can be written as

$$f(y_{t_0+1}|\theta) = \int\limits_{f_{t_0+1}} f(y_{t_0+1}|f_{t_0+1}, \theta)f(f_{t_0+1}|\theta, y^{t_0}). \tag{15}$$

We can use the posterior simulator to obtain the distribution of the model parameters and estimate the predictive likelihood by $G^{-1}\sum_{g=1}^{G} f(y_{t_0+1}|y^{t_0}, \theta^{(g)})$, where G is a large number of draws from the posterior distribution.

Finally, we use probability integral transforms (PIT) to explore whether the predicted distributions can capture the real GDP growth density in various percentiles. The PIT score is defined as

$$z_t = \int\limits_{-\infty}^{y_t} f(u|\theta, y^{t_0})du. \tag{16}$$

The intuition behind the use of PITs relies on the observation that the closer the predictive distributions to the true data generating densities, the closer these PITs, z_t, to be distributed according to a uniform distribution, see Diebold et al. (1998) for details. While the data spans the periods over the last month (last quarter) of 1968 until the end of 2019, we consider the period starting from the first month (first quarter) of 1977 until the end of

[9] In this case, marginal likelihood based comparisons might tend to favor more parsimonious models, because the marginal likelihood is computed by integrating out the entire parameter space covering insensible regions of parameter space due to improper prior specifications.

the sample as the evaluation period to compute the predictive likelihoods and PITs.[10]

4 Empirical Findings

In this section, we discuss our empirical findings based on the estimation results of various models. We first estimate the baseline statistical model as described in (5), denoted as **'BM'** referring to the Baseline Model. We extend this model using the mean predictions obtained by SPF each time adding k-quarter ahead predictions for $k = 0, 1, \ldots, 4$ as described in (10). We denote these models as **'(BM+S)-Mk'** for $k = 0, 1, 2, 3, 4$ referring to Baseline Model and Survey information with Mean aligned using up to $k^{th}-$quarter ahead predictions. Comparison of the BM with those extended using the first moment of survey-based predictive distributions would indicate, first, whether survey-based predictions of mean bear additional information and second, whether this information is embedded only in survey-based nowcasts or it is also carried over in forecasts at various horizons.

Next, we extend the BM with the stochastic volatility using the specification described in (6). This implies that we conduct a thorough density estimation with time variation in both the first and second moments. This model is denoted as **'BMSV'**. We extend this model using the mean predictions obtained by SPF as in the previous case. We denote these models as **'(BMSV+S)-Mk'** for $k = 0, 1, 2, 3, 4$. Finally, we further extend this model using the measure of disagreements among the survey participants obtained by SPF as described in (11). This implies that the last group of models exploits both the first and second moments of the predictive distributions from SPF essentially combining these with the predictive distributions obtained from the baseline statistical model. We denote these models as **'(BMSV+S)-MVk'** for $k = 0, 1, 2, 3, 4$.

4.1 Full sample findings

We start by evaluating the main estimation results of selected competing models. For visual inspection of the main findings in conforming stylized facts, Figure 4 displays the

[10] The out-of-sample analysis employs pseudo-real-time datasets in the sense that these involve revised data. We also perform the analysis for the GDP series that is based on the second release as in Faust and Wright (2009) and Carriero et al. (2015). Results are very similar to those reported in the next section and available upon request.

predicted mean, i.e. the fitted values obtained using the BMSV model.

[Insert Figure 4 about here]

As can be seen from Figure 4 the model performs very well in tracking the mean real GDP growth rate throughout the sample periods. The timing of recessions as well as recovery periods are captured by the statistical model successfully. Here we do not display the fitted values obtained from (BMSV+S) type of models for the clarity of the demonstration as these are very similar to those displayed in Figure 4.

Figure 5 displays the predicted volatility using the BMSV model. Here we also include the estimates obtained from the (BMSV+S)-MV4 that combines both the mean and the disagreements obtained from the SPF together with the BMSV model for comparison.

[Insert Figure 5 about here]

The estimation results imply that the volatile periods of the 1970s and early 1980s are followed by a rapid decline in volatility in the mid-1980s consistent with the great moderation. We observe increases in volatility during recessions in the 2000s, albeit much limited compared to the levels of volatility before great moderation. One important aspect of the volatility estimated by the BMSV model is that it evolves smoothly over time. On the contrary, volatility estimates obtained by the (BMSV+S)-MV4 model involve more variation. It seems that (BMSV+S)-MV4 model provides a compromise between the smoothly evolving volatility obtained by the BMSV model and the rapidly changing nature of ambiguity captured by the disagreement of participants of SPF. Figure 4 and 5 indicate that the main competing models can capture the moments of the real GDP growth quite successfully.

Next, we consider the model fit based on the marginal likelihood metric computed using the full sample for all competing models. For the ease of model comparison, we provide the (log-)marginal likelihood values for the baseline statistical model (BM), and for the remaining models, we display the differences between the (log-)marginal likelihood of the competing models with that of the BM. These correspond to the Bayes factors of competing models with respect to BM. We display these in the first three columns of Table 2. Different groups of models are given in rows with varying shades of grey for the ease of demonstration.

[Insert Table 2 about here]

When we consider the baseline model in the first row, it is seen that the data flow does not matter much. The (log-)marginal likelihood values are around -227 regardless of the month of the quarter the predictions are performed as can be seen in the first row of Table 2. When these predictions are aligned with the first moment of survey-based predictions displayed in the next panel with darker grey shade, we observe that survey-based nowcasts, i.e. predictions of the current quarter by the survey participants, do not cause any improvement in the marginal likelihoods. However, this picture reverses when we incorporate survey forecasts, i.e. predictions of next quarters by the survey participants. In this case, marginal likelihood values increase by around 9 points. Thus, the mean of survey-based predictions of the future periods indeed possesses additional information.

When the two baseline statistical models with and without stochastic volatility are compared, allowing for time variation in volatility deteriorates the model fit as can be seen from the negative Bayes factor values corresponding to the BMSV model. Once again, when this model is extended using the alignment with the mean of survey-based predictions of future periods we observe positive Bayes factors by around 10 points. This shows that the survey participants' predictions of future periods rather than the current period bear indeed additional information beyond that is contained in the extracted factor and/or volatility.

The largest improvement is obtained when survey-based predictions are incorporated both for the first and second moments. The last panel with the darkest shade of grey indicates that not only the disagreements among the survey participants on future periods' expectations matter but the disagreement on nowcasts also provides a significant improvement of marginal likelihood values. When the first and second moments of the predictive distributions implied by the baseline model are aligned with those of the survey-based nowcast distribution as it is the case for the (BMSV+S)-MV0 model, we observe an increase around 7 points over the conventional baseline statistical model and an increase around 6 points over the baseline statistical model with stochastic volatility. When we incorporate further the future predictions based on the survey both in terms of the predicted mean and the disagreement among the survey participants, marginal likelihood values increase by 22 points. These results decisively indicate that survey-based predictions of the full density of the real GDP growth embed pivotal information over the conventional baseline statistical models.

4.2 Predictive performance

While the marginal likelihood metric-based evaluation of the models relies on the dataset over the full sample period, it is crucial to examine the performance of the models in real-time using the information available at the time of prediction. Therefore, in this section, we evaluate the predictive performance of the models based on predictive likelihoods.

Predictive likelihood values are displayed in Table 2 starting from the fourth column. The fourth to sixth columns under the column heading 'Q0' present the predictive likelihood values computed using nowcasts performed at the end of the first, second, and third months of the current quarter, respectively. When we focus on the baseline statistical model, the effects of data flow can explicitly be seen in real-time, as the amount of increase in predictive likelihoods over the course of the quarter is as high as 22 points. This increase in the predictive likelihood values can be seen for all cases independent of whether we allow for time variation in volatility and whether survey information is added or not.

Incorporating the mean nowcasts from the survey predictions adds on top of the baseline statistical model with increases in predictive likelihoods of around 3 points for the first two months of the quarter as can be seen from the row corresponding the (BM+S)-M0 model. The impact of the survey-based predictions on the baseline model increases further when we incorporate survey-based forecasts with the amount of increase up to 6 points. This is particularly the case in the second month of the current quarter when the survey information is released. Hence, incorporating survey-based forecasts of mean in addition to nowcasts of it enhances the improvement of the predictive capability of the baseline model further.

In the previous section when we evaluate full sample results, we see that incorporating stochastic volatility to the baseline model deteriorates the model fit. This picture reverses when we evaluate models in real-time using predictive likelihoods. In fact, allowing for stochastic volatility in the baseline statistical model improves predictive likelihood by around 14 to 19 points when the row corresponding to BMSV is compared to that of BM. When the survey-based nowcasts of the mean are incorporated, we observe an improvement of 3 points. Once again, this rises to around 5 points with the incorporation of survey-based forecasts on the first moment of the predictive distribution of the real GDP growth.

The impact of the survey information on the predictive ability of the baseline statistical model soars when we also incorporate the disagreement among the survey participants as a measure of the second moment, i.e. volatility of the predictive distribution. In this case, incorporating only the nowcasts of survey-based predictive distribution including the first and the second moment, increase the predictive likelihoods for nowcasts by around 6 points. When survey-based predictive distributions using the survey participants' forecasts are incorporated, in addition to nowcasts of them, this gain increases to 8 points compared to the baseline statistical model with stochastic volatility where no use is made of survey information.

To summarize, when the (BMSV+S)-MV4 model, where we incorporate all available survey information on the full distribution of the real GDP, is compared to the plain BM model, the improvement in predictive likelihoods are as high as 26 points. For a visual representation of the real-time nowcasts obtained by the BM and (BMSV+S)-MV4 models, we display (kernel approximations of) the distributions, that is the predictive densities for the current quarter real GDP growth, obtained from these models over time in Figure 3. We display those for the BM model in the middle panel and those for the (BMSV+S)-MV4 model in the bottom panel of Figure 3. Comparison of the distributions provided by the (BMSV+S)-MV4 model with that of the survey-based nowcasts in the upper panel Figure 3 and with those obtained from the BM model reveals these predictive gains in nowcasts clearly. On the one hand, compared to the SPF, predictions offered by (BMSV+S)-MV4 model provide more smoothly evolving predictive distributions with relatively wider distributions during expansions. These distributions become thinner when we consider recessions mitigating the excessive uncertainty in the predictions of SPF. On the other hand, compared to the BM model, predictive distributions offered by (BMSV+S)-MV4 model have lighter colors indicating that these have smaller variance and therefore they are gathered densely around the mean of the nowcasts. This improves the predictive likelihood values considerably as the uncertainty around the real GDP growth is further resolved when survey nowcasts are combined with the baseline statistical model. These findings suggest that survey-based predictions deliver important and genuine information that is not contained in comprehensive datasets as ours including many forward-looking variables such as PMI and confidence indices beyond conventional indicators.

Next, we evaluate the forecasting performance of the model by focusing on 1- and 2-quarter ahead forecasts in the last six columns of Table 2. Considering the impact of the data flow on the 1-quarter ahead forecasts, we observe a sizable effect in the sense that predictive likelihood values improve by around 8 points when predictions are performed in the third month of the current quarter, M3, compared to predictions performed in the first month of the current quarter, M1. This impact of the data flow vanishes though when we consider 2-quarter ahead forecasts. In this case, the differences in predictive likelihood values drop to values around 3 points. Therefore, we conclude that, while the impact of the data flow over the course of the current quarter is largest when we consider nowcasting, this impact fades smoothly out with the increasing forecast horizon.

Regarding the impact of the survey-based information on forecasting, it is seen that in almost all of the cases, the information in survey nowcasts is sufficient to improve the predictive ability of the statistical model. In terms of the first moment, survey-based forecasts do add partly on top of the gains obtained by survey-based nowcasts. To give an example, enhancing the baseline statistical model, 'BM', by incorporating the mean extracted from survey-based nowcasts, (BM+S)-M0, escalates the predictive likelihood value around 5 points. However, incorporation of the first moments from survey-based forecasts, as it is the case for the model (BM+S)-M4, improves the likelihood value around 2 points for 1-quarter ahead predictions performed in the second month of the current quarter. As in the case of nowcasting, incorporating stochastic volatility to the baseline model, 'BMSV', enhances the predictive ability of the statistical model for both 1- and 2-quarter ahead forecasts. More importantly, aligning the first moment of the predictive distribution extracted from survey nowcasts together with that of the baseline statistical model with stochastic volatility leads to an increase in the predictive likelihood by almost 5 points for both 1- and 2-quarter ahead forecasts. Aligning the second moment from the survey information further improves predictive capability by an additional 3 points on average.

4.3 Evolution of predictive gains from survey information over time

The findings displayed in the previous section show that survey enhanced model predictions possess useful information enhancing the predictive ability of the baseline statistical model both in terms of nowcasts and forecasts. In this section, we focus on the evolution of these predictive gains to explore further whether survey information has a particular pattern over time in improving the predictive capability of the baseline model. Figure 6 displays the evolution of the Bayes factors which are computed as the differences of predictive likelihoods (computed recursively) between the competing models and the BM model. The predictive likelihoods are computed using (15) with the predictions performed in the second and third months of the current quarter, i.e. using nowcasts of the current quarter.[11] Moreover, for the models that use survey information we only include those models that incorporate all available survey information including nowcasts and up to 4-quarter ahead forecasts.

[Insert Figure 6 about here]

Before the analysis of the contribution of survey information throughout the evaluation period, we consider the evolution of the Bayes factor for the selected models at the onset of our evaluation sample until the mid-1980s. During these periods, all models but the BMSV model perform worse than the benchmark of the BM model. On the contrary, the outperformance of the survey enhanced models prevails for the period after the mid-1980s until the end of the sample. Nevertheless, this deterioration rapidly vanishes for the nowcasts of the third month of the quarter with the extension of the dataset with further data releases.

Regarding the impact of the survey information, we start to evaluate the dynamics of the value-added of incorporating the survey-based mean predictions to the statistical model. Therefore, we focus on the evolution of the Bayes factor for the (BM+S)-M4, relative to the plain baseline model, BM, as shown using the dashed line in Figure 6 for the nowcasts. We see that the Bayes factor evolves smoothly favoring the (BM+S)-M4, almost like a straight line. While the slope of the line, representing the contribution to the Bayes factor in each period, is steeper for the nowcasts performed in the second month, it gets dampened for

[11] We display the graph of the evolution of predictive likelihoods for predictions performed in the first month of the current quarter in Appendix B as it is very similar to the Figure 6.

the nowcasts performed in the third month. Some sudden changes in Bayes factors can be traced including an increase around the recession of 1990-91 and 2003 for the nowcasts performed in the second month. The increase in the Bayes factor in 1990-91 is related to the timely prediction of the short-lasting mild recession by the survey participants relative to the statistical model. The sudden increase in Bayes factor around 2003 corresponds to the start of a relatively faster growth path of US real GDP which seems to be anticipated by the survey participants swiftly relative to the statistical model leading to this improved predictive ability. Therefore, while the survey-based predictions of the real GDP are mostly aligned with that of the statistical model, these provide valuable information beyond that is contained in the statistical model during times of abrupt changes.

Considering the impact of time variation in volatility, we include the evolution of the Bayes factor of the baseline model with stochastic volatility, BMSV, with respect to the plain baseline model, BM, displayed using the dotted line. This impact can clearly be seen after the mid-1980s with the start of great moderation when the variation in real GDP growth rates has reduced considerably. While the BMSV model can accommodate this reduction throughout the 1980s and 1990s the evolution of the Bayes factor is dampened in the 2000s. Indeed, while the Bayes factor is around 10 in 2000 it increased only to 14 towards the end of the sample, with a large part of this increase taking place after 2017. When we focus on the model where the (mean) survey predictions of the real GDP growth are incorporated into the BMSV model, i.e. (BMSV+S)-M4, we observe that it performs constantly better than the BMSV model starting from the mid-1980s until the end of the sample. The inclusion of the survey information in terms of the first moment of the predictive distribution leads to some swings in the evolution of the Bayes factors as displayed using the dashed-dotted line. Strikingly, the improvement of the Bayes factors does not get dampened out towards the end of the sample contrary to the BMSV model.

Finally, when we incorporate the full predictive density in terms of the first and second moments from the survey information with the statistical model as is the case for the (BMSV+S)-MV4 represented using the solid line, we see that this model is dominant over all competing models throughout the evaluation period. We can track rapid and long-lasting increases throughout the evaluation period interrupted with rapid but short-lasting decreases mostly coinciding with depressive periods such as 2008, 2011, and 2013. These periods correspond to the collapse of Lehman Brothers, the European debt crisis, and the taper tantrum, respectively. Inclusion of the second moment of predictive density extracted from the survey information leads to more visible and potent swings in Bayes factors. Hence, the impact on predictive performance due to survey information follows a sentimental wave. During turmoil periods, excessively increasing volatility deteriorates the 'density' prediction rapidly which lasts very short. On the contrary, this deterioration is followed with a long-lasting improvement after a swift adjustment during good times following turmoil periods. Therefore, in total, it leads to a large improvement in the overall performance of survey information augmented statistical models compared to the plain statistical model.

4.4 Evaluation of the predictive performance using PITs

So far, we analyze the model fit and the predictive performance of the models in predicting the density of the real GDP using likelihood-based metrics that summarize the overall performance. In this section, we focus on the performance of the models in predicting specific percentiles of the distribution. Therefore, we compute the probability integral transforms (PITs) corresponding to the deciles of the real GDP growth sample for predictive distributions obtained from the competing models as described in (16). If the predictive distribution successfully approaches the true density, then the PITs computed for the deciles should be close 0.1 as in this case they are distributed according to a uniform distribution. We also include the credibility sets (of two standard deviations) of these PITs using horizontal dashed-dotted lines. These are displayed in Figure 7.

[Insert Figure 7 about here]

Figure 7 indicates the importance of incorporating survey information to the baseline statistical model. When we consider the baseline model with constant volatility, the PITs clearly do not follow a uniform distribution with PITs of the deciles in the middle of the distribution exceeding the value 0.1. On the contrary, PITs of the deciles at the tails are much lower than 0.1 below the lower bound. Allowing for time variation in volatility leads to some minor improvement of PITs. Still, the BMSV model displays the behavior of

PITs similar to the BM model. When we incorporate the first moment extracted from the survey information to the plain baseline model with constant volatility, we observe some moderation in PITs but it still fails for the PITs to follow a uniform distribution. However, when we incorporate the first moment extracted from the survey information to the BMSV model, we observe a major improvement. Nevertheless, the PITs of the fifth and eighth deciles approach to the upper bound and the PITs of the deciles in the tails approach to the lower bound. Finally, when we incorporate both the first and second moment of the predictive distributions from survey-based predictions to the BMSV model, the distribution of the PITs are quite close to the uniform distribution with almost all deciles being clearly inside the credible sets. Hence, we conclude that, first, allowing for time variation in volatility is crucial for density prediction and the resulting PIT based evaluation, second, incorporating survey information, most notably the second moment extracted from the survey, improves the density nowcast further throughout all deciles of the density.

5 A closer look on the predictive performance of the models during the Covid-19 pandemic

The outbreak of the novel coronavirus, Covid-19, has led to a dramatic health crisis. The spread of the virus has been at an unexpected pace since the burst of the pandemic in early January of 2020. Several countries including the US have taken measures to contain the pandemic. These measures include partial or complete closure of various businesses leading to a devastating supply shock, see Alvarez et al. (2020); Acemoglu et al. (2020). On top of that, the pandemic has altered the daily routines substantially with a fundamental change in preferences, which leads to a sizeable demand shock, see Çakmaklı et al. (2020); Eichenbaum et al. (2020). Note that, our results in previous sections are based on the dataset that excludes the first six months of 2020. Here, we would like to have a closer look at the predictive ability of the competing models during this period. We display the mean of survey participants' predictions and the (square root of) disagreement among the forecasters as computed in (12) that includes the releases up to the second quarter of 2020 in the upper panel of Table 3

[Insert Table 3 about here]

We further display the kernel approximation of the predicted distributions obtained using a selected set of competing models, in Figure 8. Each of the predicted distributions of the quarterly real GDP growth is generated using the data available at the end of each month from January until June 2020 and the predictions are performed for the corresponding quarters of the months. Therefore, these are nowcasted real GDP growth densities for the first two quarters of 2020. We also display the actual realization of the real GDP growth using the most recent release by a vertical line. For a comparison of our predictions with popular publicly available nowcasts, we further include the **point** nowcasts provided by the Federal Reserve Banks of New York and Atlanta displayed in Table 3.

[Insert Figure 8 about here]

First, we focus on the nowcasted densities of the first-quarter GDP growth using available data at the end of January, February, and March. The first three graphs in Figure 8 display these nowcasted densities obtained using selected competing models. The actual growth rate turned out to be -5.0% due to the extensive pandemic measures that become effective after mid-March. Therefore, the data that is available at the end of March still did not contain the devastating effects of the pandemic. Accordingly, as of February 14, which is the release date of SPF, the effects of the pandemic were still unforeseen. Therefore, nowcasts based on SPF are centered around 1.7% with a small standard deviation. On the other hand, the predictions obtained by the BMSV range between 2.2% and 2.4% for the successive months incompetent with the actual realization. Combining these two sources of information, the predictions provided by the (BMSV+S)-MV4 model varies between 2.1% and 2.2% over the course of the quarter which is marginally smaller than those of the BMSV model reflecting the effects of the SPF release. One minor difference is on the probability of a negative growth rate, which can be computed using the corresponding area under the curves. Reflecting the skewed density provided by the SPF, this probability is slightly larger for the (BMSV+S)-MV4 model compared to the BMSV model. Finally, end of the month predictions exhibit a similar picture for the New York and Atlanta Feds' nowcasts with the latter being slightly greater. Therefore, regardless of the source, these predictions almost uniformly fail to nowcast the economic downturn of the first quarter of 2020 mainly caused by the pandemic measures utilized in the last two weeks of March.

The picture becomes dramatic when we consider the second quarter of 2020. The uncertainty brought by the devastating shock due to the Covid-19 pandemic is overwhelming. The actual growth rate for the second quarter of 2020 turned out to be as large as -31.4% according to the third estimate by the Bureau of Economic Analysis. Considering the end of April estimates, we observe that the baseline model BMSV anticipates the downturn to some extent but the magnitude of the contraction is still profoundly far from the actual rate with the predicted mean of -11.2%. This is in parallel with the predictions of New York and Atlanta Feds that are -7.8% and -12.1%, respectively. Notice that, SPF release is still not available as of April. Therefore, when predictions of the (BMSV+S)-MV4 are performed, only 1-quarter ahead predictions of the February 14 release are used for the survey information on the second quarter of 2020 real GDP growth that is 2.1%. Consequently, the mean nowcast for the (BMSV+S)-MV4 is -7.5% that is close to the nowcast of the New York Fed.

The predictions of the end of May are quite decisive for almost all of the sources. First, the May 15 release of the SPF updates the predictions for the second quarter of 2020 real GDP growth substantially to -31.5% together with an extreme level of disagreement. In particular, participants of the SPF provide predictions ranging from -10% at the most optimist side to -50.2% at the most pessimist side as can be seen in the fifth graph of Figure 8. This sizable amount of uncertainty is also reflected in the nowcasts of New York and Atlanta Feds with the diverging predictions of -35.5% and -51.2%, respectively. The predictive distribution provided by the BMSV model has a mean of -33.2% indicating the accuracy of the BMSV model together with a low volatility level of 1.3%. Quite closely, combined with the survey information, the predictive distribution of the (BMSV+S)-MV4 model is centered around the mean of -35.2% with a higher level of volatility of 1.8% compared to the BMSV model due to the substantial disagreement in the SPF May 15 release. As a result, both distributions include the actual value inside the 95% credibility sets.

Finally, the end of June estimates exhibit another substantial shift in the location of many of the predictions. Specifically, the efforts for the reopening of business are reflected in the observations of various variables in May, thereby to the predictions performed using available data as of the end of June. This reversal can also be seen in the predictions of New York and Atlanta Feds with substantial revisions to -16.3% and -39.5%, albeit still divergent. A similar revision can also be observed for the predictive distribution obtained by the BMSV model with the distribution centered around the mean of -25.2% with a volatility of 1.3%. The prediction of the (BMSV+S)-MV4, however, have a similar predictive distribution as in the previous month, thanks to the survey information, with a distribution centered around the mean of -31.3% and with a volatility of 1.8%. Therefore, (BMSV+S)-MV4 leads to the most accurate prediction of this extreme downturn with the magnitude of contraction of -31.7%, by combining the survey information with the baseline statistical model.

This demonstration represents the underlying drivers of the findings on the predictive ability of our model. On the one hand, survey-based predictions of the first moment are mostly aligned with those of the baseline statistical model. However, it embeds relevant information in particular for extreme cases. On the other hand, survey-based predictions of the second moment react swiftly to the changing uncertainty. This leads to a more accurate adjustment of the predictive distribution obtained using the combined model to the new conditions compared to the conventional statistical model.

6 Conclusion

We propose an econometric model where we incorporate the survey-based information to the conventional dynamic factor model (together with stochastic volatility) used for nowcasting the US real GDP in a statistically coherent way. Our model effectively combines the predictive distributions of the real GDP implied by the predictions of survey participants with the predictive distributions implied by the statistical model. We do this by aligning the implied first and second moments of those predictive distributions from these two sources of information. The resulting approach fits naturally to the state-space structure sidestepping the need for the tilting methods used in similar setups.

Our model produces survey consistent measures of output growth expectations and accompanying time-varying uncertainty. We use the output projections for different horizons from the Survey of Professional Forecasters (SPF). We provide results on the accuracy of nowcasts and short-term forecasts of U.S. real GDP growth in a real-time exercise with the evaluation period starting from the first quarter of 1977 until the end of 2019. A comparison of different specifications through the predictive likelihoods and PITs reveals the importance of the survey information in predicting the density of the US real GDP. A month by month analysis in the first six months of 2020, when the economic shock generated by the Covid-19 pandemic hit the US economy deeply, confirms the outperformance of the survey enhanced model in nowcasting the density of the US real GDP.

Our model combines a competitive baseline statistical model with the survey information provided by the SPF. Although these forecasts are publicly available, they are conducted at a quarterly frequency that leads to some informational delays. For example, the first quarter of 2020 survey nowcasts (released on February 14) still do not reflect the effects of the Covid-19 pandemic, that hit the economy extensively starting from mid-March. Therefore, a monthly survey would potentially increase the accuracy of our proposed framework further. One such important source of survey information is Blue Chip Publications that provide monthly predictions but which is not publicly available as opposed to SPF.

References

Aastveit KA, Gerdrup KR, Jore AS, Thorsrud LA. 2014. Nowcasting gdp in real time: A density combination approach. *Journal of Business & Economic Statistics* **32**: 48–68.

Acemoglu D, Chernozhukov V, Werning I, Whinston MD. 2020. A multi-risk SIR model with optimally targeted lockdown. Working Paper 27102, National Bureau of Economic Research.

Çakmaklı C, Demiralp S, Kalemli-Özcan e, Yesiltas S, Yildirim MA. 2020. Covid-19 and emerging markets: An epidemiological model with international production networks and capital flows. Working Paper 27191, National Bureau of Economic Research.

Altavilla C, Giacomini R, Ragusa G. 2017. Anchoring the yield curve using survey expectations. *Journal of Applied Econometrics* **32**: 1055–1068.

Altug S, Çakmaklı C. 2016. Forecasting inflation using survey expectations and target inflation: Evidence for brazil and turkey. *International Journal of Forecasting* **32**: 138 – 153.

Alvarez FE, Argente D, Lippi F. 2020. A simple planning problem for COVID-19 lockdown. Working Paper 26981, National Bureau of Economic Research.

Ang A, Bekaert G, Wei M. 2007. Do macro variables, asset markets, or surveys forecast inflation better? *Journal of monetary Economics* **54**: 1163–1212.

Aruoba BS. 2020. Term structures of inflation expectations and real interest rates. *Journal of Business & Economic Statistics* **38**: 542–553.

Banbura M, Giannone D, Modugno M, Reichlin L. 2013. Now-casting and the real-time data flow. *Handbook of economic forecasting* **2**: 195–237.

Bańbura M, Modugno M. 2014. Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. *Journal of Applied Econometrics* **29**: 133–160.

Bloom N. 2014. Fluctuations in uncertainty. *Journal of Economic Perspectives* **28**: 153–76.

Bok B, Caratelli D, Giannone D, Sbordone AM, Tambalotti A. 2018. Macroeconomic nowcasting and forecasting with big data. *Annual Review of Economics* **10**: 615–643.

Bomberger WA. 1996. Disagreement as a measure of uncertainty. *Journal of Money, Credit and Banking* **28**: 381–392.

Campbell SD. 2007. Macroeconomic volatility, predictability, and uncertainty in the great moderation: evidence from the survey of professional forecasters. *Journal of Business & Economic Statistics* **25**: 191–200.

Carriero A, Clark TE, Marcellino M. 2015. Realtime nowcasting with a bayesian mixed frequency model with stochastic volatility. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **178**: 837–862.

Carter C, , Kohn R. 1994. On Gibbs Sampling for State Space Models. *Biometrika* **81**: 541–553.

Creal D. 2012. A survey of sequential monte carlo methods for economics and finance. *Econometric Reviews* **31**: 245–296.

D'agostino A, McQuinn K, Whelan K. 2012. Are some forecasters really better than others? *Journal of Money, Credit and Banking* **44**: 715–732.

Diebold FX, Gunther A, A T, Tay AS. 1998. Evaluating density forecasts with applications to financial risk management. *International Economic Review* **39**: 863–883.

Durbin J, Koopman S. 2012. *Time Series Analysis by State Space Methods: Second Edition*. Oxford Statistical Science Series. OUP Oxford. ISBN 9780199641178.

Eichenbaum MS, Rebelo S, Trabandt M. 2020. The macroeconomics of epidemics. Working Paper 26882, National Bureau of Economic Research.

Faust J, Wright JH. 2009. Comparing greenbook and reduced form forecasts using a large realtime dataset. *Journal of Business & Economic Statistics* **27**: 468–479.

Faust J, Wright JH. 2013. Forecasting inflation. In *Handbook of economic forecasting*, volume 2. Elsevier, 2–56.

Foroni C, Ghysels E, Marcellino M. 2013. Mixed-frequency vector autoregressive models. *Advances in econometrics* **32**: 247 – 272.

Frühwirth-Schnatter S. 1994. Data augmentation and dynamic linear models. *Journal of Time Series Analysis* **15**: 183–202.

Geisser S. 1965. A bayes approach for combining correlated estimates. *Journal of the American Statistical Association* : 602–607.

Giannone D, Reichlin L, Simonelli S. 2009. Nowcasting euro area economic activity in real time: the role of confidence indicators. *National Institute Economic Review* **210**: 90–97.

Giordani P, Söderlind P. 2003. Inflation forecast uncertainty. *European Economic Review* **47**: 1037 – 1059.

Grischenko O, Mouabbi S, Renne JP. 2019. Measuring inflation anchoring and uncertainty: A u.s. and euro area comparison. *Journal of Money, Credit and Banking* **51**: 1053–1096.

Kim CJ, Nelson CR, et al. 1999. State-space models with regime switching: classical and gibbs-sampling approaches with applications. *MIT Press Books* **1**.

Koop GM. 2003. *Bayesian econometrics*. John Wiley & Sons Inc.

Kozicki S, Tinsley P. 2006. Survey-based Estimates of the Term Structure of Expected US Inflation. Working Paper Series 46, Bank of Canada.

Kozicki S, Tinsley P. 2012. Effective use of survey information in estimating the evolution of expected inflation. *Journal of Money, Credit and Banking* **44**: 145–169.

Krüger F, Clark TE, Ravazzolo F. 2017. Using entropic tilting to combine bvar forecasts with external nowcasts. *Journal of Business & Economic Statistics* **35**: 470–485.

Lahiri K, Sheng X. 2010. Measuring forecast uncertainty by disagreement: The missing link. *Journal of Applied Econometrics* **25**: 514–538.

Marcellino M, Porqueddu M, Venditti F. 2016. Short-term gdp forecasting with a mixed-frequency dynamic factor model with stochastic volatility. *Journal of Business & Economic Statistics* **34**: 118–127.

Mariano RS, Murasawa Y. 2003. A new coincident index of business cycles based on monthly and quarterly series. *Journal of Applied Econometrics* **18**: 427–443.

McConnell MM, Perez-Quiros G. 2000. Output fluctuations in the united states: What has changed since the early 1980's? *American Economic Review* **90**: 1464–1476.

Omori Y, Chib S, Shephard N, Nakajima J. 2007. Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics* **140**: 425–449.

Schumacher C. 2016. A comparison of midas and bridge equations. *International Journal of Forecasting* **32**: 257 – 270.

Tallman EW, Zaman S. 2019. Combining survey long-run forecasts and nowcasts with bvar forecasts using relative entropy. *International Journal of Forecasting* ISSN 0169-2070.

Tanner MA, Wong WH. 1987. The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association* **82**: 528–550.

Table 1: Dataset: Frequencies, transformations, periods and publication delays

No	Series	F	T	Start	Delay
1	Real Gross Domestic Product	Q	1	1968:IV	1M
2	Industrial Production Index	M	1	1968:12	1M
3	Purchasing Manager Index: Manufacturing	M	2	1968:12	1M
4	Real Disposable Personal Income	M	1	1968:12	1M
5	Unemployment Rate	M	2	1970: 1	1M
6	All Employees: Total Nonfarm Payrolls	M	1	1968:12	1M
7	Personal Consumption Expenditures	M	1	1968:12	1M
8	Housing Starts: Total: New Privately Owned Housing	M	1	1968:12	1M
9	New One Family Houses Sold	M	1	1968:12	1M
10	Manufacturers' New Orders: Durable Goods	M	1	1992: 2	1M
11	Producer Price Index: Finished Goods	M	1	1968:12	1M
12	Consumer Price Index for All Urban Consumers	M	1	1968:12	1M
13	Imports of Goods and Services	M	1	1992: 1	2M
14	Exports of Goods and Services	M	1	1992: 1	2M
15	Philadelphia Fed Survey: General Business Conditions	M	2	1968:12	-
16	Retail Sales: Retail and Food Services	M	1	1992: 1	1M
17	Conference Board Consumer Confidence	M	2	1968:12	-
	Survey of Professional Forecasters				
	Real Gross Domestic Product	Q	1	1968:IV	-1M

Note: T indicates the type of transformation of variables to ensure stationarity (1=first difference of logarithm, 2=first difference) and F indicates frequency. Series at higher frequencies are converted to monthly frequency by using corresponding frequency averages. Our reference point for delays is the end of months in which we produce the nowcasts. The variable highlighted with the dark grey is the data obtained from the predictions of the Survey of Professional Forecasters for real GDP. "-1M" indicates that predictions up to four quarters ahead are available in the second month of the current quarter.

Table 2: The marginal and predictive likelihoods of competing models

	Marginal Likelihoods			**Predictive Likelihoods**								
	Q0			**Q0**			**Q1**			**Q2**		
	M1	**M2**	**M3**	**M1**	**M2**	**M3**	**M1**	**M2**	**M3**	**M1**	**M2**	**M3**
BM	-227.06	-226.83	-227.14	-176.95	-162.07	-154.93	-195.74	-194.90	-187.02	-204.23	-203.28	-201.47
(BM+S)-M0	-0.25	0.75	0.37	4.09	2.93	0.80	4.62	5.02	3.77	4.09	4.56	4.46
(BM+S)-M1	7.58	5.59	7.71	5.26	5.84	3.07	5.01	7.18	4.77	4.19	4.80	4.62
(BM+S)-M2	9.42	9.19	9.10	5.41	6.02	3.17	4.15	7.32	4.59	3.27	4.62	3.65
(BM+S)-M3	9.53	9.73	9.01	5.51	4.80	2.70	3.87	6.81	4.64	1.58	3.84	3.62
(BM+S)-M4	9.18	9.31	9.02	5.30	3.27	2.06	3.95	7.35	5.36	1.77	4.04	2.98
BMSV	0.78	1.12	0.65	14.43	16.86	19.75	12.19	12.97	13.12	9.62	9.46	9.61
(BMSV+S)-M0	-0.37	2.71	6.84	18.56	19.36	20.63	17.43	17.25	18.05	14.18	14.11	14.66
(BMSV+S)-M1	0.66	1.39	10.07	16.95	20.45	19.55	15.24	17.10	15.59	11.56	11.84	11.95
(BMSV+S)-M2	10.95	6.98	10.65	16.65	19.82	18.51	14.65	16.88	15.04	10.82	11.79	10.57
(BMSV+S)-M3	13.60	2.34	8.26	16.17	18.88	18.28	13.79	17.66	15.02	9.68	10.95	10.42
(BMSV+S)-M4	9.88	9.76	9.33	19.11	21.19	22.38	15.37	20.00	18.39	10.95	13.10	12.93
(BMSV+S)-MV0	6.42	7.13	6.65	20.80	23.49	23.49	18.85	19.23	19.12	15.47	15.95	15.87
(BMSV+S)-MV1	18.32	17.82	19.02	21.43	26.50	26.22	18.84	21.20	20.17	15.04	16.26	15.62
(BMSV+S)-MV2	21.66	22.29	21.27	21.39	26.00	25.99	18.08	21.16	20.19	13.57	15.84	15.69
(BMSV+S)-MV3	19.11	19.24	17.38	22.03	25.14	25.55	18.37	21.69	20.75	12.80	15.53	16.07
(BMSV+S)-MV4	20.10	17.41	18.07	21.98	24.98	25.97	19.21	21.33	20.86	13.59	16.23	17.01

Note: The marginal likelihoods are calculated for the first, second, and third months of the current quarter for the period starting from 1968 until 2019. The predictive likelihoods are computed at the end of each month in the current quarter, for the current quarter (Q0-M1, Q0-M2, Q0-M3), which implies nowcasts, for one quarter ahead (Q1-M1, Q1-M2, Q1-M3) and for two quarters ahead (Q2-M1, Q2-M2, Q2-M3), which imply forecasts, over the evaluation period starting from 1977 until 2019 recursively. The estimation sample starts from 1968 as described in (15). BM stands for the baseline model as described in (5) and (BM+S)-Mk stands for the model where we extend the BM model with the first moment of the survey-based predictions for $k = 0, 1, \ldots, 4$ as described in (10). BMSV stands for the baseline model together with stochastic volatility as described in (6). Finally,(BM+SV)-MVk stands for the model where we extend the BMSV model with the first and the second moment of the survey-based predictions for $k = 0, 1, \ldots, 4$ as described in (10) and (11).

Table 3: The first and second moments of the predictive distributions of 2020-Q1 and 2020-Q2 real GDP growth rates based on predictions from various sources

		2020-Q1			2020-Q2		
		Jan	Feb	Mar	Apr	May	Jun
SPF-Feb 14	Mean		1.7			2.1	
	$\sqrt{Dis}$		0.6			0.5	
SPF-May 15	Mean					-31.5	
	$\sqrt{Dis}$					7.1	
BMSV	Mean	2.2	2.4	2.2	-11.2	-33.2	-25.2
	Volatility	1.1	1.1	1.1	1.1	1.3	1.3
(BMSV+S)-MV4	Mean	2.2	2.2	2.1	-7.5	-35.2	-31.3
	Volatility	1.5	1.5	1.5	1.4	1.8	1.8
New York Fed		1.6	2.1	1.7	-7.8	-35.5	-16.3
Atlanta Fed		2.7	2.6	2.7	-12.1	-51.2	-39.5
		Advance Release	Second Release	Third Release	Advance Release	Second Release	Third Release
Actual		**-4.8**	**-5.0**	**-5.0**	**-32.9**	**-31.7**	**-31.4**

Note: The first four rows display the mean of the predictions performed by the survey participants and the square root of the disagreement between those, $\sqrt{Dis}$, for the first and second quarters of 2020 released on February 14, 2020, and May 15, 2020, denoted as SPF-Feb 14 and SPF-May 15. BMSV stands for the baseline model with stochastic volatility as described in (6). (BM+SV)-MV4 stands for the model where we extend the BMSV with the first and the second moment of the survey-based predictions for $k = 0, 1, \ldots, 4$ as described in (10). New York and Atlanta Feds stand for the Federal Reserve Banks of New York and Atlanta and the numbers correspond to the point nowcasts released at the end of the month displayed on top of the table, see https://www.newyorkfed.org/research/policy/nowcast and https://www.frbatlanta.org/cqer/research/gdpnow. Actual stands for the realization of the real GDP growth provided by the advance, second and third releases, respectively, see https://www.bea.gov/news/current-releases.

Figure 1: The mean of SPF based predictions

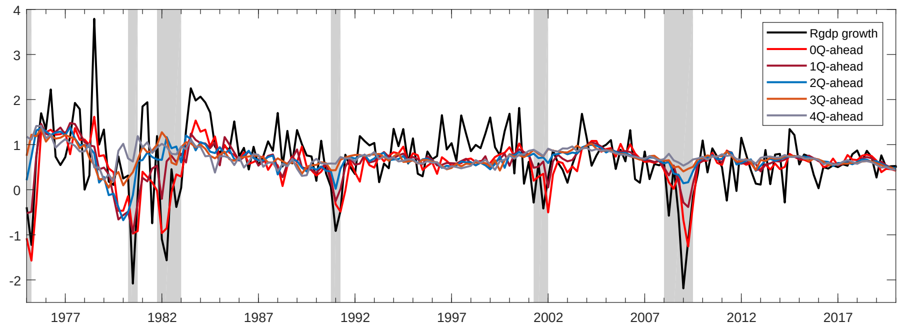


Figure 2: The disagreement in SPF based predictions

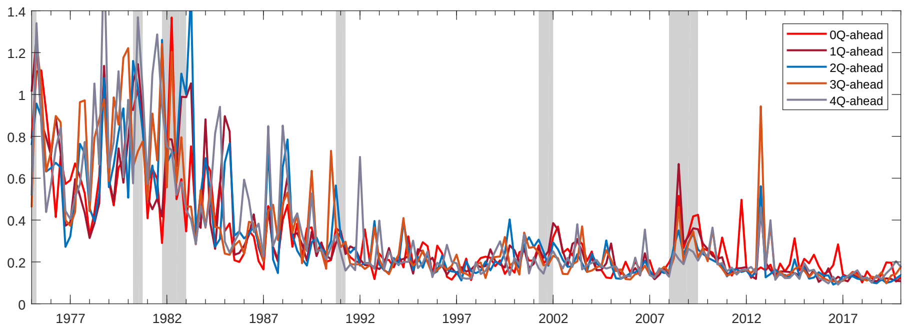
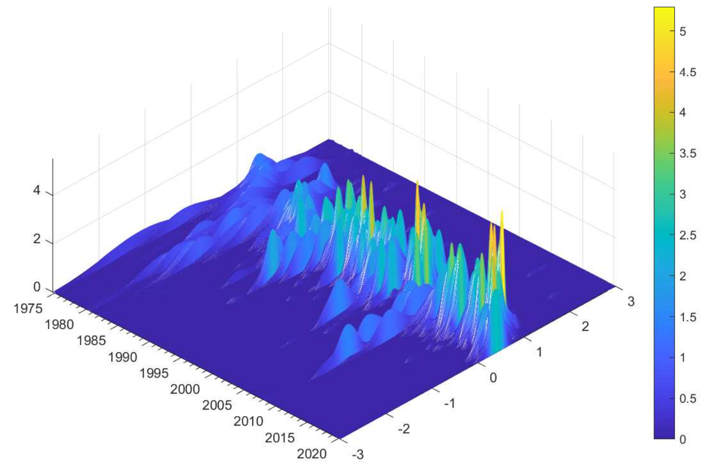
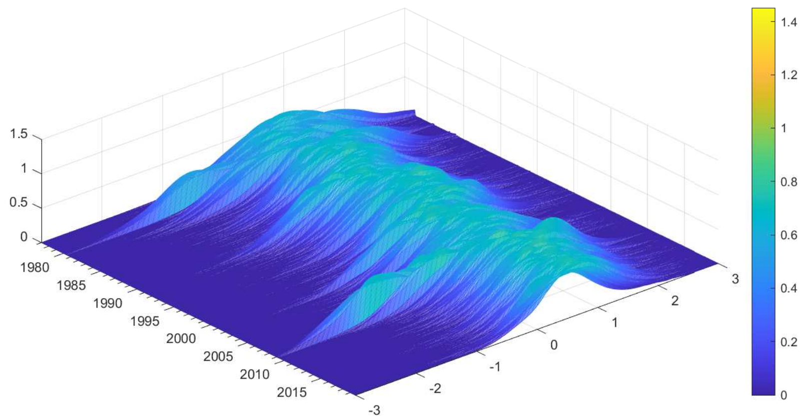


Figure 3: The distribution of the SPF and model based predictions for the current quarter real GDP growth over the period starting from 1977 until 2019

SPF based predictions



Model (BM) based predictions



Model ((BMSV+S)-MV4) based predictions

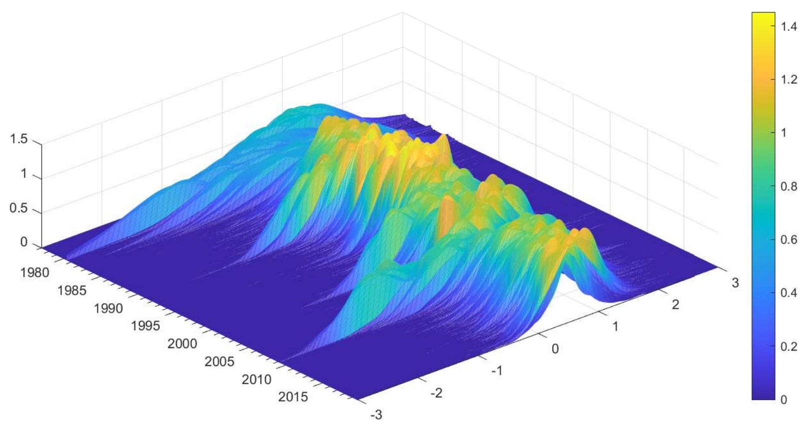


Figure 4: Fitted values of GDP growth obtained using BMSV model

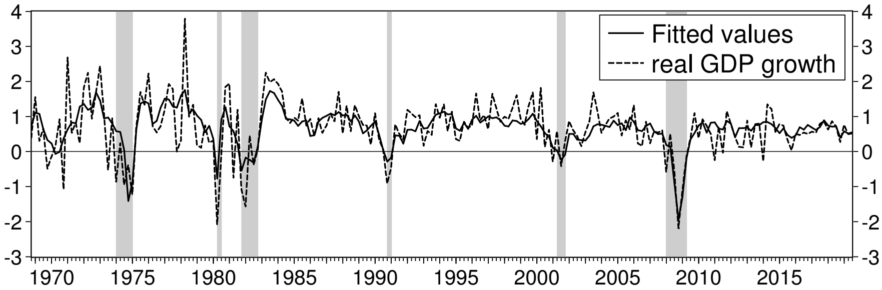


Figure 5: Comparison of volatility estimates

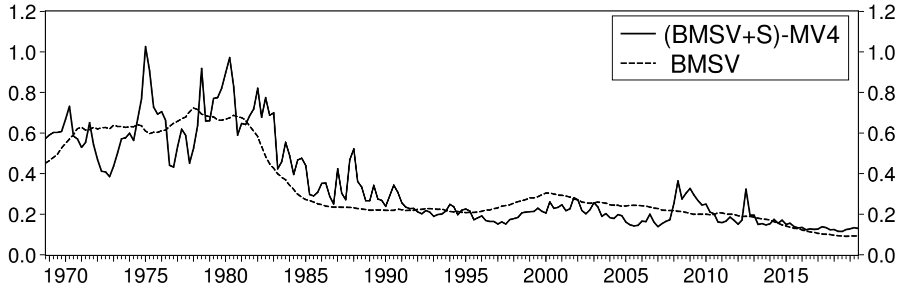


Figure 6: The evolution of Bayes factors over time computed using predictions from the selected models obtained at the end of the second and third months of the quarters

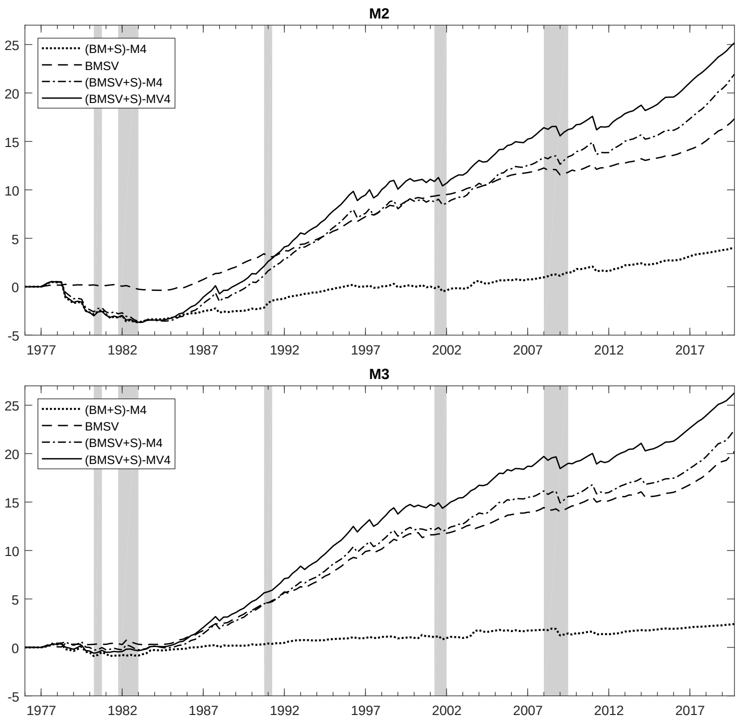
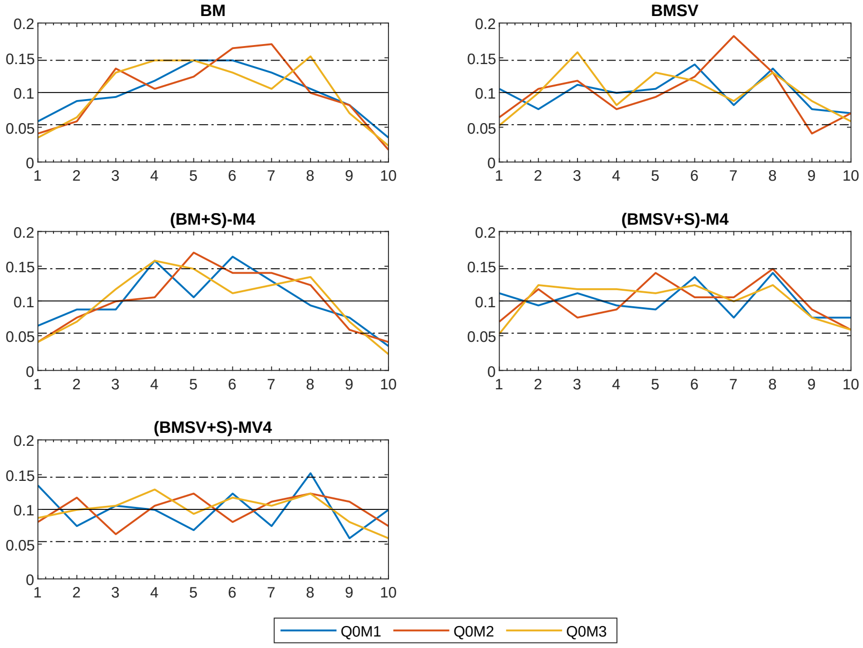
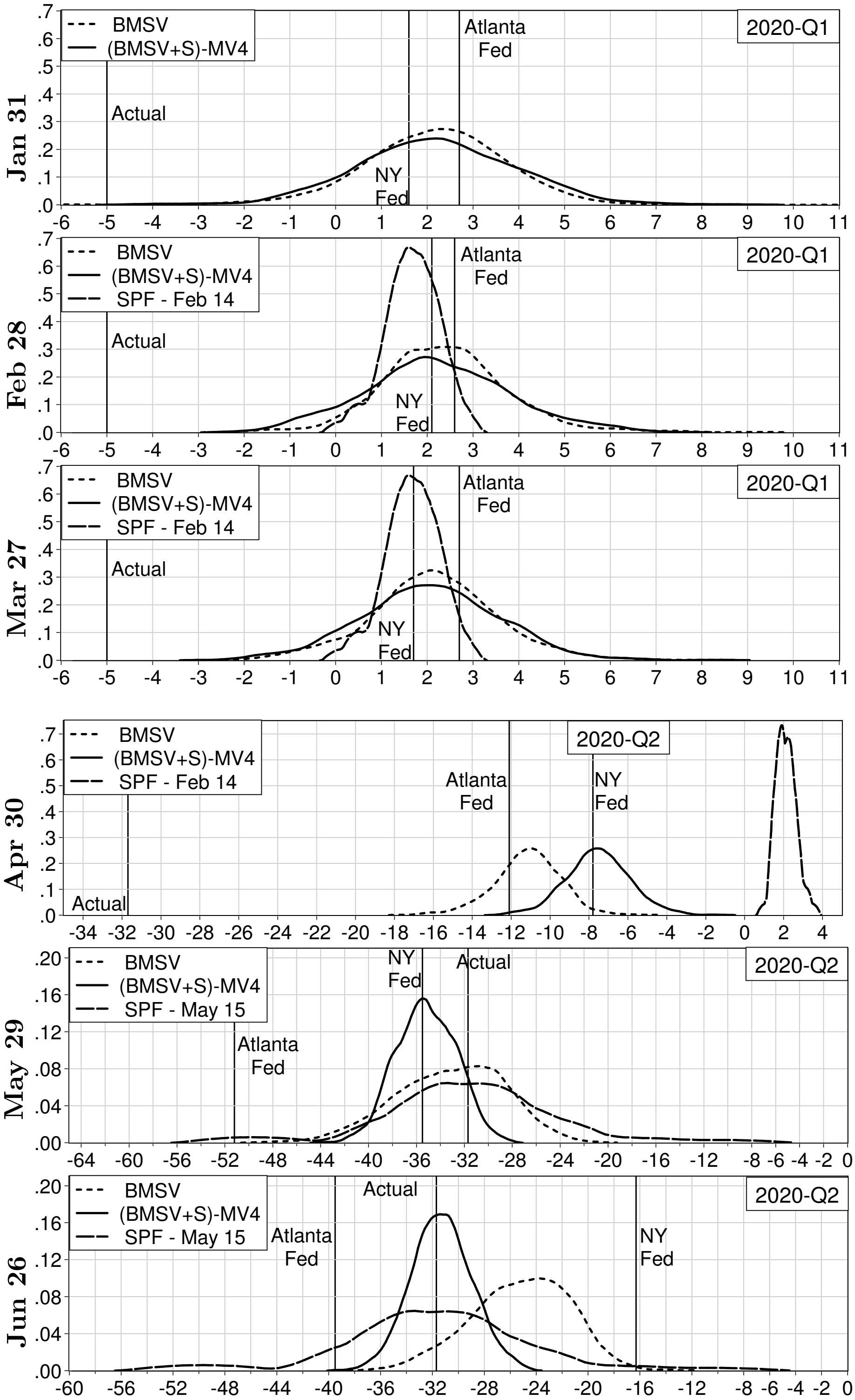


Figure 7: PITs computed using nowcasts from competing models



Note: The blue, red, and yellow lines correspond to the PITs computed using the predictive densities of real GDP growth nowcasts performed at the end of the first, second, and third month, denoted as Q0M1, Q0M2, and Q0M3, respectively.

Figure 8: Predictive distributions of 2020-Q1 and 2020-Q2 real GDP growths based on predictions from the selected models performed at the end of first six months of 2020



Note: The dotted line corresponds to the densities obtained by the BMSV model. The black line corresponds to the model enhanced with survey information, (BMSV+S)-MV4. The dashed line corresponds to the survey-based density of real GDP growth. The first column displays the dates when the predictions are performed using the available dataset. We also display the actual realization of the real GDP growth using the most recent release by a vertical line. We include the **point** nowcasts provided by the Federal Reserve Banks of New York and Atlanta displayed in Table 3 by vertical lines.

Appendix A Econometric Model and Details on Estimation

A.1 Econometric model

This section provides the details about the econometric specifications including the stochastic volatility component and the specification for enhancing the statistical model with the survey base predictive distributions. All competing models are the specific cases of the general form as displayed below

$$\begin{aligned}
y_t^q &= \lambda^q f_t^q + \varepsilon_t^q \\
y_{i,t}^m &= \lambda_i^m f_t^m + \varepsilon_{i,t}^m \quad \varepsilon_{i,t}^m \sim N(0,\sigma_i^2) \text{ for } i=1,2,..,n_m \\
f_t^m &= \phi f_{t-1}^m + u_t^m \quad u_t^m \sim N(0,\sigma_u^2) \\
f_t^q &= \sum_{s=0}^{4} w_s f_{t-s}^q \text{ for } t=3k,\ k=1,2,..,K \\
\varepsilon_t^q &= \exp(\tfrac{h_t}{2})\epsilon_t \quad \epsilon_t \sim N(0,1) \\
h_t &= h_{t-3} + \eta_t \quad \eta_t \sim N(0,\sigma_\eta^2) \\
E_t^S[y_{t+1}^q] &= \lambda^q(\phi f_t^m + 2f_t^m + 3f_{t-1}^m + 2f_{t-2}^m + 1f_{t-3}^m) + \psi_{1,t} \\
E_t^S[y_{t+h}^q] &= \lambda^q \left(\sum_{k=0}^{4} w_k \phi^{h-k}\right) f_t^m + \psi_{h,t} \quad \psi_{h,t} \sim N(0,\sigma_{\psi_h}^2) \\
\log(D_{t,h}) &= h_t + \xi_{t,h} \quad \xi_{h,t} \sim N(0,\sigma_{\xi_h}^2) \quad \text{with} \quad h=3k+1 \text{ for } k=0,1,2,3,4.
\end{aligned} \tag{A.1}$$

The first part of the model excluding stochastic volatility and the measure of disagreement can be cast into a state-space form as

$$\begin{aligned}
\mathbf{y}_t &= \mathbf{H_1}\boldsymbol{\beta}_t + \boldsymbol{\varepsilon}_t \qquad \boldsymbol{\varepsilon}_t|h_t \sim N(\mathbf{0},\mathbf{R}_{\mathbf{1},t}) \\
\boldsymbol{\beta}_t &= \mathbf{F}_t\boldsymbol{\beta}_{t-1} + \boldsymbol{\zeta}_t \qquad \boldsymbol{\zeta}_t \sim N(\mathbf{0},\mathbf{\Omega_t}),
\end{aligned} \tag{A.2}$$

$$\mathbf{y}_t = \begin{bmatrix} y_t^q & E_t^S[y_{t+1}^q] & E_t^S[y_{t+4}^q] & E_t^S[y_{t+7}^q] & E_t^S[y_{t+10}^q] & E_t^S[y_{t+13}^q] & y_{1,t}^m & y_{2,t}^m & \dots & y_{16,t}^m \end{bmatrix}',$$

$$\mathbf{H_1} = \begin{bmatrix} \lambda^q & \lambda^q & 0 & & & \dots & & & & 0 \\ 0 & & & & \dots & & & & & 0 \\ 0 & \phi & \lambda^q\mathbb{A}_1 & \lambda^q\mathbb{A}_2 & \lambda^q\mathbb{A}_3 & \lambda^q\mathbb{A}_4 & \lambda_1^m & \lambda_2^m & \dots & \lambda_{16}^m \end{bmatrix}', \ \boldsymbol{\beta}_t = \begin{bmatrix} f_t^q & \bar{f}_t^q & f_t^m \end{bmatrix}',$$

$\mathbf{R}_{\mathbf{1},t}$ is a diagonal matrix with the vector of diagonal elements as

$$\begin{bmatrix} \exp(\tfrac{h_t}{2}) & \sigma_{\psi_0}^2 & \sigma_{\psi_1}^2 & \dots & \sigma_{\psi_4}^2 & \sigma_1^2 & \dots & \sigma_{16}^2 \end{bmatrix}',$$

$$\mathbf{F_{1}}_{,t} = \begin{cases} \begin{bmatrix} 0 & I_r & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \phi \end{bmatrix}, & t = 3k+1, \\ \begin{bmatrix} I_r & 0 & 0 \\ 0 & I_r & 0 \\ 0 & 0 & \phi \end{bmatrix}, & \text{otherwise,} \end{cases} \quad , \mathbf{F_{2}}_{,t} = \begin{cases} \begin{bmatrix} I_r & 0 & -w_{R(3-t,3)} \\ 0 & I_r & 0 \\ 0 & 0 & I_r \end{bmatrix}, & t = 3k+1, \\ \begin{bmatrix} I_r & 0 & -w_{R(3-t,3)} \\ 0 & I_r & -w_{R(3-t,3)+3} \\ 0 & 0 & I_r \end{bmatrix}, & \text{otherwise,} \end{cases} ,$$

and $\mathbf{Q}$ is the square matrix of zeros of size three with the third diagonal element is replaced by σ_u^2. Moreover, $\mathbf{F}_t = \mathbf{F}_{\mathbf{1},t}^{-1}\mathbf{F}_{\mathbf{2},t}$ and $\mathbf{\Omega_t} = \mathbf{F}_{\mathbf{1},t}^{-1}\mathbf{Q}(\mathbf{F}_{\mathbf{1},t}^{-1})'$.

Stochastic volatility component can further be incorporated as

$$\begin{aligned} \varepsilon_t^q &= \exp(\tfrac{h_t}{2})\epsilon_t\ , \quad \epsilon_t \sim N(0,1) \\ h_t &= h_{t-3} + \eta_t \quad \eta_t \sim N(0,\sigma_\eta^2) \end{aligned} \tag{A.3}$$

where $t = 3k,\ k = 1,2,..,K$, defined at quarterly intervals. Squaring and taking the logarithm of the first equation in (A.3), we obtain the following:

$$\log((\varepsilon_t^q)^2) = h_t + log(\epsilon_t^2)\ , \quad \epsilon_t \sim N(0,1) \tag{A.4}$$

$$\log(D_{t,h}) = h_t + \xi_{t,h} \quad \text{with} \quad h = 3k+1 \text{ for } k = 0,1,2,3,4. \tag{A.5}$$

These equations can be cast into the state-space form as

$$\begin{aligned} \mathbf{\Upsilon}_t &= \boldsymbol{\mu}_t + h_t + \boldsymbol{\xi}_t \\ h_t &= h_{t-3} + \eta_t \qquad \eta_t \sim N(0,\sigma_\eta^2), \end{aligned} \tag{A.6}$$

$$\mathbf{\Upsilon}_t = \begin{bmatrix} \log((\varepsilon_t^q)^2) \\ \log(D_{t,0}) \\ \log(D_{t,3}) \\ \log(D_{t,6}) \\ \log(D_{t,9}) \\ \log(D_{t,12}) \end{bmatrix}, \mathbf{R_{2}}_{,t} = \begin{bmatrix} s_{w_t}^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_{\xi_0}^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_{\xi_1}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_{\xi_2}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_{\xi_3}^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_{\xi_4}^2 \end{bmatrix}, \boldsymbol{\mu}_t = \begin{bmatrix} \mu_{w_t} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

The distribution of $\boldsymbol{\xi}_t$ and the parameters $s^2_{w_t}$ and μ_{w_t} will explicitly be analyzed in the next section when we discuss the inference on stochastic volatility component. The dynamic factor model is constituted by (A.2) together with (A.6). The joint posterior distribution of the parameters is constructed by the product of prior together with the likelihood. In Section A.2, we discuss the prior distributions and in Section A.3, we discuss the full posterior conditional distributions together with the resulting simulation scheme.

A.2 Prior distributions

We use diffuse priors for most of the parameters to let the data be decisive for estimation results. For the autoregressive coefficients of common (monthly) factor, ϕ we use flat prior

$$f(\phi) \propto 1 \tag{A.7}$$

in a restricted set where characteristic roots of ϕ lie outside the unit circle.

For the monthly factor loading parameters we also use flat priors of form

$$f(\lambda_i^m) \propto 1 \quad \text{for } i = 1, \ldots, n_m. \tag{A.8}$$

For the variance parameters of variables as well as factors and volatility, we use noninformative Jeffreys priors. These correspond to the following for a representative variance parameter, σ^2

$$f(\sigma^2) \propto \sigma^{-2}, \tag{A.9}$$

see Geisser (1965).

A.3 Posterior inference

The posterior distribution is proportional to the product of the complete data likelihood, for which we use a multivariate Normal distribution conditional on the volatility component, together with the prior specifications described in (A.9). For the inference of the posterior distribution, we use Metropolis within the Gibbs algorithm that leads to the following sampling scheme. We start with initializing the parameters, at step (s) of the iteration

1. Sample f^T from $p(f^T|y^T, \phi^{(s-1)}, \sigma_u^{2(s-1)})$
2. Sample ϕ from $f(\phi|y^T, f^{T(s)}, \sigma_u^{2(s-1)})$

3. Sample σ_u^2 from $f(\sigma_u^2|y^T, f^{T(s)}, \phi^{(s)})$
4. Sample λ_i^m from $f(\lambda_i^m|y^T, f^{T(s)}, \sigma_i^{2(s-1)})$ for $i = 1, \ldots, 16$,
5. Sample σ_i^2 from $f(\sigma_i^2|y^T, f^{T(s)}, \lambda_i^{m(s)})$ for $i = 1, \ldots, 16$,
6. Sample λ^q from $f(\lambda^q|y^T, f^{T(s)}, h^{T(s-1)}, \phi^{(s)}, \sigma_{nf}^{2(s-1)})$,
7. Sample $\sigma_{\psi_h}^2$ from $f(\sigma_{\psi_h}^2|y^T, f^{T(s)}, \phi^{(s)}, \lambda^{q(s)})$ for $h = 0, 1, \ldots, 4$,
8. Sample h^T from $p(h^T|y^T, f^{T(s)}, \lambda^{q(s)}, \sigma_\eta^{2(s-1)})$,
9. Sample σ_η^2 from $f(\sigma_\eta^2|y^T, f^{T(s)}, h^{T(s)}, \lambda^{q(s)})$,
10. Sample $\sigma_{\xi_h}^2$ from $f(\sigma_{\xi_h}^2|y^T, f^{T(s)}, h^{T(s)}, \lambda^{q(s)})$ for $h = 0, 1, \ldots, 4$,
11. Repeat (1)-(10) S times.

Our model specifications imply that the parameter, ϕ, in our measurement equations corresponding to the incorporation of survey information admits a nonlinear form in multiple equations. This leads to the use of the Metropolis-Hastings algorithm to sample these parameters in steps (2) and (6). Conditional on these parameters, sampling of remaining parameters can be conducted using standard distributions via Gibbs steps. In the next section, we derive the conditional distributions used in the posterior sampling scheme.

A.4 Conditional Posterior Distributions

A.4.1 Sampling of f_t

The framework (A.2) is a linear Gaussian state-space model conditional on the time-varying variance and model parameters. Therefore, we first run the Kalman filter forwards and then run the simulation smoother backwards We modify the system for handling missing observations due to ragged edge, unbalanced data, and mixed frequency of variables before running the filter. Let $\mathbf{W}_t$ be a diagonal selection matrix, with i^{th} diagonal element takes the value 1 if $y_{i,t}$ is observed and 0 otherwise. Before running the filter we replace $\mathbf{y}_t$, $\mathbf{H_1}$ and $\mathbf{R}_{\mathbf{1},t}$ with $\mathbf{y}_t^* = \mathbf{W}_t\mathbf{y}_t$, $\mathbf{H_1}^* = \mathbf{W}_t\mathbf{H_1}$ and $\mathbf{R}_{\mathbf{1},t}^* = \mathbf{W}_t\mathbf{R}_{\mathbf{1},t}\mathbf{W}_t^{'}$. Once the Kalman filter is run forward, a simulation smoother is run backward using the filtered values for drawing smoothed states as in Carter et al. (1994) and Frühwirth-Schnatter (1994). For a textbook exposition we refer to Kim et al. (1999) and Durbin and Koopman (2012) for further details.

A.4.2 Sampling of ϕ and σ_u^2

We use a Metropolis-Hastings step to sample ϕ. To obtain an efficient candidate density, we consider the transition equation of the factor

$$f_t^m = \phi f_{t-1}^m + u_t^m \quad u_t^m \sim N(0, \sigma_u^2) \tag{A.10}$$

as a natural choice. To sample ϕ from the candidate density, we use a normal distribution with mean $(\sigma_u^2 \sum_{t=2}^T (f_{t-1}^m)^2)^{-1} \sum_{t=2}^T (f_t^m f_{t-1}^m)$ and variance $(\sigma_u^2 \sum_{t=2}^T (f_{t-1}^m)^2)^{-1}$.

We evaluate the probabilities conditional on the data to compute the acceptance probability of the Metropolis-Hastings sampler using the Kalman filter by directly conditioning on the data rather than factors to improve efficiency. We calculate the average rate of acceptance probability with its standard deviation to check whether candidate draws are not always accepted or rejected, making sure that the draws sufficiently visit the entire surface of the posterior distribution, see the diagnostics discussed in Koop (2003).

Once we sample ϕ, we proceed by sampling σ_u^2 from an inverse-Gamma distribution with scale parameters $\left(\sum_{t=2}^T (f_t^m - \phi f_{t-1}^m)^2\right)$ and degrees of freedom $(T-1)$.

A.4.3 Sampling of λ_i^m and σ_i^2

To sample λ_i^m we consider the monthly measurement equation in (A.1) as a form of the following:

$$Y_t = X_t \lambda_i^m + \varepsilon_{i,t}^m \quad \varepsilon_{i,t}^m \sim N(0, \sigma_i^2) \tag{A.11}$$

We sample λ_i^m using a normal distribution with mean $(X'X)^{-1}X'Y$ and variance $(\sigma_i^2 X'X)^{-1}$, where $Y = (Y_1, \ldots, Y_T)'$ and $X = (X_1', \ldots, X_T')'$.

Once we sample λ_i^m, we proceed by sampling σ_i^2 from an inverse-Gamma distribution with scale parameters $\left(\sum_{t=1}^T (\varepsilon_{i,t}^m)^2\right)$ and degrees of freedom T.

A.4.4 Sampling of λ^q and $\sigma_{\psi_h}^2$

To sample λ^q, we use another Metropolis-Hastings step similar to sampling ϕ. In order to obtain a candidate generating density, we consider the equation

$$y^q_{i,t} \;=\; \lambda^q f^q_t + \varepsilon^q_t \qquad \text{(A.12)}$$

Let $Y_t = {y^q_t}/{\exp(h_t/2)}$ and $X_t = {f^q_t}/{\exp(h_t/2)}$. Our candidate generating density is normal distribution with mean $(X'X)^{-1}X'Y$ and variance $(X'X)^{-1}$.

Conditional on λ^q and ϕ, we sample $\sigma^2_{\psi_h}$ from an inverse-Gamma distribution with scale parameters $\left(\sum_{k=1}^{K}(\psi_{h,3k-1})^2\right)$ and degrees of freedom $K(={T}/{3})$.

A.4.5 Sampling of h_t, σ^2_η and $\sigma^2_{\psi_h}$

Once f^T and λ^q are sampled we compute ε^q_t in (A.1), and use the logarithmic transformation that leads to following form

$$\log((\varepsilon^q_t)^2) \;=\; h_t + \log(\epsilon^2_t)\ , \quad \epsilon_t \sim N(0,1) \qquad \text{(A.13)}$$

Sampling of volatilities, h_t, is conducted following Omori et al. (2007). This involves approximating the log-χ^2 distribution of $\log(\epsilon^2_t)$, due to the nonlinear transformation of the model, using a mixture of ten Gaussian distributions, denoted as $N(\mu_{w_t}, s^2_{w_t})$. Hence, conditional on the mixtures the model becomes linear and Gaussian where standard inference can be carried out easily. Here, we proceed with the same strategy as in Omori et al. (2007) and refer to this study for details. For the remaining measurement equations regarding survey information, we use a Normal distribution for the error terms to facilitate the inference of the volatility terms. This implies sampling the variances from an inverse-Gamma distribution with scale parameters constructed by the squared sum of corresponding error terms and degrees of freedom $K(={T}/{3})$. Alternatively, sequential Monte Carlo methods serve as an attractive strategy that provides exact inference. Indeed, Rao-Blackwellized particle filters are very suitable for the problem at hand, see, for example, Creal (2012) for a comprehensive survey. However, our system has a relatively large dimension together with the requirement of running the filter multiple times for estimating the remaining model parameters. Therefore, we do not perform sequential Monte Carlo methods as they are computationally costly, especially, when smoothed estimates are required. Given that we also perform a recursive forecasting exercise, such computationally costly approaches are not feasible.

Appendix B Evolution of the Bayes factors computed using nowcasts performed in each month of the quarter

Figure B.1: The evolution of Bayes factors over time computed using predictions from the models obtained in the first, second, and third month of the quarters