TI 2022-069/III
Tinbergen Institute Discussion Paper

Does economic uncertainty predict real activity in real-time?

Bart Keijsers[1,3]
Dick van Dijk[2,3]

1 University of Amsterdam
2 Erasmus University Rotterdam
3 Tinbergen Institute

Does economic uncertainty predict real activity in real-time?*

Bart Keijsers†
University of Amsterdam
Tinbergen Institute

Dick van Dijk
Erasmus University Rotterdam
Tinbergen Institute

September 26, 2022

Abstract

We assess the predictive ability of 15 economic uncertainty measures in a real-time out-of-sample forecasting exercise for the quantiles of The Conference Board's coincident economic index and its components (industrial production, employment, personal income, and manufacturing and trade sales). The results show that the measures hold (real-time) predictive power for quantiles in the left tail. Because uncertainty measures are all proxies of an unobserved entity, we combine their information using principal component analysis. A large fraction of the variance of the uncertainty measures can be explained by two factors. First, a general economic uncertainty factor with a slight tilt toward financial conditions. Second, a consumer/media confidence index which remains elevated after recessions. Using a predictive regression model with the factors from the set of uncertainty measures yields more consistent gains compared to a model with an individual uncertainty measure. Further, although often better forecasts are obtained using the National Financial Conditions Index (NFCI), the uncertainty factor models are superior when forecasting employment and in general the uncertainty factors have predictive content that is complementary to the NFCI.

Keywords: Economic uncertainty, real-time forecasting, quantile forecasting, factor analysis

JEL classification: E27, C21, C38

*We would like to thank Sander Barendse and seminar participants at the University of Amsterdam, University of New South Wales, Monash University, University of Sydney, and the IAAE Conference 2022 in Cyprus for helpful discussions and feedback. We thank ICE for access to the BofA US Bond Market Option Volatility Estimate Index, and Jeremy Piger for sharing real-time real activity data.

†Corresponding author. E-mail address `b.j.l.keijsers@uva.nl`.

1 Introduction

Understanding the fundamental causes of business cycles has intrigued macroeconomists for decades, if not centuries. According to real option theory (Bernanke, 1983; Dixit and Pindyck, 1994), uncertainty is one of the key drivers of such cyclical fluctuations: as uncertainty increases, businesses hold off on investment and consumers postpone large purchases, thus reducing economic activity, such as output and employment. Bloom (2009) sparked a new line of research, on empirically measuring economic uncertainty and assessing its relationship with real macroeconomic variables, see Bloom (2014) for an overview. This is not a straightforward exercise, because uncertainty is a latent concept and its exact definition can be debated. Not surprisingly then, a variety of measures of economic uncertainty has been proposed over the last decade. Examples include financial volatility (Bloom, 2009), news based indices (Baker et al., 2016), dispersion in micro data (Bloom, 2009), and disagreement among professional forecasters (Rossi et al., 2016).

On the introduction of a new measure of economic uncertainty, it is usually added to a vector autoregressive model to assess its impact on macroeconomic variables, typically by means of impulse response functions. The comparison to other measures is usually limited to simple correlations, a visual comparison of extremes, and of impulse response functions. All uncertainty measures are proxies of a latent entity, which makes it difficult to assess their validity. This partly explains why a thorough (statistical) comparison of the proposed measures is lacking. Furthermore, evidence of the (dynamic) relationship between uncertainty and economic activity thus far is based on in-sample analysis only. It is important for the validity of these findings to test whether the relationship holds out-of-sample.

In this paper we address both open issues identified above. First, we collect a comprehensive set of different uncertainty measures and conduct a factor analysis. This allows us to examine the similarities and differences between the various measures. Furthermore, the resulting factors, essentially combining the information in the different measures, might provide more comprehensive and accurate proxies of (different aspects of) the underlying notion of 'uncertainty.' Second, we conduct a real-time out-of-sample forecasting analysis to assess whether a forecaster is able to take advantage of the implied relationship between uncertainty and economic activity. This is important to gain insight into the practical usefulness of the various uncertainty measures. Though in-sample tests have more power (Inoue and Kilian, 2005), a forecasting analysis is relevant for policy makers. They might be able to use this extra source of information to improve decision making.

For the first part of our analysis, we identify 15 monthly uncertainty measures that comply with a number of restrictions such as being freely and directly available for a substantial time period. An important additional restriction is that the uncertainty measure should not be subject to revisions, such that we can exploit it in our real-time forecasting exercise. We show that the various measures can be categorized into five categories, based on their source: (i) volatility, (ii) cross-sectional dispersion, (iii) news, (iv) surveys, and (v) forecast errors. The collected measures are spread quite evenly across these categories.

The factor analysis shows that there is indeed a relatively strong common component. The first principal component explains about 40% of total variation for the period 1989-2021. It can be interpreted as general economic uncertainty, because it loads positively on all measures, though it loads slightly more strongly on financial information. Interestingly, the importance of the factor increases during periods of financial stress. Additionally, we identify a second factor, which can be interpreted as media/consumer uncertainty. It loads most heavily on news based and consumer survey based uncertainty measures. This factor remains elevated after recessions, reflecting that the media and consumers need more time to become confident about the recovery than reflected by the fundamentals. Finally, the factors are robust over time. While the COVID-19 period does lead to some differences, the factors remain clearly identified.

For the second part of our analysis, we set up an extensive real-time out-of-sample analysis to forecast The Conference Board's US coincident economic index, and its components: industrial production, employment, manufacturing and trade sales, and personal income excluding transfer payments. Note that these variables are also taken into account by the NBER business cycle dating committee, confirming their importance as measures of real economic activity. In contrast with pseudo out-of-sample analyses, we use different data vintages to take into account that publications of macroeconomic variables are revised multiple times after the first release. Using these vintages allows us to assess whether a forecaster is able to gain from using the values that are available at that point in time. The importance of employing real-time data in forecasting analyses is discussed in Croushore (2006), among others. The increased interest in computing the downside risk of macroeconomic growth, also known as growth at risk (Prasad et al., 2019), motivates us to forecast quantiles. This provides insight into possible asymmetries in the relationship between uncertainty and macroeconomic activity. Forecasts are produced for the period 2000 to 2021, based on an expanding window starting in 1990. We consider multiple forecasting horizons, from nowcasting up to 24 months ahead.

We find that at the lower quantiles, the uncertainty measures have forecasting power. This mirrors the asymmetric relationship between output growth and the Chicago Fed's National Financial Conditions Index (NFCI) documented by Adrian et al. (2019), among others. In comparison with the uncertainty measures, we in fact find that generally the NFCI is a stronger predictor. When forecasting employment though, it is better to use a factor model with uncertainty factors. Interestingly, Bloom (2009) finds that employment responds negatively to an uncertainty shock, and uses this to build a labor-capital model. From the individual uncertainty measures, the Jurado et al. (2015) measures – the volatility of forecast errors from a large set of macroeconomic and financial variables – and financial volatility perform best. The performance of individual media and news based measures is disappointing. Forecasting accuracy is better and more consistent when using a factor model instead of individual uncertainty measures. So in that sense it is recommended to combine information from multiple uncertainty measures.

Our paper provides three main contributions. First, we add to the literature by conducting a real-time quantile forecasting exercise. Second, we contribute to the empirical uncertainty literature by showing how the different uncertainty measures are related and that they can largely be summarized by two common factors. Third, we provide further empirical evidence of the relationship between economic uncertainty and the labor market.

Multiple papers are interested in forecasting economic output using financial volatility or conditions. For example, Chauvet et al. (2015) perform a real-time forecasting exercise linking multiple financial volatility measures in a Markov-switching dynamic factor model and find evidence for nonlinearities. Further, Adrian et al. (2019) consider the impact of financial conditions on economic output. They allow for asymmetry across the density based on quantile forecasts and find that especially the left tail is affected by financial conditions. Most closely related to our paper in spirit is Giglio et al. (2016), who do a similar analysis for systemic risk measures. They gather multiple systemic risk measures and conduct a quantile forecasting exercise. They find that a single common factor improves their forecast accuracy, and that predictive power for the mean is limited. In that light, our results are comparable. The main difference is – other than using uncertainty measures instead of systemic risk measures – that we conduct a real-time rather than a pseudo out-of-sample forecasting exercise, taking into account revisions. Systemic risk and financial conditions are close in concept to economic uncertainty. Hence, the findings by Giglio et al. (2016) and Adrian et al. (2019) are consistent with our finding that economic uncertainty is useful in forecasting the lower quantiles of economic output.

There is little research on the forecasting performance of uncertainty measures. Concurrently with this paper, contributions are made by Hengge (2019) and Rogers and Xu (2019). Rogers and Xu (2019) predict only the GDP growth rate with a smaller subset of uncertainty measures. Hengge (2019) investigates whether the macro uncertainty measure of Jurado et al. (2015) predicts the GDP growth rate. Here, we differ in that we perform a real-time forecasting exercise, and focus on different measures of economic output. Importantly, Hengge uses the probabilistic score as evaluation measure, while we consider the tick loss to evaluate quantile forecasts. Diks et al. (2011) show that using a different scoring rule can substantially affect the conclusions. Indeed, Hengge (2019) is more positive about the predictive ability of uncertainty, and in particular in comparison to NFCI. Our results show that NFCI is hard to beat, but the uncertainty measures hold additional predictive information at the medium horizon, and outperform it when forecasting employment.

The paper is structured as follows. Section 2 describes the uncertainty measures, the selection criteria and the different categories, followed by the factor analysis in Section 3. Section 4 provides the methodology and implementation details of the real-time forecasting analysis. The forecasting results are presented in Section 6. Section 7 compares the uncertainty measures with financial conditions, and Section 8 concludes.

2 Uncertainty measures

Our selection of uncertainty measures is based on a number of criteria. First, we restrict to US data such that all measures aim to capture the same entity. By far the largest number of measures is available for the US and it makes the results better comparable to the existing literature. Second, to match the frequency of the economic activity variables used in the second part of our analysis we focus on monthly data. Measures available at a higher frequency are transformed to monthly frequency appropriately. Measures reported at a lower frequency are excluded. They can be included using mixed frequency methods, see e.g. Carriero et al. (2018), but we choose to focus the analysis on a single frequency. Third, the data should be available in real-time, because we are interested in whether forecasters had been able to take advantage of the information. This excludes measures that are estimated using ex post data, such as forecast error distributions and many other decompositions. This does not exclude filtered time series though. Fourth, we require a sufficient time series length such that we have reasonable power for the forecast evaluation. Fifth, on a more practical note, the data should be feasible to collect or compute.

Table A.1 lists the selected uncertainty measures, including a brief description, the source and sample size. It is a reasonably sized set of 15 measures from October 1989 to December 2021, and includes most of the popular ones that have been proposed thus far.

Notable exclusions are cross-sectional dispersion of firm level profit growth (Bloom, 2009), total factor productivity growth (Bloom, 2009; Kehrig, 2015), Livingstone survey GDP forecasts (Bloom, 2009), price changes (Vavra, 2013), and employment growth (Bachmann and Bayer, 2014). Conditional volatility from decomposing financial volatility into risk aversion and uncertainty (Bekaert et al., 2013), shocks from political turmoil, natural disasters or terrorist attacks (Baker and Bloom, 2013), Fama-French factor residual variance (Gilchrist et al., 2014), and fiscal volatility shocks (Fernández-Villaverde et al., 2015) are excluded as well, either because they need to be computed ex post, or because they are only available at a lower frequency. Furthermore, we ignore measures based solely on the Survey of Professional Forecasters (SPF) because these are of quarterly frequency (Lahiri and Sheng, 2010; Rossi et al., 2016).

The descriptions in Table A.1 show that economic uncertainty can be proxied in a variety of ways and from multiple sources. We identify five categories related to how economic uncertainty is measured.[1] First, a volatility estimate of some underlying, often a financial asset. Times of high conditional volatility are assumed to be related to times of high uncertainty. In our set, the underlying assets are stocks (VIX; Bloom, 2009), long-term bonds (MOVE) and the WTI oil price (OVX; Kellogg, 2014).

The second type of uncertainty measure utilizes micro data by estimating cross-sectional dispersion in each time period for a set of individuals, forecasters or firms. More dispersed outcomes suggest higher economic uncertainty. We consider cross-sectional dispersion in stock returns (CSDR and CSDRsic; Bloom, 2009), and forecast disagreement between respondents from the Philadelphia Fed's Manufacturing Business Outlook Survey (FDISP; Bachmann et al., 2013) and from Consensus Economics GDP growth forecasts (CEgdp; Dovern et al., 2012).

The third source is news, as conveyed via newspaper articles or Bloomberg announcements, among others. In uncertain times, newspapers publish (more) articles to report on the uncertainty and Bloomberg announcements deviate more from expectations. The most prominent measures are the indexes from Baker et al. (2016), based on

[1]Kozeniauskas et al. (2018) also categorize uncertainty measures. They distinguish between macro uncertainty, micro uncertainty, and higher-order uncertainty.

newspaper article counts – we select general economic policy uncertainty (EPU and EPU+) and monetary policy uncertainty (MPU).

Fourth, outcomes of polls or surveys taken among consumers, professional forecasters or firms gauge their expectations for the coming period. It is a direct way of measuring the uncertainty perceived by economic agents. For example, the Reuters/University of Michigan Survey of Consumers includes the response "uncertain times" for not buying a vehicle (LLv; Leduc and Liu, 2016) or large household goods (LLh; Fajgelbaum et al., 2017). FDISP, CEgdp, and EPU+ can also be counted to this category.

Fifth, uncertainty can be inferred from the volatility of forecast errors. This is to distinguish the uncertainty measure from 'forecastable' time-varying volatility. Jurado et al. (2015) construct uncertainty measures based on this principle. They pool a large set of macroeconomic and financial variables, remove the forecastable part, and compute measures as the stochastic volatility of the forecast errors. The volatility is calculated on subsets of macroeconomic (JLNm), financial (JLNf) and real variables (JLNr).

The data set is well balanced across the categories, see Table A.1. There are three measures based on conditional volatility, news, or forecast errors, four measures based on cross-sectional dispersion, and five measures based on survey data.

The EPU and JLN measures are subject to revisions and therefore not, strictly speaking, available in real-time. Both Baker et al. (2016) and Jurado et al. (2015) argue that revisions are small though and capture timely events sufficiently well. Due to the limited size of the revisions, and the popularity of these measures, they are included. The issue is further alleviated by lagging the measures one month.

3 Commonality in uncertainty measures

The comparison of uncertainty measures in the literature is thus far limited to comparing the pattern of the different time series or computing correlations. Furthermore, it usually involves quite a small set of about four uncertainty measures. We analyze the commonalities for our more extensive set of measures, and assess the underlying factor structure. Haddow et al. (2013) and Charles et al. (2018) also use a factor model, but on a set of 4 or 6 measures only, and their sample excludes the recent COVID-19 period.

Figure A.1 presents the time series of all 15 selected uncertainty measures. They are largely similar in that all uncertainty measures peak around the time of recessions. Nevertheless, the average correlation between the measures is quite modest at 37.7%. This seems less than expected given that all measures aim to capture US economic uncertainty. The modest average correlation could be due to measurement error, or because the measures account for different aspects of economic uncertainty. The correlation matrix in Figure A.2 indicates multiple blocks of more strongly correlated measures. The average correlation between categories is 34.0%, while the average correlation within a category is 47.5%. In particular the survey based measures (excluding FDISP) and the forecasting error based measures of Jurado et al. (2015) are closely connected with correlations of 72.5% and 60.9%, respectively. Across categories, measures based on stock market data are similar. For example, the correlation between VIX and JLNf is 79.7%.

To determine the commonality between the uncertainty measures more formally, we extract factors using principal components analysis. Table 1 presents the factor loadings, and explained fraction of total variance for the first five principal components. The principal component analysis suggests the presence of two common factors. Together, they explain more than 60% of the total variance and both have a clear interpretation. The first factor represents average (economic) uncertainty, with a slight emphasis on financial uncertainty. The loadings are all positive, and it explains about 40% of the total variance, see Table 1. The factor level spikes during recessions and periods of financial turmoil, such as Black Monday in October 1987, the Russian financial crisis in 1998, and the Greek government debt crisis in 2012, see Figure 1. Figure 2 shows the explanatory power for the first three recursively estimated factors. It is interesting to observe that the explanatory power of the first factor increases during recessions. A large part of the variation occurs during those periods as most measures increase during recessions, which is captured by the first factor.

The second factor loads most heavily and positively[2] on consumer confidence measures LLv and LLh and news based uncertainty measures EPU and EPU+, see Table 1. This factor can be interpreted as a consumer/media uncertainty factor. Consumers rely on media outlets for economic news, which explains why they are linked. It is interesting to see in Figure 1 that its value remains relatively high after the recession has ended. Apparently, while fundamentals are recovering, the public is still uncertain. This can be because the recovery still has to feed back to consumers, e.g. in the form of new jobs – the unemployment rate typically lags other output variables. It is in line with the jobless recoveries that characterize the periods following the recessions in the 1990s and 2000s (Groshen and Potter, 2003; Jaimovich and Siu, 2020). Further, consumer spending probably lags as well, as their savings might be depleted or at least diminished at the end of a downturn so they probably want to save before spending again. Alternatively, consumers and the media are simply not confident whether the recovery has fully started or if it is simply just a coincidental good output number. This is plausible, given that even the NBER's Business Cycle Dating Committee has a delay of several months in 'officially' calling the end of recessions.

The other factors lack a clear interpretation or explain only a single measure, see Table 1. The third factor explains the variance of a number of variables, but there is no clear link between them. The fourth factor loads heavily of MPU, and the fifth explains most of FDISP.

There is fairly strong commonality between the measures. FDISP is somewhat different from most measures though. The correlations for FDISP with other measures are small in general, with the maximum correlation with the VIX at 23.6%. FDISP is even (weakly) negatively correlated with the consumer survey based measures, with correlations of -10.3% (with LLv) and -10.7% (with LLh). That FDISP is different from most measures is also clear from Table 1. The loading is small for the first factor compared to the others, and less than 20% of FDISP's variance is explained by the first two common factors. The time series in Figure A.1f further confirms its idiosyncratic

[2] The second factor is multiplied by -1. This does not matter for how much of the variance is explained or for forecasting, but makes it more intuitive to explain our interpretation of the factor.

behavior, which could be due its regional focus or because business surveys capture a unique part of uncertainty.

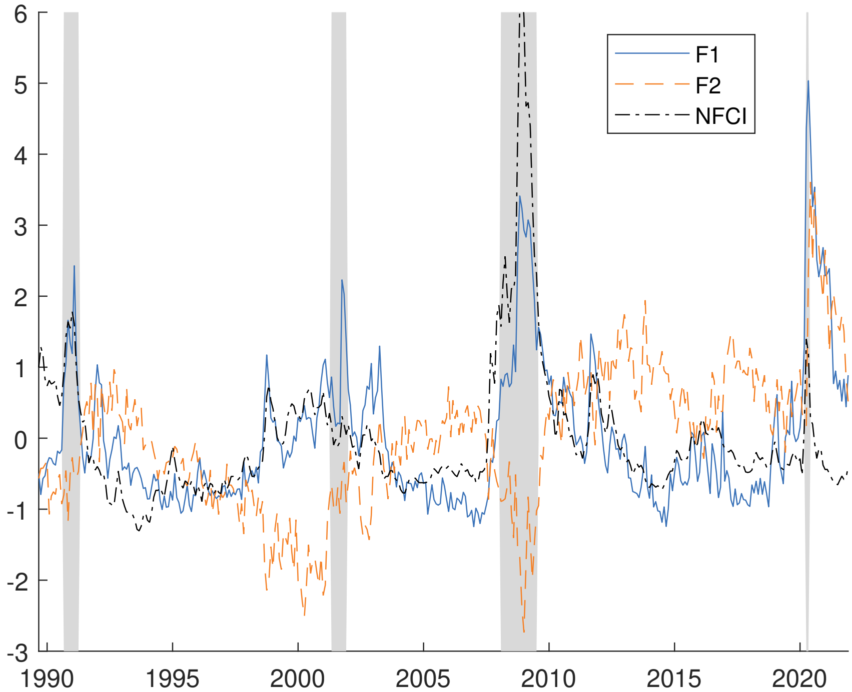
Table 1: Factor loadings and marginal R^2

r	1	2	3	4	5
VIX	0.295	−0.230	0.031	0.095	−0.047
MOVE	0.181	−0.413	0.321	−0.134	0.355
OVX	0.270	−0.154	−0.134	0.287	−0.132
CSDR	0.276	−0.307	−0.081	−0.185	−0.263
CSDRsic	0.185	−0.363	−0.022	−0.240	−0.175
FDISP	0.062	−0.234	−0.055	0.567	0.616
CEgdp	0.263	0.069	−0.030	−0.353	0.333
LLv	0.254	0.318	0.220	−0.186	0.181
LLh	0.274	0.231	0.275	−0.276	0.235
EPU+	0.285	0.307	0.172	0.183	−0.121
EPU	0.273	0.287	0.105	0.300	−0.256
MPU	0.237	−0.057	0.508	0.336	−0.141
JLNm	0.319	0.055	−0.421	0.013	0.179
JLNf	0.301	−0.259	−0.138	−0.019	−0.191
JLNr	0.283	0.251	−0.498	0.071	0.095
R^2	0.443	0.173	0.083	0.078	0.057

The table presents the factor loadings (top panel) and the marginal R^2, the fraction of total variance explained by the r-th principal component, for the first five principal components for the sample 1989M10–2021M11. The second factor is rotated (loadings multiplied by −1) for interpretation purposes. See Table A.1 for an explanation of the abbreviations.

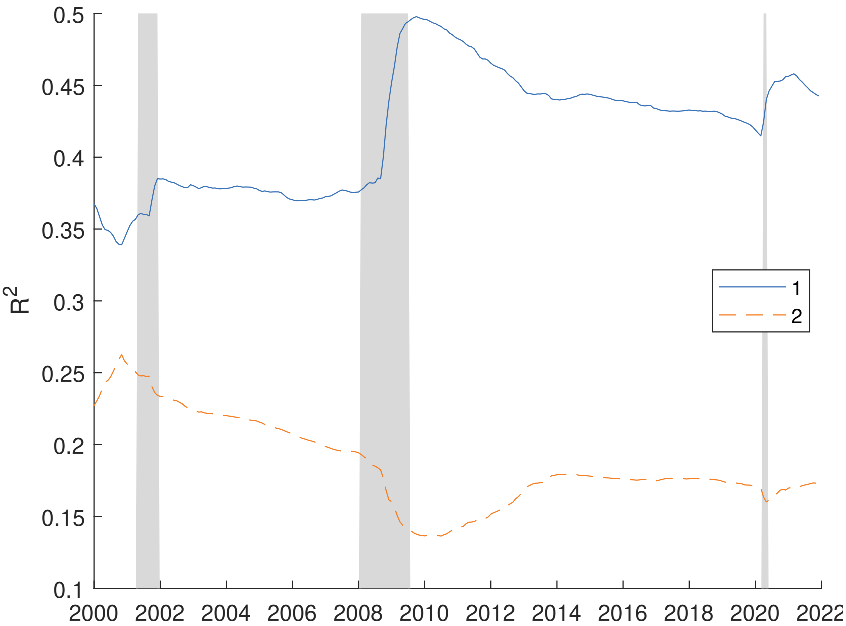
As a final point, the COVID-19 period deserves extra attention. Intuitively the uncertainty increased, but for reasons different from other recessionary periods. Uncertainty indeed increases during the recession according to all measures, although the increase is muted for FDISP and MPU. Afterwards there are more pronounced differences between the measures: while uncertainty quickly decreases according to most measures, it remains high for others (e.g. JLNm, JLNr and CEgdp). Despite a mixed response of the measures to the pandemic, the identification and interpretation of the factors is robust over time. The first and second factor are clearly identified both in the full sample and when excluding 2020–2021, and follow the same general pattern, see Figure B.1. The first factor is nearly identical, while the correlation is also high (84.2%) for the second factor between the different samples. The difference between the samples is due to MPU deviating more from the average uncertainty measure. This results in a weaker loading on the first factor and an increase in the second factor's loading. In turn, the loadings

Figure 1: Uncertainty factors and NFCI



The figure presents the time series of the first (solid blue line) and the second factor (dashed orange line) from the full sample principal components analysis, and the NFCI (dash-dotted black line). The second factor is rotated (multiplied by -1) for interpretation purposes. The NFCI are the end-of-month values from the 16/02/2022 vintage. The gray bars are recessions as determined by NBER's Business Cycle Dating Committee. All series are standardized.

Figure 2: Explanatory power over time



The figure presents the marginal explanatory power of the first (solid blue line) and second principal component (dashed orange line) of the uncertainty factor model, estimated recursively.

of various news and consumer based measures (LLh, LLv and EPU) in the second factor decrease.

4 Methodology

This section describes the target variables, the real-time forecasting setup, and explains the quantile regression model as well as the forecast evaluation.

4.1 Coincident variables

Theory suggests a link of economic uncertainty with the business cycle (Bernanke, 1983). For this reason, we consider The Conference Board's Coincident Economic Index (CEI) as our main target variable. In addition, we analyze the predictive ability of uncertainty for the four CEI constituents, that is, industrial production (IP), nonfarm payroll employment (EMP), manufacturing and trade industries sales (MTS), and personal income excluding current transfer receipts (PIX).

In order to asses whether a forecaster is able to improve the accuracy of her predictions, real-time data should be used. That is, the vintages with values that were available to the forecaster at the time the forecasts are made. This is relevant, because macroeconomic variables are reported with a delay and are subject to revisions. Relying on final vintage data would misrepresent the forecaster's information set, see e.g. Croushore (2006). Real-time data of CEI is obtained from The Conference Board. The four component variables are obtained from the data set of Chauvet and Piger (2008).[3] The data set is updated using the Philadelphia Fed's Real-Time Data Set for Macroeconomists (Croushore and Stark, 2001) for industrial production and employment. The most recent vintages for sales and personal income are taken from St. Louis Fed's ALFRED. For more details, see Appendix C.

The coincident economic index as well as its constituents are treated as integrated of order 1 and we transform them into annualized growth rates:

$$y_{t+h}^{h,t+h+1} = (1200/h)\log(\mathrm{Y}_{t+h}^{t+h+1}/\mathrm{Y}_{t}^{t+h+1}), \tag{1}$$

[3] To be precise, it is an updated version of the Giusto and Piger (2017) data set, which updates the Chauvet and Piger (2008) data set to 2013. Thanks to Jeremy Piger for uploading the raw data set on his website: `https://pages.uoregon.edu/jpiger/research/published-papers/`.

where Y_t^v is the original variable at time t from vintage v.

4.2 Quantile forecast

Based on the link with financial conditions and findings by Giglio et al. (2016) and Adrian et al. (2019), economic uncertainty is expected to mainly affect the left tail of the distribution of the coincident variables. To examine whether the predictive ability of uncertainty indeed varies across the distribution of output growth, we construct quantile forecasts.[4]

Before stating the forecasting model, there are two things to consider. First, following Giglio et al. (2016), we are interested in the quantiles of the shocks to the growth rates of economic activity rather than the growth rates themselves. These shocks are approximated by residuals from an autoregressive (AR) model with p lags. Second, the aim is to forecast output in real-time, emulating reality as close as possible. Therefore, we use the 'real-time vintage' approach (Koenig et al., 2003; Clements and Galvão, 2013), instead of using end of sample data. That is, we use the first release of the data for estimation when available, matching the release maturity of the leading observations on the left- and right hand side,[5]

$$y_{t+h}^{h,t+h+1} = \beta_0^h + \sum_{j=1}^{p} \beta_j^h y_{t-j+1}^{1,t+1} + u_{t+h}^{h,t+h+1}, \tag{2}$$

for $t = 1, \ldots, T-h$, where $y_t^{h,v}$ is defined in (1). The number of lags $0 \leq p \leq 6$ is selected using BIC. The lags on the right hand side are from the same vintage as the first lag, and can be lightly revised. For example, the second lag will be the second release of that observation. Then, after estimating (2), the shocks are defined as the first release

[4] In preliminary research, we also checked the ability of uncertainty measures to forecast the mean. The overall results were rather negative. There is little to no forecasting power.

[5] Equation (2) is slightly different for MTS, because there is a two month rather than a one month reporting lag. So to only use data available at the time of forecasting, the vintage is $t + 2$ instead of $t + 1$. The equation for MTS becomes

$$y_{t+h}^{h,t+h+2} = \beta_0^h + \sum_{j=1}^{p} \beta_j^h y_{t-j+1}^{1,t+2} + u_{t+h}^{h,t+h+1}.$$

residuals $\hat{u}_{t+h}^{h,t+h+1}$.

Quantile regression is a semiparametric method dating back to the seminal work by Koenker and Bassett (1978). The estimate of α-quantile q for variable y is the solution to the optimization

$$Q_\alpha(y) = \arg\inf_q \mathrm{E}\left[\rho_\alpha\left(y - q\right)\right], \tag{3}$$

where $\rho_\alpha(x) = \big(\alpha - \mathbf{1}(x \leq 0)\big)x$ the tick loss function, and we specify q as a linear function of (exogenous) regressors. Then, the α-quantile forecasts can be written as

$$Q_\alpha(\hat{u}_{t+h}^{h,t+h+1}|\Omega_{t+1}) = \psi_{\alpha,0}^h + \boldsymbol{\psi}_{\alpha,1}^h \boldsymbol{w}_t^{t+1}, \tag{4}$$

with Ω_{t+1} the information set at time $t+1$, and $\boldsymbol{w}_t^{t+1}$ the set of regressors, so the uncertainty measures, or factors, or NFCI, at time t from vintage $t+1$. Most regressors are not revised, and can therefore be denoted without a vintage superscript. Financial data is available instantly, but this is not the case for the survey data or the forecast error based measures. To be consistent and to ensure that the information is available to the forecaster, we use lagged values for all uncertainty measures. Adding more than just the first lag of the regressors $\boldsymbol{w}_t^{t+1}$ did not yield better forecasts.[6]

For values of α, we focus on 0.2, but also analyze results for 0.1, 0.5 (the median), and 0.8. The parameters $\boldsymbol{\psi}_{\alpha,j}^h$ are estimated using the interior point algorithm. For a review on quantile forecasting, see Komunjer (2013).

4.3 Models

As regressors in (4), we consider the following variables. First, the predictive ability of each uncertainty measure is considered individually, that is $\boldsymbol{w}_t^v = z_{i,t}$ is the i-th uncertainty measure at time t.

Second, we consider the factors extracted from the uncertainty measures using PCA,

[6] In particular, we considered adding up to three lags and selection using BIC, in line with literature on diffusion forecasting, see e.g. Stock and Watson (2002) or McCracken and Ng (2016).

$\boldsymbol{w}_t^v = \boldsymbol{f}_t$, the vector of k uncertainty factors at time t. Following the results in Section 3, we consider models with a fixed number of $k = 1$ up to 3 factors. This allows us to assess the relevance of adding a second or third factor, and compare against the same model over time. We refrain from estimating the number of factors k because methods of Bai and Ng (2002), Onatski (2010) and Ahn and Horenstein (2013) are for larger panels (larger N) than in our case and the alternative of a rank test (Cragg and Donald, 1997; Kleibergen and Paap, 2006) is not suitable for covariance matrices, see Donald et al. (2007).

Third, we compare the results from the uncertainty measures and factors to a model with NFCI as predictor, with $\boldsymbol{w}_t^v = \text{NFCI}_t^v$ the end-of-month NFCI at time t from vintage v. Real-time data for NFCI is from ALFRED, and available from May 2011 onwards. Before that, we use data from the oldest vintage (25/05/2011). The impact of not using real-time date in the first periods is likely to be minor. An inspection of the absolute size of revisions shows that they are small, and part of the revisions can be explained by the fact that each vintage is standardized.

Finally, as a benchmark, the performance of the models is compared to the historical quantile estimate $\hat{q}_{\alpha,t}$, the empirical quantile based on data up to and including time t.[7]

4.3.1 Sample

In the forecasting exercise, we recursively estimate all models. That is, at each time t, we first estimate the factors and models using data from 1989M12 to time $t - h$. Earlier (initial) observations are included if the lag order is larger than one. We start in 1989M12 because that is the first period where at least three months of data is available for all variables. Recursive estimation is in line with other diffusion forecasting literature, see e.g. Stock and Watson (2002) and McCracken and Ng (2016). Using all available information improves convergence of the factor estimates. Moreover, results using a rolling window did not indicate the presence of a structural break, while the forecasting

[7]As another benchmark, we considered a factor model with factors from the FRED-MD dataset (McCracken and Ng, 2016) as predictors, where real-time data is available from the 1999M08 vintage and the number of factor is selected using BIC, with a maximum of 8 factors. The FRED-MD factors' predictions are a bit disappointing at the lower quantiles and oftentimes don't even beat the historical quantile. The results are included in Appendix D.

results deteriorate in some cases.

Second, the parameter estimates and time t observations are used to construct the forecast for the $t+h$ value $y_{t+h}^{h,t+h+1}$. We imagine a forecaster, who starts forecasting in January 2000. The first forecast is made for period 1999M12 + h, and the final one for 2021M12, for horizons h of 1 (nowcast), 3, 6, 12, and 24 months. This yields a sample of 120 initial in-sample and $265-h$ out-of-sample observations.

4.4 Evaluation

A relevant question with real-time data is which values to use as 'actuals' to evaluate the forecasts. Preferably, these are the true values that are no longer revised. This is impossible however, because of benchmark revisions. For example, due to a change of the index year, annual updating following the consensus numbers, and redefinitions or measurement changes. Broadly, there are three alternatives.

One option is to use the x-th release observations $y_{t+h}^{h,t+h+x}$ for some $x \geq 1$. Many empirical studies use x-th release data to evaluate their forecasts, see e.g. Romer and Romer (2000), Groen et al. (2013) and D'Agostino et al. (2013). Selecting x requires some knowledge on the revision process. For quarterly data the second revision (third release) is often used because this is usually the 'final' revision from the statistical agency.

A second option is to use the final vintage observations $y_{t+h}^{h,T+1}$. The final vintage is the most recent publication of the numbers. For example Koenig et al. (2003) and Clements and Galvão (2013) use the vintage published about a year and a half after the end of their sample. An advantage is that it incorporates the latest available information and are currently closest to the true values as a single time series.

The third option is to use the prebenchmark observations $y_{t+h}^{h,\mathrm{PBM}}$ as actuals. Prebenchmark values are the final observation before the first benchmark revision after a first value for a given date has been reported. We choose to use this third option. In contrast to regular non-benchmark revisions, benchmark revisions can and should not be predictable to the forecaster (Croushore, 2006). An additional argument to use prebenchmark values as actuals is that the actuals should represent the forecasters' target rather than be the closest possible to the current truth. Both x-th release and final vintage observations are subject to benchmark revisions. Some prebenchmark observations are still subject to regular revisions though.

To evaluate the quantile forecasts, the prebenchmark residuals that we use as actuals are defined as

$$\hat{u}_{t+h}^{h,PBM} = y_{t+h}^{h,PBM} - \left(\hat{\beta}_0^h + \sum_{j=1}^{p} \hat{\beta}_j^h y_{t-j+1}^{1,t+1} \right), \tag{5}$$

with p the same value as in Equation (2) and $\hat{\beta}_j^h$ estimated from the model in Equation (2). Because the estimated model using first vintage data is an efficient estimate of the actual mean (Koenig et al., 2003), we can use it to 'demean' the actuals too to get the actual shocks. So the mean is consistent across estimation and evaluation.

4.4.1 Evaluation measures

The relative forecast accuracy of the quantile forecasts is evaluated by comparing the mean tick loss (MTL). Statistical significance is tested using one-sided Diebold and Mariano (1995) tests, where we test the null of equal predictive accuracy, versus the alternative of smaller loss compared to the benchmark model. The Diebold-Mariano (DM) test is defined for a general loss function. The tick loss function can be used to compare quantile forecasts.

Some econometric difficulties arise because of our setup of comparing (partially) nested models, estimated using an expanding window with real-time data, see Clark and McCracken (2013) for an overview. Clark and McCracken (2009) derive the limiting distribution of tests of equal predictive accuracy when data is subject to revisions. Their setting ignores benchmark revisions, which is in line with our data as we use prebenchmark observations as actual values. However, their test is for comparing predictive accuracy in population, while we are interested in the finite sample performance. Recently, Amburgey and McCracken (2022) propose a finite sample correction when evaluating quantile forecasts with data subject to revisions. One difference to our setting is that we predict quantiles of shocks rather than the levels, though.

We follow the arguments by Faust and Wright (2013) by relying on the Monte Carlo evidence presented by Clark and McCracken (2013). Their simulation study shows that the Diebold-Mariano test statistic with standard normal critical values and the corrections by Harvey et al. (1997) yields satisfactory size, even for nested models.

Next to the relative performance, we test the absolute performance. This is typically done by inspecting the sequence of violations or hits, the observations that fall below the quantile forecast. For correct coverage, the number of hits should be approximately equal to what is expected from the quantile level. Further, the hits should not be forecastable. Therefore, we apply Engle and Manganelli's (2004) dynamic quantile (DQ) test. It tests the coverage conditional on Ω_t, the information set at time t. Define $e_{t+h} = \mathbf{1}(y_{t+h} \leq q_{t+h}) - \alpha$ the 'demeaned' hits for quantile level α, and the vector of k instruments $\boldsymbol{x}_t$, which are in the information set at time t. It may contain q_t or its lags, and lags of e_t for example. The null hypothesis is $\mathrm{E}[\boldsymbol{x}_t e_{t+h}] = 0$. The out-of-sample DQ test statistic is[8]

$$DQ_{OOS} = \boldsymbol{e}'\boldsymbol{X}\left(\boldsymbol{X}'\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{e}/\left(\alpha(1-\alpha)\right), \tag{6}$$

where $\boldsymbol{e} = (e_{T+h}, \ldots, e_{T+T_p+h})'$, and $\boldsymbol{X} = (\boldsymbol{x}'_T, \ldots, \boldsymbol{x}'_{T+T_p})'$, with T and T_p the number of in-sample and out-of-sample observations. DQ_{OOS} follows a χ^2 distribution with k degrees of freedom. We apply the test with two sets of instruments $\boldsymbol{x}_t$. First, $\boldsymbol{x}_t = 1$ for an unconditional coverage test. Second, $\boldsymbol{x}_t = (1, q_t)'$ for a conditional coverage test. The latter is equivalent to a Wald test on a quantile version of the Mincer-Zarnowitz regression.

[8]Equation (6) only applies when $h = 1$. For longer horizon forecasts the variance of $X'e$ is computed using a HAC covariance matrix with a rectangular kernel of width $h-1$ to correct for the dependence in overlapping forecasts. If the resulting covariance matrix is not PSD, it is computed using a Bartlett kernel of width $1.5h$. This is similar to how we compute the DM test statistic. For more information on backtesting quantiles, see Barendse et al. (2021).

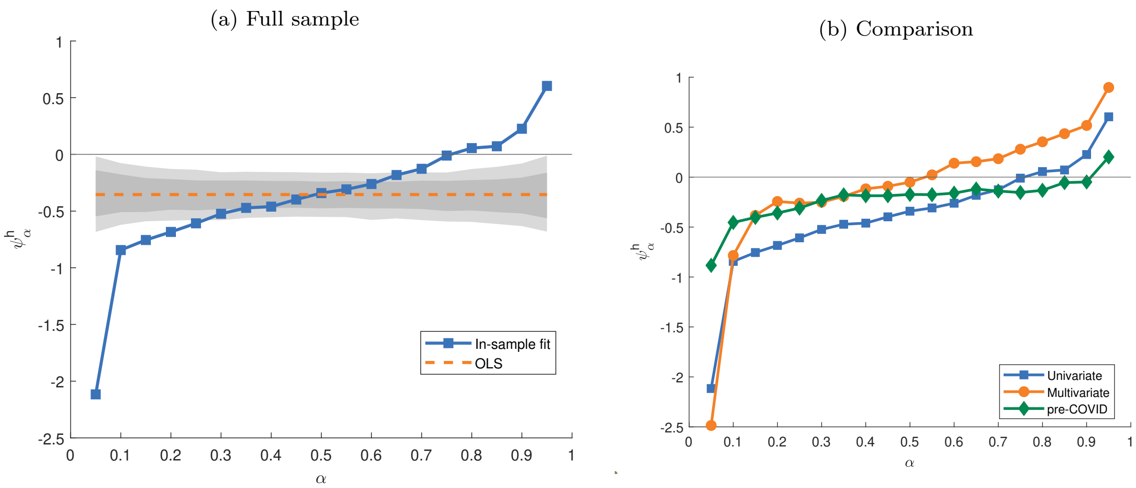
5 Full sample quantile regressions

Before turning to the forecasting exercise, we briefly consider in-sample quantile regressions of economic activity on the first uncertainty factor for evidence of a non-linear relationship. We perform quantile regressions based on the full sample, from 1990M1 to 2021M12, on the 0.05 to 0.95 quantiles.

The in-sample quantile estimates in Figure 3 show that there is substantial evidence that the relationship between economic activity and uncertainty is unlikely to be linear. In particular, the impact of the first uncertainty factor is stronger at the lower quantiles and is close to zero at higher quantiles. This holds not only for the coincident economic index, but also for quantile regressions on the components of the index. Moreover, Figure 3b shows that these results are robust to including NFCI as additional regressor, and considering only pre-COVID data.

The pattern over the quantiles is in line with how systemic risk affects quantiles of output shocks (Giglio et al., 2016). It provides empirical evidence to investigate the real-time out-of-sample predictive power of uncertainty measures for quantiles of shocks of the coincident economic index, with a focus on the lower quantiles.

Figure 3: In-sample quantile estimates for first uncertainty factor on CEI



The left figure presents the in-sample quantile estimates for the full sample (1990M1–2021M12) of the first uncertainty factor on the coincident economic index, at the 3 month forecasting horizon. The shaded areas are bootstrapped confidence bounds at the 90% and 95% level for a linear model based on 1000 bootstrap samples. The right figure presents the in-sample quantile estimates of the first uncertainty factor on the coincident economic index for the univariate regression, the multivariate regression (where NFCI is included as additional regressor), and the univariate regression using only the pre-COVID sample (up to 2019).

6 Forecasting results

To assess the forecasting performance per quantile, we compute the mean tick loss (MTL) relative to the historical quantile's mean tick loss (RMTL). This means that if the RMTL is below 1 for a model, it has a smaller loss and thus yields on average better quantile forecasts than the historical quantile benchmark.

We start off by looking at the performance in forecasting the coincident economic index, the aggregate business cycle indicator. The RMTLs from predicting the coincident economic index are plotted in Figure 4 for different horizons and quantiles. It is evident that the NFCI beats the models with uncertainty measures and factors in most cases. For now though, we turn our attention to the uncertainty measures and leave the comparison with NFCI to Section 7.

6.1 Uncertainty measures

Figure 4a immediately shows that the models yield better forecasts at lower quantiles than at higher quantiles. The average RMTL over the uncertainty measures is 0.962 for quantile 0.1 and all but 2 measures yield an RMTL below 1, while it is on average 1.007 at quantile 0.5 with only 4 out of 15 measures outperforming the benchmark.

A few uncertainty measures stand out in terms of predictive power: OVX, VIX and the JLN measures. First, the model with OVX as regressor yields the smallest tick loss for the uncertainty measures at the short horizon (up to 3 months) for the quantiles 0.1 and 0.2. At the 3 month horizon, the model with OVX significantly outperforms the historical quantile by 12.7% for quantile 0.1 and by 6.2% for quantile 0.2.

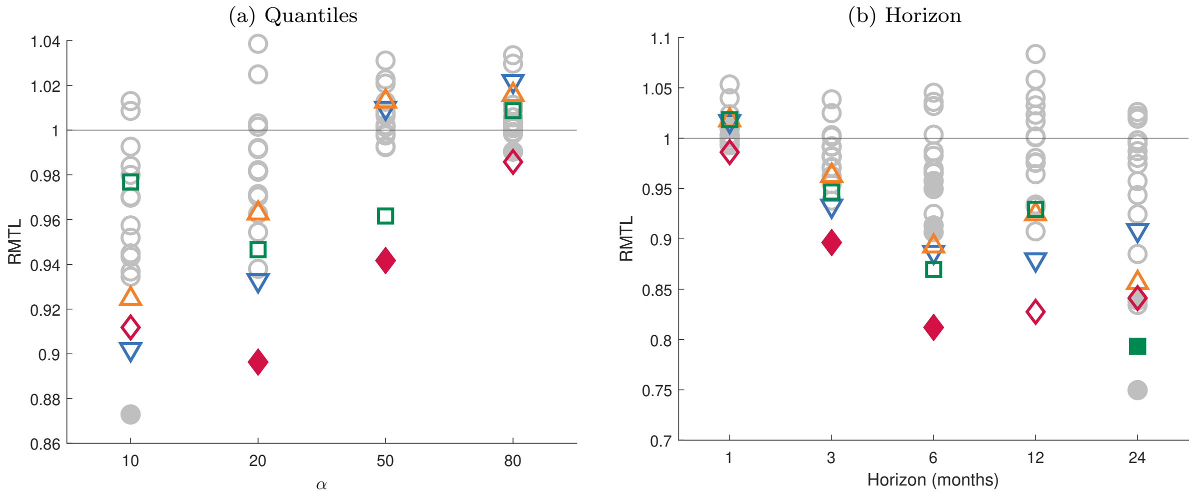
Second, the VIX yields good forecasts at the medium horizon, with p-values under 10% between 3 and 12 month horizons at quantile 0.2. The gains in terms of tick loss are between 3% (3 month horizon) and 9% (6 month horizon) compared to the historical quantile.

Third, the JLN measures yield good predictions at the short to medium-long horizon, from 3 months to 12 (and even 24) months ahead. JLNm has a slightly lower loss than

JLNf with 0.935 compared to 0.945 at quantile 0.1 and 0.954 to 0.970 at quantile 0.2 for the 3 month horizon forecast, but no significant improvement over the benchmark – the p-values are between 20% and 30%. JLNf, on the other hand, does yield significant gains, with p-values of 6.1%, 2.6% and 7.4% at quantile 0.2 for the 3, 6 and 12 month horizon, respectively.

Models with other uncertainty measures do not perform as well consistently. Uncertainty measures that perform surprisingly poorly are EPU and EPU+. Despite their popularity, they do not yield better forecasts than the benchmark in most of the cases when predicting CEI.

Figure 4: RMTL from forecasting coincident economic index



The figures present the relative mean tick loss (RMTL), with the historical quantile as benchmark, from forecasting the coincident economic index for multiple quantiles and forecast horizons. The forecasting horizon for the different quantiles is 3 months. The quantile for the different horizons is $\alpha = 0.2$. Gray circles are models with a single uncertainty measure. Blue down-pointing triangles, orange up-pointing triangles, and green squares are models with one, two and three uncertainty factors, respectively. Red diamonds are models with NFCI. Filled symbols indicate significance of the one-sided DM test against the historical quantile at the 5% significance level. Significance testing is not done for the recession and expansion subsets, because this is not a single consecutive period, required for the kernel for estimating the HAC standard errors. The RMTL values are presented Appendix D in tables.

6.2 Uncertainty factors

Turning to the uncertainty factor models, Figure 4 shows that they generally perform on par with the best or better than the best models where a single individual uncertainty measure is included. In particular at the medium horizon, the uncertainty factor models perform well, with gains in tick loss of 11.3% and 12.1% compared to the benchmark at the 6 and 12 month horizon for the 0.2 quantile. The forecasts are (close to) significantly better than the benchmark with p-values of 6.3% (6 month horizon) and 6.7% (12 month horizon). Importantly, the gains when using a factor are more consistent across the horizon and (lower) quantiles when compared to the individual uncertainty measures.

While the number of factors needed to explain most of the commonality in the uncertainty measures seems to be 2, this is not necessarily the number that yields the best forecasts. In the case of the coincident economic index, the second factor does not hold much additional predictive power over the first factor. In fact, the two factor model actually leads to a slightly higher tick loss when forecasting CEI, although the differences are small.

6.3 Tick loss over time and impact of COVID-19

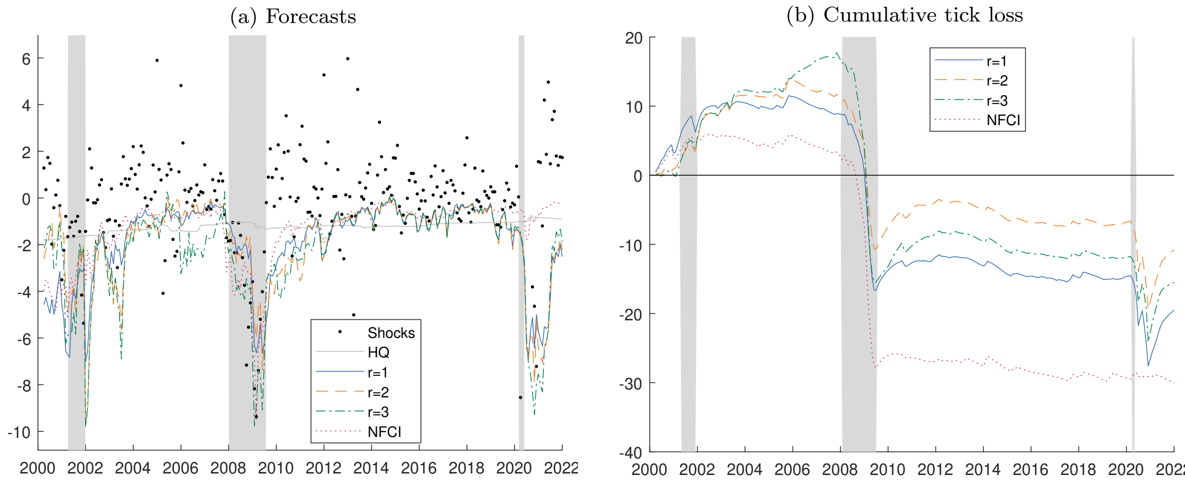
Perhaps unsurprisingly, most of the gains compared to the benchmark are achieved during recessionary periods and specifically during the financial crisis in 2008, see Figure 5a. This holds for both the models with individual uncertainty measures and the factor models. The historical quantile does not capture the downturns. The quantile is overestimated, and due to the tick loss function, this leads to a relatively large loss. Outside of recessions, the historical quantile is hardly outperformed by the models under consideration.

The COVID-19 period differs from the other recessions in that the underlying reason is not economic and therefore hard to predict using economic uncertainty and financial conditions.[9] Indeed, including 2020 and 2021 substantially increases the MTL for all models. At the 3 month horizon for quantile 0.2, the MTL of the historical quantile increases from 0.566 to 1.105 in the case of the coincident index. Other models are impacted even more, as the relative tick loss is also worse than compared to the pre-COVID period. Again for the 3 month horizon and the 0.2 quantile, the RMTL increases with 1.1 percentage points on overage over all measures when including 2020 and 2021 in the sample period. Still, the uncertainty factor models at least somewhat capture the downturn. This is not true for NFCI, see Figure 5a, which leads to an

[9]For discussions on how to treat the COVID-19 period when modelling or forecasting macroeconomic variables, see e.g. Carriero et al. (2021), Lenza and Primiceri (2022) and Schorfheide and Song (2021).

increase in RMTL from 0.782 to 0.896. Ultimately though, the ordering of the models' performance and the significance levels of the forecasting performance are not affected much.

Figure 5: Tick loss for coincident index over time



The left figure presents the realized values and the quantile forecasts for the coincident economic index at quantile 0.2 for a 3 month horizon from the historical quantile (solid gray line), factor models (with r factors) and the model with NFCI. The right figure presents the cumulative tick loss over the evaluation period minus the cumulative tick loss from the historical quantile.

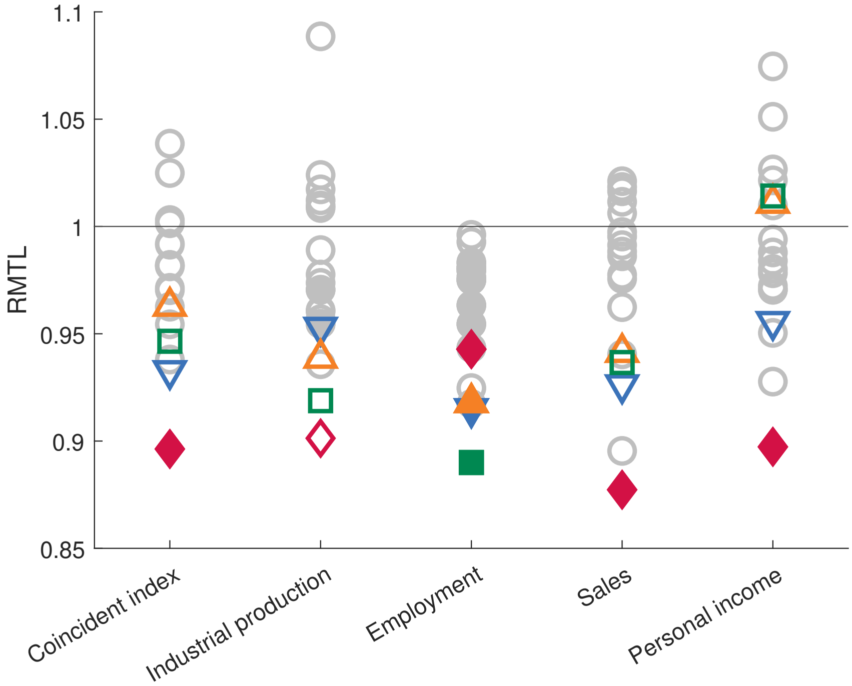
6.4 Forecasting the components of the coincident index

The findings in forecasting the components of the coincident index (industrial production, employment, manufacturing and trade sales, and personal income) are largely consistent with the index' forecasting results. The most important and consistent finding is that uncertainty measures and factors outperform the historical quantile when forecasting the left tail, see Figure 6 for an example at the 3 month horizon.

The forecasting results are most convincing for employment. At the 3 month horizon, for quantile 0.2 all uncertainty measures beat the historical quantile, and the uncertainty factors render the best performance with improvements of 8.2% to 11% in mean tick loss compared to the benchmark. Moreover, the difference in tick loss is significant for all factor models at the 5% significance level. The uncertainty measures that yield significant gains are mostly those based on financial information: VIX, MOVE, CSDR(sic), MPU, and JLNf. This is in line with Bloom (2009), who considers the effect of uncertainty on the labor market using the VIX (and CSDR(sic)).

The gains are similar for the 0.1 quantile forecasts of employment, see Figure 7a, while there is little predictive power at the middle and higher quantiles. Further, Figure 7b shows that the factor models perform best at the medium forecasting horizon of 3 to 12 months.

Figure 6: Tick loss across target variables



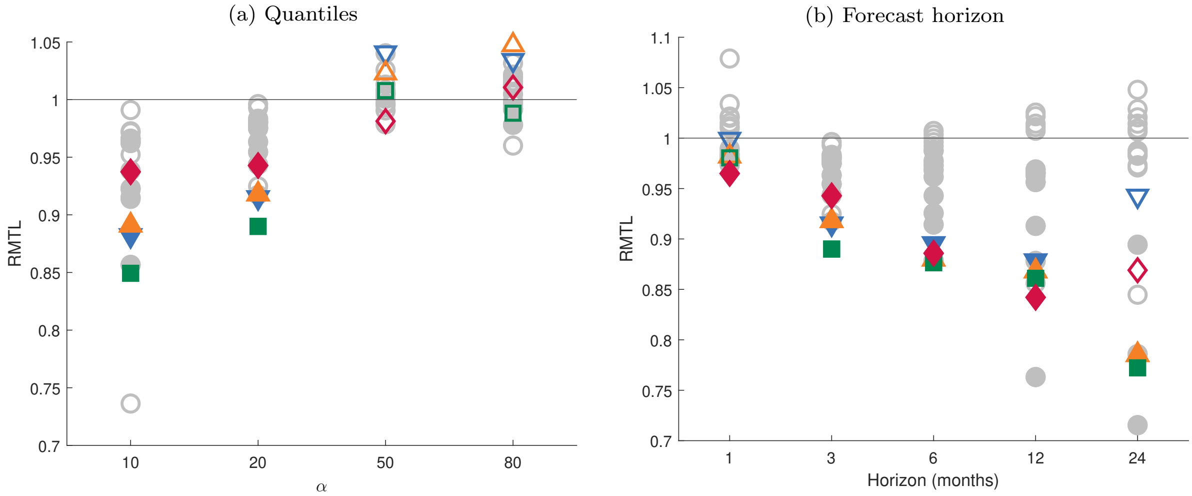
The figure presents the relative mean tick loss (RMTL), with the historical quantile as benchmark, from forecasting the coincident economic index and its component: industrial production, employment, manufacturing and trade sales, and personal income. The forecasting horizon is 3 months. The quantile is $\alpha = 0.2$. Gray circles are models with a single uncertainty measure. Blue down-pointing triangles, orange up-pointing triangles, and green squares are models with one, two and three uncertainty factors, respectively. Red diamonds are models with NFCI. Filled symbols indicate significance of the one-sided DM test against the historical quantile at the 5% significance level. The RMTL values are presented Appendix D in tables.

6.5 Coverage

Next, we evaluate the absolute performance of the quantile forecasts by means of Engle and Manganelli's (2004) DQ tests to verify if the coverage is in line with the expected level, and whether the hits – observations smaller than the predicted quantile – are not forecastable.

In general, the hit rates in Table 2 indicate that the coverage is good. Hit rates for the individual uncertainty measures and other target variables are in Appendix E. At shorter horizons, they are slightly below the expected level of 0.2, but the null of correct coverage

Figure 7: Tick loss in forecasting employment



The figures present the relative mean tick loss (RMTL), with the historical quantile as benchmark, from forecasting non-farm payroll employment for multiple quantiles and forecast horizons. The forecasting horizon for the different quantiles is 3 months. The quantile for the different horizons is $\alpha = 0.2$. Gray circles are models with a single uncertainty measure. Blue down-pointing triangles, orange up-pointing triangles, and green squares are models with one, two and three uncertainty factors, respectively. Red diamonds are models with NFCI. Filled symbols indicate significance of the one-sided DM test against the historical quantile at the 5% significance level. The RMTL values are presented Appendix D in tables.

is not rejected for most models. As the forecast horizon increases, the hit rates increase and match expectations best at the 3 to 12 month horizon. At the 24 month horizon, the number of hits is often too high. The coverage of 1 month horizon employment quantile forecasts is low, with hit rates from 0.102 (EPU+) to 0.155 (CSDRsic), and correct coverage is rejected by the DQ test for most models. Again, at medium horizons the coverage is up to the expected level. So in short, the models are well specified at horizons up to a year, but some caution is advised when forecasting longer horizons.

7 Economic uncertainty and financial conditions

Adrian et al. (2019) show that the NFCI has predictive power for the left tail of the US GDP growth rate. Additionally, Adams et al. (2021) find that financial conditions improve forecasts of employment, industrial production and inflation. Financial conditions are closely related to economic uncertainty, or at least capture part of it. The NFCI is constructed to only reflect the financial conditions, not the general economic conditions. Therefore, it is interesting to compare with the uncertainty measures.

Table 2: Hit rates

Horizon (months)	1	3	6	12	24
Panel A: Coincident economic index					
HQ	0.182	0.195	0.197	0.237	0.336 ‡
NFCI	0.186	0.179	0.189	0.304 ‡	0.390 † ‡
Factor models					
$r = 1$	0.182	0.168	0.174	0.229	0.349
$r = 2$	0.182	0.187	0.216	0.320 † ‡	0.344 † ‡
$r = 3$	0.163	0.134 † ‡	0.178	0.273	0.311 ‡
Panel B: Employment					
HQ	0.136 † ‡	0.179	0.201	0.281	0.344 ‡
NFCI	0.114 † ‡	0.164	0.193	0.293	0.423 † ‡
Factor models					
$r = 1$	0.114 † ‡	0.172	0.185	0.269	0.365 †
$r = 2$	0.117 † ‡	0.179	0.216	0.308	0.378 † ‡
$r = 3$	0.125 † ‡	0.164 ‡	0.197	0.300	0.398 † ‡

The table presents hit rates for various forecasting horizons, for the full sample and quantile $\alpha = 0.2$. The † denotes rejection of the null hypothesis of correct unconditional coverage, and ‡ denotes rejection of the null hypothesis of correct coverage conditional on an intercept and the quantile estimates q_t, all at a 5% significance level, based on the DQ test with HAC standard errors.

The correlations between the uncertainty measures and NFCI are all positive. As expected, it is quite strongly correlated with the financial measures (69.8% with VIX, 53.3% with MOVE, and 63.4% with JLNf). The strongest correlation pre-COVID was actually with the forecasting error based measures JLNm (84.7%) and JLNr (71.9%), but the response to COVID-19 was so different that the correlations drop to 54.9% and 20.3% when including 2020 and 2021. Still positive but substantially smaller. It is interesting that the correlation with JLNm is high since JLNm ought to describe macro uncertainty, while NFCI captures financial conditions unrelated to other economic conditions. The correlation with consumer survey based and news-based measures are more modest and in the range of 24.2% to 47.0%.

Next comparing to the uncertainty factors, Figure 1 shows that the NFCI closely resembles the first factor. There are some deviations – for example the 2001 recession is not captured by the NFCI and the impact of COVID-19 is relatively small – but the general pattern is very similar. The correlation between the first factor and NFCI is quite high at 62.8%, and even 82.4% if we exclude 2020 and 2021. This is not surprising given the correlations with the individual uncertainty measures, and that the first factor loads somewhat more on the uncertainty measures based on financial information. The correlation of NFCI with the second factor is negative and moderate at -38.5% (-15.9% pre-COVID).

Section 6.2 shows good forecasting performance of the uncertainty factors. The question from the correlations is how this compares to NFCI, whether the predictive power is due to the relationship with NFCI, or whether it reflects additional relevant information that is not captured by the NFCI.

7.1 Forecasting comparison

From the RMTL plots in Figures 4, 6 and 7 it is clear that NFCI is a very strong competitor for the uncertainty measures and factors, as expected from the findings by Adrian et al. (2019) and Adams et al. (2021). The model with NFCI often has the smallest tick loss. When forecasting employment, the uncertainty factor models do have an edge over NFCI. At the 3 month horizon for the 0.2 quantile, the uncertainty factor models have a 2.9% (one factor) to 5.6% (three factors) lower tick loss compared to the NFCI – according to the DM tests there is no significant improvement though.

Even if NFCI outperforms the uncertainty measures and factor models most of the time, there may still be relevant information in the uncertainty measures that is not present in the NFCI. Table 3 presents the results from the Giacomini and Komunjer (2005) encompassing test.[10] The table shows whether none, both or only one of the variables should be included. In the latter case, one encompasses the other and that variable is preferred to use. Where it says both, a combination of the quantiles is better than only one of the two.

Two conclusions follow from the encompassing tests. First, the NFCI encompasses the uncertainty factor model in many cases, especially at short horizons of 1 and 3 months. This is somewhat contrary to Hengge (2019), who finds using the predictive score that the predictive power of JLNm is not impacted by including NFCI. Second, Table 3 shows

[10] The unconditional version. To allow for misspecification, we implement the test with an intercept and allow the competing quantile's coefficient to differ from 1. That is, the combined quantiles are $\hat{q}_{ct} = \theta_0 + \theta_1 \hat{q}_{1t} + \theta_2 \hat{q}_{2t}$, where $\hat{q}_{1t}$ and $\hat{q}_{2t}$ are the quantile predictions from the competing models, and we test whether $\theta_1 = 0$ and/or $\theta_2 = 0$. Standard errors are computed using a Newey-West estimator with $h - 1$ kernel width, following Giacomini and Komunjer (2005).

that it is valuable to add uncertainty factors when forecasting at medium horizons of 6 and 12 months. For employment, there is even evidence that the uncertainty factor models encompasses NFCI.

Table 3: Preferred model from encompassing test

	Horizon (months)				
Target variable	1	3	6	12	24
Coincident index	NFCI	NFCI	Both	Both	NFCI
Industrial production	NFCI	NFCI	Both	Both	Uncertainty
Employment	Both	Uncertainty	Both	Both	Uncertainty
Manufacturing and trade sales	NFCI	NFCI	NFCI	Both	Uncertainty
Personal income	NFCI	NFCI	Both	Both	NFCI

The table presents what model is preferred, based on the Giacomini and Komunjer (2005) encompassing test at a 5% significance level, for forecasts of quantile $\alpha = 0.2$. The competing models are the factor model with the two first uncertainty factors and the model with NFCI.

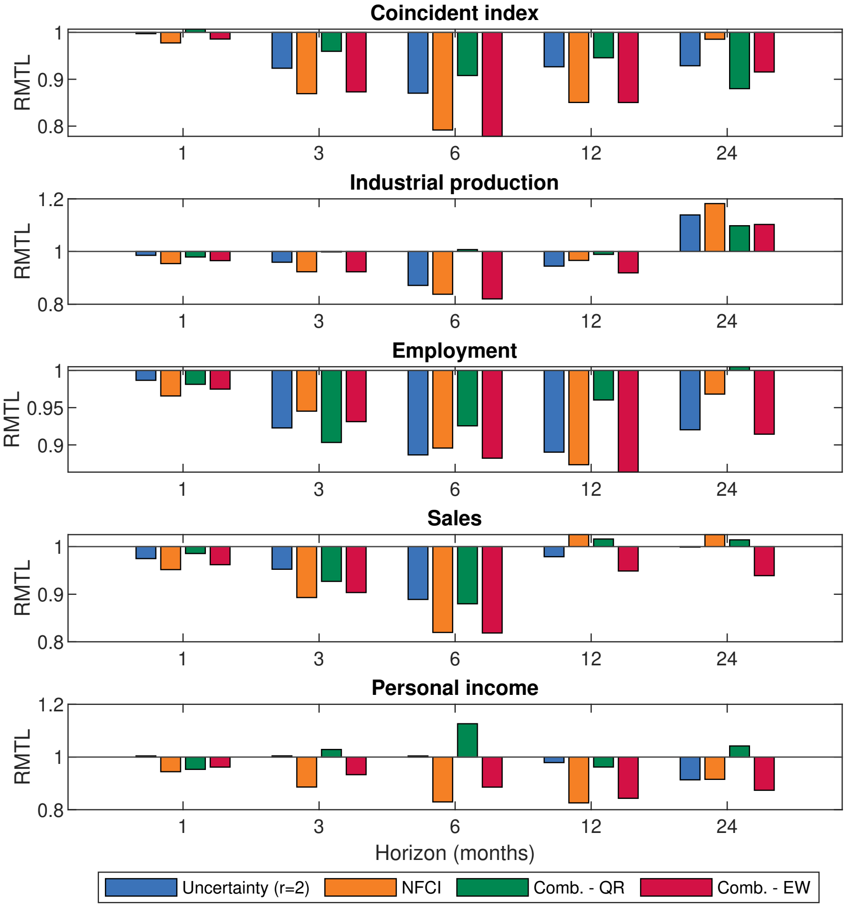
To verify the encompassing results, we combine the quantile forecasts from the model with NFCI and the uncertainty factor model in two ways. First, we include both forecasts as regressors in a quantile regression including an intercept and estimate the weights recursively – the evaluation sample starts at 2002M12+h, to allow for a burn-in period for the estimated combination weights. This specification corresponds to the model in the encompassing test. Second, we also consider an equally weighted combination, which is known to work well in general (see e.g. Timmermann, 2006).

The encompassing results are largely corroborated by the performance of the forecast combinations, see Figure 8. As expected from Table 3, the model with only NFCI is the best model at the short horizons of 1 and 3 months, except for employment at the 3 month horizon. The relative performance of the uncertainty two factor model does deviate a bit from the encompassing results. When forecasting employment at the 3 month horizon there is a small gain from combining with NFCI forecasts: the RMTL decreases from 0.923 to 0.902.

At longer horizons, combining forecasts yields a smaller RMTL, also in agreement with the encompassing test results. However, this is only when considering an equally weighted combination of the uncertainty factor model and NFCI model forecasts. The forecast combination based on quantile regression is actually the worst model in most cases. On average, the RMTL at the 6 and 12 month horizon is 13.3 and 9.0 percentage points smaller when using equal weights instead of estimated weights. This is in line with the known robustness of equally weighted forecast combinations (Timmermann, 2006).

So, economic uncertainty does hold relevant predictive information that is not captured by the financial conditions.

Figure 8: RMTL of forecast combinations



The figure presents the relative mean tick loss (RMTL), with the historical quantile as benchmark, from forecasting the coincident economic index and its component: industrial production, employment, manufacturing and trade sales, and personal income. The quantile is $\alpha = 0.2$. Models include the uncertainty factor model with $r = 2$ factors (blue bars), the model with NFCI as regressor (orange bars), and two forecast combinations that combines the forecasts from the 2 uncertainty factor model and the model with NFCI, where the quantiles forecasts are weighted by recursively estimating a quantile regression (QR, green bars) or equally weighted (EW, red bars). The evaluation period is 2002M12+h–2021M12, to allow for a two year burn-in period for the weights in the forecasting combination.

7.2 Financial and non-financial based uncertainty measures

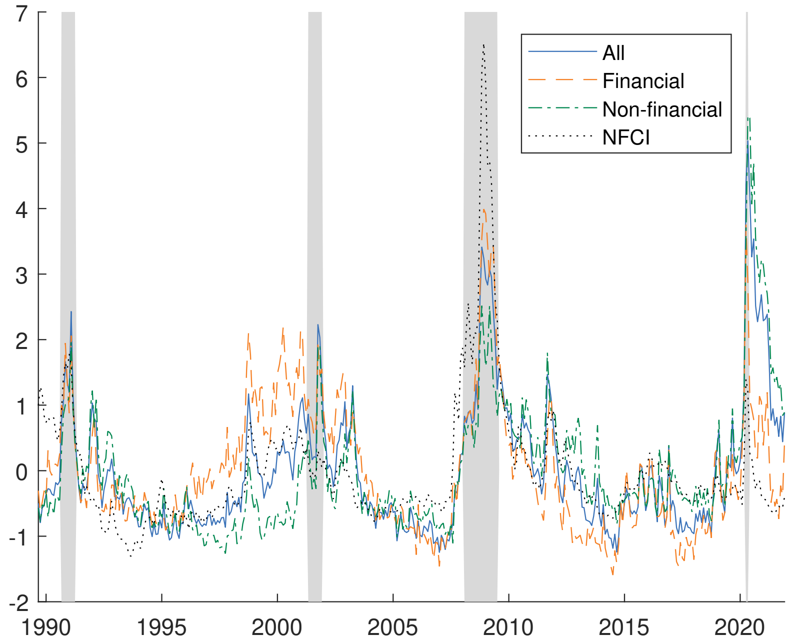
The comparison with NFCI raises the question what information is relevant in forecasting the coincident economic index and its components, whether this is related to the financial conditions or macroeconomic uncertainty. Ludvigson et al. (2021) also differentiate between macroeconomic and financial uncertainty. They find that high financial uncertainty is a cause of real activity shocks, while increases in macroeconomic uncertainty are a response to it. Therefore, we split the uncertainty measures into two groups: those based on financial information (VIX, MOVE, OVX, CSDR, CSDRsic and JLNf) and those based on non-financial information. Then, we compute the first factor from the principal components analysis for each group separately.

At first sight the first factors from both subsets of uncertainty measures seem to be very similar, see Figure 9; with correlations of 80.8% and 90.7% with the first factor obtained from the the full set. The correlation between the factors of the subsets is only 48.7% though. Where the factors deviate most is during the period 1996–2001 – likely due to the Asian and Russian financial crises and the dot-com bubble. Also, the factor from financial-based uncertainty measures attains a higher value during the financial crisis in 2007–2008, and responds less strongly to the COVID-19 crisis compared to the factor from non-financial uncertainty measures. As expected, this pattern for the factor from financial-based uncertainty measures is more in line with the NFCI and is also seen in the correlations: 71.8%, which is higher than the factor based on the full set of measures (62.8%) or the non-financial based measures (41.9%).

The second factor from the full set is not captured well by the factors from the subsets. The largest (absolute) correlation is 56.8%, with the first factor from the financial subset, and 40.8% for the non-financial subset.

Since the NFCI outperforms the uncertainty factor models in most cases, we would expect that the factors based on the financial-based uncertainty measures outperform the other factors models too. Table 4 shows that this is partly true. Using the first financial-based uncertainty factor leads to a smaller MTL compared to using the first factor based on non-financial information. However, neither outperforms the factor model based on the full set of uncertainty measures. Except for industrial production, where the single factor model based on the financial uncertainty measures performs close to the model with NFCI.

Figure 9: First factor from subsets of measures



The figure presents the time series of the first factor from the principal components analysis based on the full set of uncertainty measures (solid blue line), the subset of financial-based uncertainty measures (dashed orange line), and the subset of non-financial-based uncertainty measures (dash-dotted green line) and the NFCI (dotted black line). The NFCI are the end-of-month values from the 16/02/2022 vintage. The gray bars are recessions as determined by NBER's Business Cycle Dating Committee. All series are standardized.

Interestingly, the second factor contains important predictive power when considering the non-financial set, in particular when forecasting employment. Even in the first ten years of the sample, from January 2001 to December 2009, the non-financial factor models yield the smallest tick loss with an RMTL of 0.850 (one factor) and 0.816 (two factors) when forecasting the coincident index. The second factor loads heavily on FDISP and MPU. Individually, those measures don't decrease the MTL by much compared to the benchmark, up to 3.2% for a 12 month horizon or less, though the gains are strongly significant.

So it seems that most relevant information comes from the financial uncertainty measures, but not all predictive power.

Table 4: RMTL for factor models on subsets

Horizon (months)	1	3	6	12	24
Panel A: Coincident economic index					
Financial information based measures					
$r = 1$	1.011	0.977	0.925**	0.929**	0.947
$r = 2$	1.015	0.984	0.962	0.926**	0.905
Non-financial information based measures					
$r = 1$	1.042	0.974	0.963	0.978	0.973
$r = 2$	1.054	0.928	0.882	0.919	0.832*
Panel B: Industrial production					
Financial information based measures					
$r = 1$	0.983	0.920**	0.897**	0.956	1.035
$r = 2$	0.985	0.949	0.916*	0.912*	0.935
Non-financial information based measures					
$r = 1$	1.000	1.012	0.980	1.017	0.971
$r = 2$	0.953	0.924	0.825	0.822	0.808*
Panel C: Nonfarm payroll employment					
Financial information based measures					
$r = 1$	0.972***	0.954***	0.936***	0.916***	0.974
$r = 2$	0.975**	0.952**	0.921***	0.898***	0.958
Non-financial information based measures					
$r = 1$	1.040	0.962	0.955	0.967	0.971
$r = 2$	1.013	0.886**	0.867**	0.806**	0.744**
Panel D: Manufacturing and trade sales					
Financial information based measures					
$r = 1$	0.995	0.959	0.956	0.958	1.052
$r = 2$	1.001	0.989	0.979	1.022	1.101
Non-financial information based measures					
$r = 1$	1.024	0.968	0.966	1.029	0.939**
$r = 2$	1.018	0.929	0.863	0.925	0.847*
Panel E: Personal income excluding transfer receipts					
Financial information based measures					
$r = 1$	1.000	0.969	0.986	1.011	0.999
$r = 2$	0.998	0.996	1.021	1.037	0.967
Non-financial information based measures					
$r = 1$	0.982	1.028	1.060	0.999	1.006
$r = 2$	1.005	0.968	0.998	0.998	0.829

The table presents the relative mean tick loss for various forecasting horizons and target variables, for the full sample and quantile $\alpha = 0.2$. ***,**, and * denote significance of a one-sided Diebold-Mariano test at the 1%, 5% and 10%, respectively.

8 Conclusion

Many economic uncertainty measures have been proposed over the last 15 years. We show that they share a factor structure. The first common component explains over 40% of the total variation. The second factor can be interpreted as a media/consumer uncertainty, which tends to remain high after officially leaving recessions.

The results of our real-time forecasting analysis show that there is a non-linear relation between the uncertainty measures and factors and the coincident economic index, that can be utilized to forecast the lower quantiles of the index. The VIX, OVX and JLN measures are recommended individual measures, but using the factors is preferred for more consistent gains. The predictive power of economic uncertainty is relevant for professional forecasters and policy makers to keep an eye on, in particular when interested in the labor market, where economic uncertainty is more important than financial conditions.

References

Adams, P. A., Adrian, T., Boyarchenko, N., and Giannone, D. (2021). Forecasting macroeconomic risks. *International Journal of Forecasting*, 37(3):1173–1191.

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

Ahn, S. C. and Horenstein, A. R. (2013). Eigenvalue ratio test for the number of factors. *Econometrica*, 81(3):1203–1227.

Amburgey, A. and McCracken, M. W. (2022). On the real-time predictive content of financial conditions indices for growth. FRB St. Louis Working Paper.

Bachmann, R. and Bayer, C. (2014). Investment dispersion and the business cycle. *American Economic Review*, 104(4):1392–1416.

Bachmann, R., Elstner, S., and Sims, E. R. (2013). Uncertainty and economic activity: Evidence from business survey data. *American Economic Journal: Macroeconomics*, 5(2):217–49.

Bai, J. and Ng, S. (2002). Determining the number of factors in approximate factor models. *Econometrica*, 70(1):191–221.

Baker, S. R. and Bloom, N. (2013). Does uncertainty reduce growth? Using disasters as natural experiments. NBER Working Paper No. 19475.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *Quarterly Journal of Economics*, 131(4):1593–1636.

Barendse, S., Kole, E., and Van Dijk, D. (2021). Backtesting Value-at-Risk and expected shortfall in the presence of estimation error. *Journal of Financial Econometrics*.

Bekaert, G., Hoerova, M., and Duca, M. L. (2013). Risk, uncertainty and monetary policy. *Journal of Monetary Economics*, 60(7):771–788.

Bernanke, B. S. (1983). Irreversibility, uncertainty, and cyclical investment. *Quarterly Journal of Economics*, 98(1):85–106.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica*, 77(3):623–685.

Bloom, N. (2014). Fluctuations in uncertainty. *Journal of Economic Perspectives*, 28(2):153–175.

Carriero, A., Clark, T. E., and Marcellino, M. (2018). Measuring uncertainty and its impact on the economy. *The Review of Economics and Statistics*, 100(5):799–815.

Carriero, A., Clark, T. E., Marcellino, M., and Mertens, E. (2021). Addressing COVID-19 outliers in BVARs with stochastic volatility. *The Review of Economics and Statistics*.

Charles, A., Darné, O., and Tripier, F. (2018). Uncertainty and the macroeconomy: Evidence from an uncertainty composite indicator. *Applied Economics*, 50(10):1093–1107.

Chauvet, M. and Piger, J. (2008). A comparison of the real-time performance of business cycle dating methods. *Journal of Business & Economic Statistics*, 26(1):42–49.

Chauvet, M., Senyuz, Z., and Yoldas, E. (2015). What does financial volatility tell us about macroeconomic fluctuations? *Journal of Economic Dynamics and Control*, 52:340–360.

Clark, T. and McCracken, M. (2013). Advances in forecast evaluation. In Elliott, G. and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 2, pages 1107–1201. Elsevier.

Clark, T. E. and McCracken, M. W. (2009). Tests of equal predictive ability with real-time data. *Journal of Business & Economic Statistics*, 27(4):441–454.

Clements, M. P. and Galvão, A. B. (2013). Real-time forecasting of inflation and output growth with autoregressive models in the presence of data revisions. *Journal of Applied Econometrics*, 28(3):458–477.

Cragg, J. G. and Donald, S. G. (1997). Inferring the rank of a matrix. *Journal of Econometrics*, 76(1-2):223–250.

Croushore, D. (2006). Forecasting with real-time macroeconomic data. In Elliott, G. and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 1, pages 961–982. Elsevier.

Croushore, D. and Stark, T. (2001). A real-time data set for macroeconomists. *Journal of econometrics*, 105(1):111–130.

D'Agostino, A., Gambetti, L., and Giannone, D. (2013). Macroeconomic forecasting and structural change. *Journal of Applied Econometrics*, 28(1):82–101.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–263.

Diks, C., Panchenko, V., and Van Dijk, D. (2011). Likelihood-based scoring rules for comparing density forecasts in tails. *Journal of Econometrics*, 163(2):215–230.

Dixit, A. K. and Pindyck, R. S. (1994). *Investment under uncertainty.* Princeton, NJ: Princeton University Press.

Donald, S. G., Fortuna, N., and Pipiras, V. (2007). On rank estimation in symmetric matrices: The case of indefinite matrix estimators. *Econometric Theory*, 23(6):1217–1232.

Dovern, J., Fritsche, U., and Slacalek, J. (2012). Disagreement among forecasters in G7 countries. *The Review of Economics and Statistics*, 94(4):1081–1096.

Engle, R. F. and Manganelli, S. (2004). CAViaR: Conditional autoregressive Value at Risk by regression quantiles. *Journal of Business & Economic Statistics*, 22(4):367–381.

Fajgelbaum, P. D., Schaal, E., and Taschereau-Dumouchel, M. (2017). Uncertainty traps. *Quarterly Journal of Economics*, 132(4):1641–1692.

Faust, J. and Wright, J. H. (2013). Forecasting inflation. In Elliott, G. and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 2, pages 2–56. Elsevier.

Fernández-Villaverde, J., Guerrón-Quintana, P., Kuester, K., and Rubio-Ramírez, J. (2015). Fiscal volatility shocks and economic activity. *American Economic Review*, 105(11):3352–84.

Giacomini, R. and Komunjer, I. (2005). Evaluation and combination of conditional quantile forecasts. *Journal of Business & Economic Statistics*, 23(4):416–431.

Giglio, S., Kelly, B., and Pruitt, S. (2016). Systemic risk and the macroeconomy: An empirical evaluation. *Journal of Financial Economics*, 119(3):457–471.

Gilchrist, S., Sim, J. W., and Zakrajšek, E. (2014). Uncertainty, financial frictions, and investment dynamics. NBER Working Paper No. 20038.

Giusto, A. and Piger, J. (2017). Identifying business cycle turning points in real time with vector quantization. *International Journal of Forecasting*, 33(1):174–184.

Groen, J. J., Paap, R., and Ravazzolo, F. (2013). Real-time inflation forecasting in a changing world. *Journal of Business & Economic Statistics*, 31(1):29–44.

Groshen, E. L. and Potter, S. M. (2003). Has structural change contributed to a jobless recovery? *Current Issues in Economics and Finance*, 9(8):1–7.

Haddow, A., Hare, C., Hooley, J., and Shakir, T. (2013). Macroeconomic uncertainty: What is it, how can we measure it and why does it matter? *Bank of England Quarterly Bulletin*, 53(2):100–109.

Harvey, D., Leybourne, S., and Newbold, P. (1997). Testing the equality of prediction mean squared errors. *International Journal of forecasting*, 13(2):281–291.

Hengge, M. (2019). Uncertainty as a predictor of economic activity.

Inoue, A. and Kilian, L. (2005). In-sample or out-of-sample tests of predictability: Which one should we use? *Econometric Reviews*, 23(4):371–402.

Jaimovich, N. and Siu, H. E. (2020). Job polarization and jobless recoveries. *The Review of Economics and Statistics*, 102(1):129–147.

Jurado, K., Ludvigson, S. C., and Ng, S. (2015). Measuring uncertainty. *American Economic Review*, 105(3):1177–1216.

Kehrig, M. (2015). The cyclical nature of the productivity distribution. Available at SSRN: https://ssrn.com/abstract=1854401.

Kellogg, R. (2014). The effect of uncertainty on investment: Evidence from Texas oil drilling. *American Economic Review*, 104(6):1698–1734.

Kleibergen, F. and Paap, R. (2006). Generalized reduced rank tests using the singular value decomposition. *Journal of Econometrics*, 133(1):97–126.

Koenig, E. F., Dolmas, S., and Piger, J. (2003). The use and abuse of real-time data in economic forecasting. *The Review of Economics and Statistics*, 85(3):618–628.

Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica*, 46(1):33–50.

Komunjer, I. (2013). Quantile prediction. In Elliott, G. and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 2, pages 961–994. Elsevier.

Kozeniauskas, N., Orlik, A., and Veldkamp, L. (2018). What are uncertainty shocks? *Journal of Monetary Economics*, 100:1–15.

Lahiri, K. and Sheng, X. (2010). Measuring forecast uncertainty by disagreement: The missing link. *Journal of Applied Econometrics*, 25(4):514–538.

Leduc, S. and Liu, Z. (2016). Uncertainty shocks are aggregate demand shocks. *Journal of Monetary Economics*, 82:20–35.

Lenza, M. and Primiceri, G. E. (2022). How to estimate a vector autoregression after March 2020. *Journal of Applied Econometrics*, 37(4):688–699.

Ludvigson, S. C., Ma, S., and Ng, S. (2021). Uncertainty and business cycles: Exogenous impulse or endogenous response? *American Economic Journal: Macroeconomics*, 13(4):369–410.

McCracken, M. W. and Ng, S. (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics*, 34(4):574–589.

Onatski, A. (2010). Determining the number of factors from empirical distribution of eigenvalues. *The Review of Economics and Statistics*, 92(4):1004–1016.

Prasad, A., Elekdag, S., Jeasakul, P., Lafarguette, R., Alter, A., Feng, A. X., and Wang, C. (2019). Growth at Risk: Concept and application in IMF country surveillance. IMF Working Papers WP/19/36.

Rogers, J. H. and Xu, J. (2019). How well does economic uncertainty forecast economic activity? Finance and Economics Discussion Series 2019-085.

Romer, C. D. and Romer, D. H. (2000). Federal Reserve information and the behavior of interest rates. *American Economic Review*, 90(3):429–457.

Rossi, B., Sekhposyan, T., and Soupre, M. (2016). Understanding the sources of macroeconomic uncertainty. Available at SSRN: https://ssrn.com/abstract=2780213.

Schorfheide, F. and Song, D. (2021). Real-time forecasting with a (standard) mixed-frequency VAR during a pandemic. NBER Working Paper No. 29535.

Stock, J. H. and Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association*, 97(460):1167–1179.

Timmermann, A. (2006). Forecast combinations. In Elliott, G. and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 1, pages 135–196. Elsevier.

Vavra, J. (2013). Inflation dynamics and time-varying volatility: New evidence and an Ss interpretation. *Quarterly Journal of Economics*, 129(1):215–258.

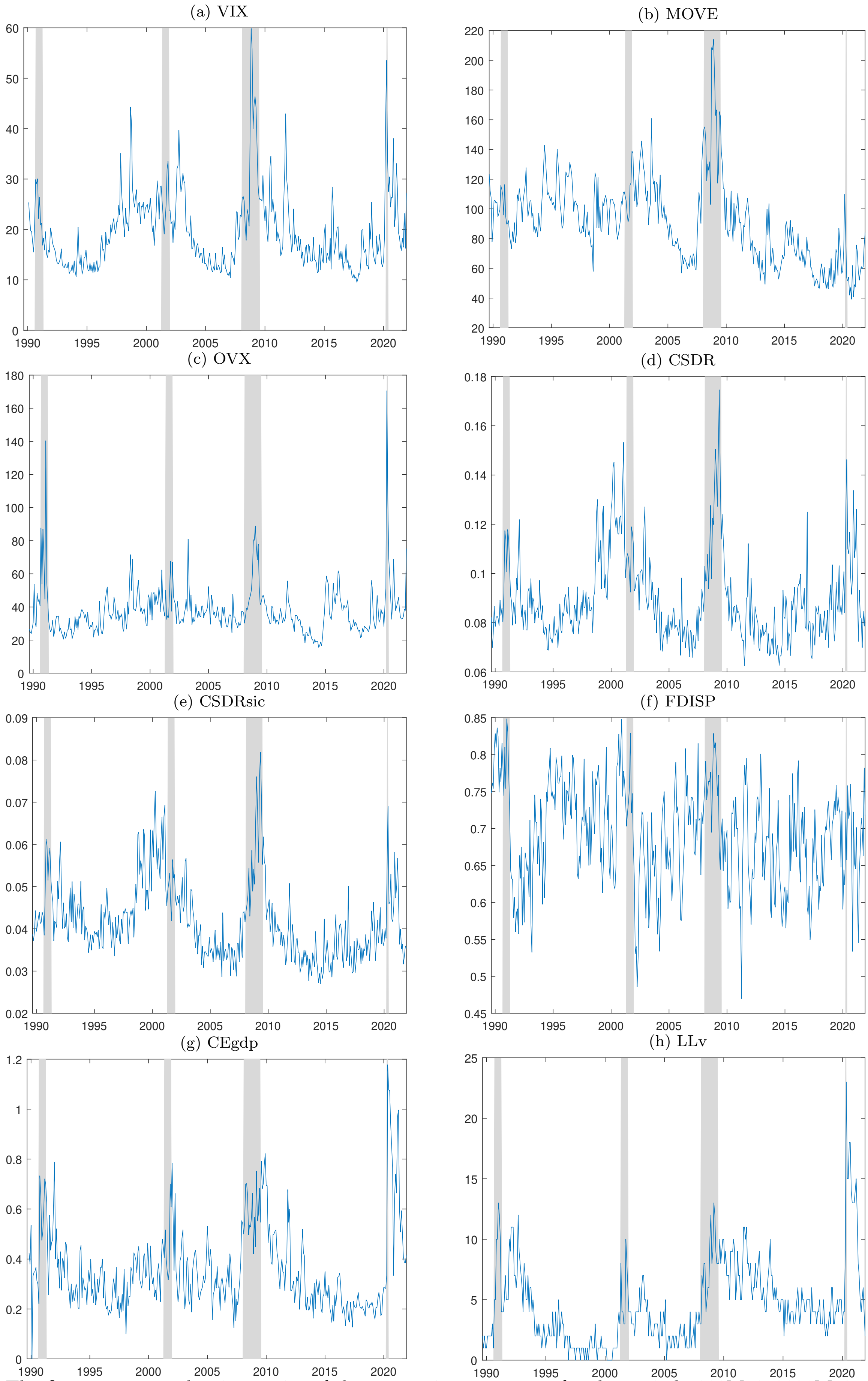
Appendix A Uncertainty measures

Table A.1: Uncertainty measures

#	Abbrev.	Description	Reference	Source	Type	Start	Merge start
1	VIX	End of month closing price of VIX	Bloom (2009)	CBOE	A	1990M1	
2	MOVE	ICE BofA US Bond Market Option Volatility Estimate Index		ICE	A	1988M4	
3	OVX	(1) Realized volatility of daily WTI returns; (2) End of month closing price of OVX	Kellogg (2014)	FRED (daily returns) and CBOE (OVX)	A	1986M2	2007M5
4	CSDR	Cross sectional standard deviation of stocks with 500+ month observations	Bloom (2009)	CRSP	B	1926M1	
5	CSDRsic	Cross sectional standard deviation of stocks with 500+ month observations; mean over dispersion in SIC3 code	Bloom (2009)	CRSP	B	1926M1	
6	FDISP	Ex ante forecast dispersion of general business conditions 6 months ahead; not seasonally adjusted	Bachmann et al. (2013)	Philadelphia Fed's Manufacturing Business Outlook Survey	B, D	1968M5	
7	CEgdp	Consensus economics forecaster interquartile range on GDP forecasts	Dovern et al. (2012)	Consensus Economics	B, D	1989M10	
8	LLv	Consumer confidence (personal vehicle)	Leduc and Liu (2016)	Thomson Reuters/University of Michigan Survey of Consumers	D	1978M2	
9	LLh	Consumer confidence (large households)	Fajgelbaum et al. (2017)	Thomson Reuters/University of Michigan Survey of Consumers	D	1978M1	
10	EPU+	Economic policy uncertainty; combination of newspaper counts, tax code provisions and forecaster disagreement in the survey of professional forecasters (SPF)	Baker et al. (2016)	Economic Policy Uncertainty website	C, D	1985M1	
11	EPU	(1) Historical economic policy uncertainty based on 6 to 10 major US newspapers; (2) Newspaper based economic policy uncertainty (original)	Baker et al. (2016)	Economic Policy Uncertainty website	C	1900M1	1985M1
12	MPU	Monetary Policy Uncertainty; category of EPU, counts of articles additionally containing monetary policy related keywords	Baker et al. (2016)	Economic Policy Uncertainty website	C	1985M1	
13	JLNm	Macroeconomic variables' forecast error variance based on large factor model, horizon=12	Jurado et al. (2015)	Sydney Ludvigson's website	E	1960M7	
14	JLNf	Financial variables' forecast error variance based on large factor model, horizon=12	Jurado et al. (2015)	Sydney Ludvigson's website	E	1960M7	
15	JLNr	Real activity variables' forecast error variance based on large factor model, horizon=12	Jurado et al. (2015)	Sydney Ludvigson's website	E	1960M7	

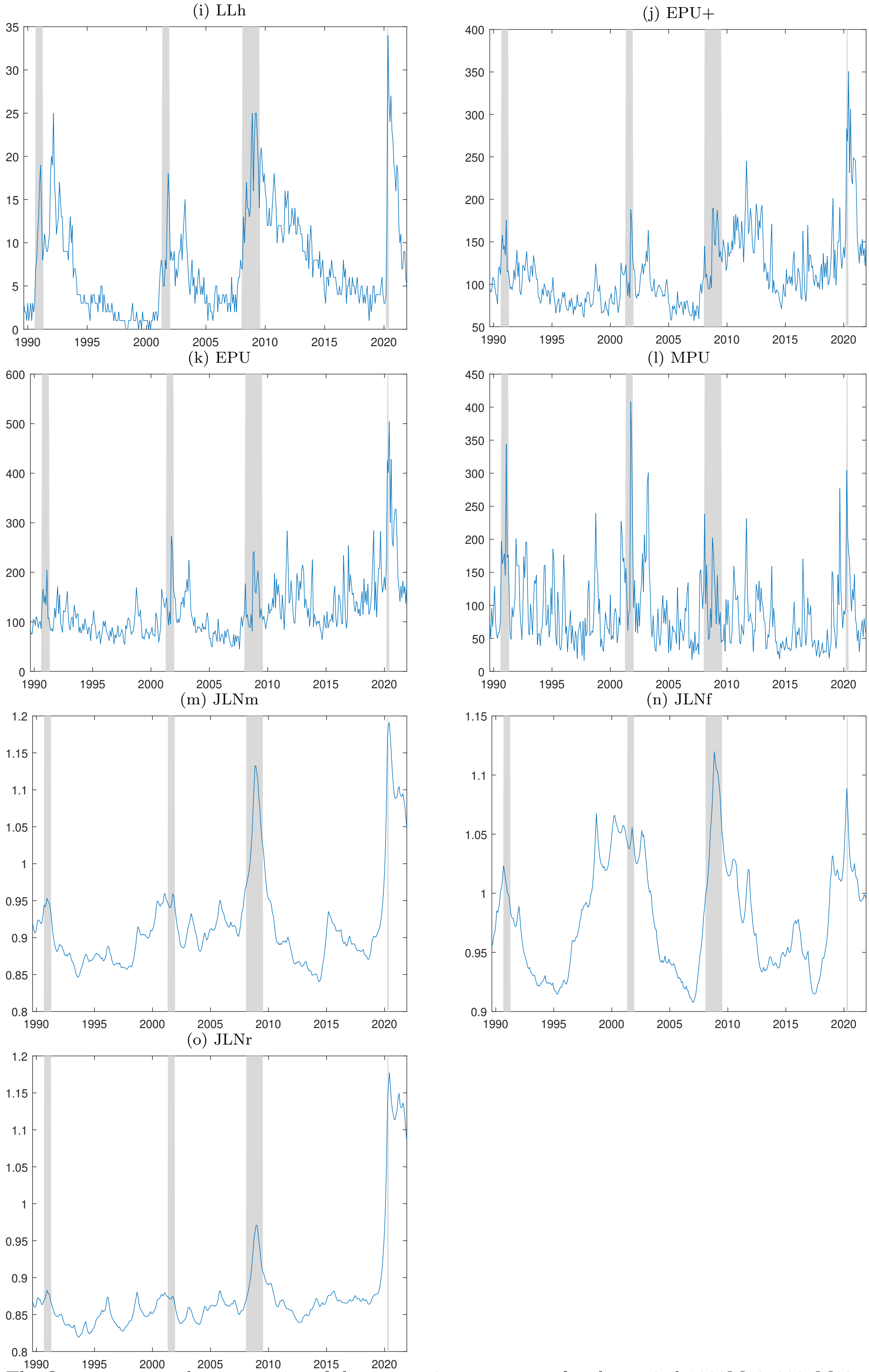
The table presents the uncertainty measures, their source, reference, uncertainty type (conditional volatility [A], cross-sectional dispersion [B], news [C], surveys [D], forecast errors [E]), start of the sample and when the data merge starts, where relevant. Websites: Economic Policy Uncertainty (`http://www.policyuncertainty.com/index.html`), Sydney Ludvigson's website (`https://www.sydneyludvigson.com/`).

Figure A.1: Uncertainty measure time series



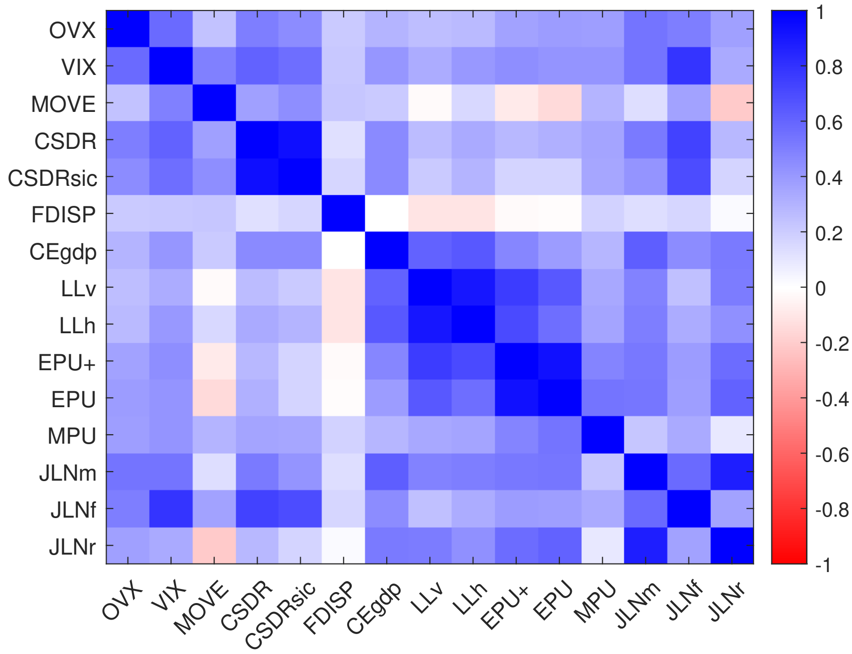
The figure presents the time series of the uncertainty measures for the period 1989M10–2021M12.

Figure A.1: Uncertainty measure time series (continued)



The figure presents the time series of the uncertainty measures for the period 1989M10–2021M12.

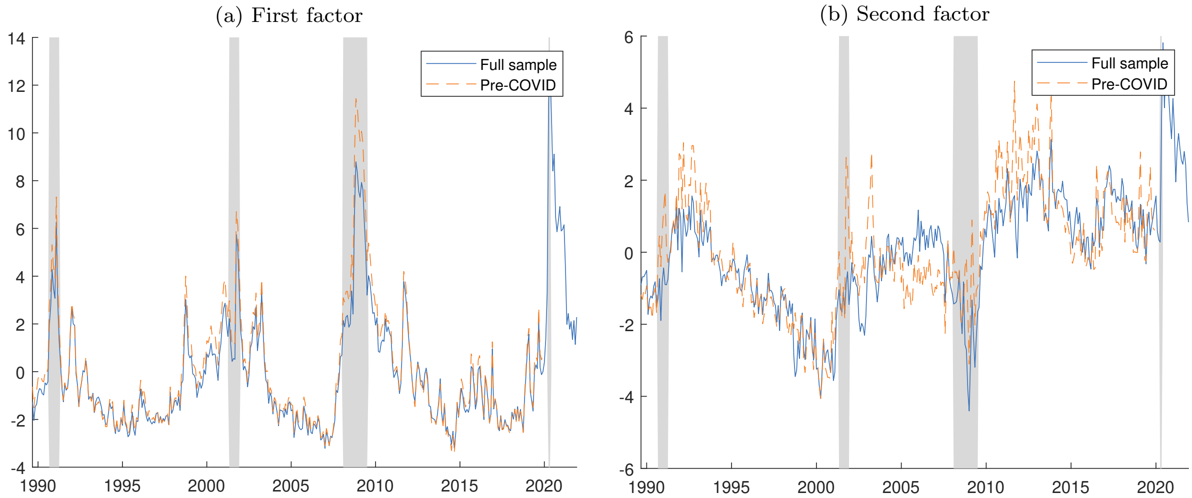
Figure A.2: Correlation matrix



The figure presents the correlation matrix of the uncertainty measures for the period 1989M10–2021M12.

Appendix B Uncertainty factors and COVID-19

Figure B.1: Impact of COVID-19 on the uncertainty factors



The figures present the time series of the first two factors estimated on the full sample, 1989M10–2021M12, (solid blue line) and the first two factor estimated on the pre-COVID sample, 1989M10–2019M12, (dashed orange line). The second factor based on the full sample is rotated (multiplied by -1) for interpretation purposes. The gray bars are recessions as determined by NBER's Business Cycle Dating Committee.

Appendix C Real-time data: coincident variables

Real-time data of the coincident economic index (CEI) is obtained from The Conference Board. The four component variables – industrial production (IP), nonfarm payroll employment (EMP), manufacturing and trade industries sales (MTS), and personal income excluding current transfer receipts (PIX) – are obtained from the data set of Chauvet and Piger (2008). It is an updated version of the Giusto and Piger (2017) data set, which updates the Chauvet and Piger (2008) data set to 2013.[11] The data set is updated using the Philadelphia Fed's Real-Time Data Set for Macroeconomists (Croushore and Stark, 2001) for industrial production and employment. The most recent vintages for sales and personal income are taken from St. Louis Fed's ALFRED. For manufacturing and trade industries sales, we use real manufacturing and trade industries sales (CMRMTSPL). Three vintages of MTS are missing (2013M10, 2014M01, and 2015M09), and we use vintages from the Conference Board to fill post-1996 observations.

For personal income excluding current transfer receipts, we follow Giusto and Piger (2017) by computing the real personal income excluding transfer receipts as the difference between personal income (PI) and personal current transfer receipts (PCTR), and dividing by the ratio of nominal (DSPI) to real disposable income (DSPIC96). Three vintages are missing of PIX due to a large (NIPA) revision at the end of 1995. Imputing the values as Chauvet and Piger (2008) is not possible because it requires observations before and after the missing sample, from the same vintage. In fact, Chauvet and Piger (2008) and Giusto and Piger (2017) skip the 1995M11–1996M1 vintages. Similarly, we delete the rows corresponding to the missing values before estimation. Since it involves only three vintages, it costs at most six observations. Additionally, there are five (additive) outliers in the level of PIX (1992M12, 1993M12, 2004M12, 2005M08, and 2012M12). We still include them in the estimation sample – it is probably difficult for the forecaster to identify outliers in real-time. However, due to the autoregressive model, the outliers also affect other forecasts. As an alternative, we impute the outliers by the

[11] Thanks to Jeremy Piger for uploading the raw data set on his website: `https://pages.uoregon.edu/jpiger/research/published-papers/`.

final vintage's unconditional median growth rate. In both cases, the periods at which the outliers are observed are excluded in the evaluation.

Benchmark revision dates are from the documentation of Philadelphia Fed's Real-Time Data Set for Macroeconomists (Croushore and Stark, 2001), from the Federal Reserve Board of Governors (`https://www.federalreserve.gov/releases/g17/`), from the Bureau of Labor Statistics (`https://www.bls.gov/web/empsit/cestn.htm#section7`), and from the Bureau of Economic Analysis (Page 1-10, note 22, of the November 2017 edition of the NIPA handbook, `https://www.bea.gov/resources/methodologies/nipa-handbook`). Further, we check for revisions in the data by looking at non-zero revisions of the sixth up to the twelfth release per vintage to identify remaining revisions. Though the reporting of revisions is quite accurate, we do identify some additional ones. But these are mostly in the in-sample period.

Appendix D Relative mean tick loss

Table D.1: **Relative mean tick loss by forecast horizon, $\alpha = 0.2$**

Horizon (months)	1	3	6	12	24
		Panel A: Coincident economic index			
MTL of HQ	1.518	1.105	0.937	0.856	0.800
NFCI	0.986	0.896**	0.812**	0.828	0.841
FRED-MD	1.055	1.019	0.989	0.949	0.789**
Factor models					
$r = 1$	1.017	0.933	0.887*	0.879*	0.909
$r = 2$	1.018	0.963	0.893*	0.925	0.856*
$r = 3$	1.018	0.947	0.870	0.929	0.793**
Uncertainty measures					
VIX	1.003	0.971*	0.913**	0.932*	0.943
MOVE	0.997	0.992	0.987	1.023	1.022
OVX	0.993	0.938*	0.965*	1.001	0.924
CSDR	1.011	1.002	0.970	1.002	0.957
CSDRsic	1.001	0.981*	0.966*	0.976	0.981
FDISP	0.994	0.992	0.958***	0.964	0.839***
CEgdp	1.024	0.982	0.969	1.017	0.995
LLv	1.009	0.982	1.032	1.040	0.974
LLh	0.998	0.963	1.004	1.058	0.987
EPU+	1.010	1.025	1.045	1.084	1.026
EPU	1.015	1.003	0.983	1.033	1.019
MPU	1.002	0.982	0.950***	0.980	0.998
JLNm	1.040	0.954	0.925	0.934	0.834*
JLNf	1.003	0.970*	0.907**	0.907*	0.885
JLNr	1.053	1.039	1.036	0.976	0.750**
		Panel B: Industrial production			
MTL of HQ	3.080	2.064	1.882	1.590	1.356
NFCI	0.954**	0.901*	0.802*	0.887	1.017
FRED-MD	0.978	1.015	0.916	0.953	0.776***
Factor models					
$r = 1$	0.957*	0.953	0.866	0.964	0.989
$r = 2$	0.968	0.939	0.834**	0.886	0.883
$r = 3$	0.925*	0.919	0.811*	0.920	0.876
Uncertainty measures					
VIX	0.962**	0.936*	0.920*	0.957	1.015
MOVE	0.992	1.024	0.995	1.019	0.983
OVX	0.973	0.961*	0.934*	0.987	1.012
CSDR	1.000	0.974	0.943*	0.962	1.045
CSDRsic	0.999	0.978*	0.952*	0.984	1.061
FDISP	0.990	0.971**	0.964**	0.994	0.960
CEgdp	1.007	1.009	0.995	1.043	1.007
LLv	1.012	1.010	1.023	0.982	0.941**
LLh	1.001	1.017	1.036	1.007	0.951
EPU+	0.990	1.012	1.035	1.031	1.018
EPU	0.981	0.989	0.994	1.007	1.050
MPU	0.970**	0.955**	0.969	0.994	1.020
JLNm	0.960	0.960	0.818	0.856	0.860
JLNf	0.977**	0.957	0.872**	0.879*	0.898
JLNr	1.013	1.089	0.982	0.996	0.798**

The table presents the relative mean tick loss (RMTL) with the historical quantile (HQ) as benchmark for various forecasting horizons for the out-of-sample period 1999M12+h–2021M12 and quantile level $\alpha = 0.2$. Further, the table presents the mean tick loss (MSPE) of the historical quantile. ***,**, and * denote significance of a one-sided Diebold-Mariano test at the 1%, 5% and 10%, respectively. See Table A.1 for an explanation of the abbreviations.

Table D.1: Relative mean tick loss by forecast horizon, $\alpha = 0.2$ (continued)

Horizon (months)	1	3	6	12	24
Panel C: Nonfarm payroll employment					
MTL of HQ	1.330	1.157	0.900	0.839	0.728
NFCI	0.965***	0.943**	0.886**	0.842**	0.869
FRED-MD	0.967*	0.968	0.929*	0.932**	0.858***
Factor models					
$r = 1$	0.999	0.915**	0.896**	0.879**	0.943
$r = 2$	0.982*	0.918**	0.880***	0.868***	0.785**
$r = 3$	0.980	0.890**	0.876**	0.861**	0.772**
Uncertainty measures					
VIX	0.975**	0.943**	0.926***	0.913***	0.987
MOVE	0.974***	0.981	0.943**	0.966	1.048
OVX	0.988	0.963**	0.962***	0.965**	0.973
CSDR	0.990*	0.980**	0.978*	0.963*	0.971
CSDRsic	0.980***	0.981**	0.973**	0.956**	1.010
FDISP	0.984***	0.996	0.988**	0.969*	0.894***
CEgdp	1.020	0.983	0.977	1.011	1.014
LLv	1.021	0.993	1.007	1.014	0.985
LLh	1.010	0.983	1.000	1.022	0.983
EPU+	1.015	0.977	1.004	1.025	1.021
EPU	1.012	0.975	0.996	1.007	1.029
MPU	0.981***	0.975***	0.968***	0.969*	1.007
JLNm	1.034	0.925*	0.879**	0.763**	0.715**
JLNf	0.978***	0.955**	0.914***	0.878***	0.845*
JLNr	1.079	0.917	0.992	0.856*	0.786**
Panel D: Manufacturing and trade sales					
MTL of HQ	3.419	1.884	1.420	1.224	1.065
NFCI	0.956**	0.877**	0.780*	0.913	0.894
FRED-MD	1.034	1.020	0.972	0.944	0.902
Factor models					
$r = 1$	0.974	0.926	0.916	0.974	1.009
$r = 2$	0.993	0.941	0.892	0.894	0.895
$r = 3$	1.001	0.937	0.829	0.892	0.744**
Uncertainty measures					
VIX	0.974**	0.962	0.952	0.949	1.039
MOVE	0.988	0.986	1.049	1.061	1.043
OVX	0.982	0.978	0.962	0.964	0.990
CSDR	1.006	0.988	0.982	0.994	1.010
CSDRsic	0.988*	0.976	0.984	0.992	1.046
FDISP	0.995	1.012	0.999	0.966	0.894***
CEgdp	0.993	0.991	0.990	1.024	1.009
LLv	1.022	1.006	1.046	1.075	0.946
LLh	1.012	0.997	1.053	1.075	0.974
EPU+	1.025	1.016	1.027	1.064	0.970
EPU	1.029	1.021	1.017	1.032	0.968
MPU	1.016	0.995	0.998	1.016	0.995
JLNm	0.958	0.895	0.850	0.918	0.950
JLNf	0.988*	0.940	0.948	0.944	1.041
JLNr	0.999	1.018	1.015	0.989	0.889

The table presents the relative mean tick loss (RMTL) with the historical quantile (HQ) as benchmark for various forecasting horizons for the out-of-sample period 1999M12+h–2021M12 and quantile level $\alpha = 0.2$. Further, the table presents the mean tick loss (MSPE) of the historical quantile. ***,**, and * denote significance of a one-sided Diebold-Mariano test at the 1%, 5% and 10%, respectively. See Table A.1 for an explanation of the abbreviations.

Table D.1: Relative mean tick loss by forecast horizon, $\alpha = 0.2$ (continued)

Horizon (months)	1	3	6	12	24
Panel E: Personal income excluding transfer receipts					
MTL of HQ	1.563	1.062	0.773	0.733	0.723
NFCI	0.955***	0.897**	0.856*	0.849	0.829
FRED-MD	1.025	0.953	0.943	1.004	0.952
Factor models					
$r = 1$	0.952*	0.956	0.968	0.961	0.912
$r = 2$	1.012	1.011	1.015	0.981	0.849
$r = 3$	1.020	1.014	0.998	0.835	0.780**
Uncertainty measures					
VIX	0.990	0.971	1.038	1.008	0.936
MOVE	0.963***	0.978	0.984	1.051	1.013
OVX	0.968**	0.928*	0.998	0.995	0.962
CSDR	0.975**	0.971	1.033	1.051	1.034
CSDRsic	0.972**	0.980	1.019	1.047	1.028
FDISP	1.004	0.994	0.990	1.018	0.881**
CEgdp	0.988	1.010	1.007	1.037	0.998
LLv	1.006	1.027	1.033	1.048	1.005
LLh	0.978	0.984	1.012	1.062	1.004
EPU+	1.028	1.075	1.074	1.040	1.097
EPU	1.033	1.022	1.066	1.012	1.048
MPU	0.994	0.988	0.997	0.994	1.019
JLNm	0.981	0.950	0.969	0.868	0.822
JLNf	0.996	0.972	0.979	0.972	0.896
JLNr	1.017	1.051	1.101	0.991	0.802*

The table presents the relative mean tick loss (RMTL) with the historical quantile (HQ) as benchmark for various forecasting horizons for the out-of-sample period 1999M12+h–2021M12 and quantile level $\alpha = 0.2$. Further, the table presents the mean tick loss (MSPE) of the historical quantile. ***,**, and * denote significance of a one-sided Diebold-Mariano test at the 1%, 5% and 10%, respectively. See Table A.1 for an explanation of the abbreviations.

Table D.2: Relative mean tick loss by evaluation period, $h = 3$ and $\alpha = 0.2$

Evaluation period	2000M1–2021M12	Recessions	Expansions	2000M1–2019M12
	Panel A: Coincident economic index			
MTL of HQ	1.105	3.596	0.807	0.566
NFCI	0.896**	0.714	0.994	0.782*
FRED-MD	1.019	0.712	1.184	0.985
Factor models				
$r = 1$	0.933	0.743	1.034	0.891
$r = 2$	0.963	0.784	1.058	0.950
$r = 3$	0.947	0.692	1.083	0.912
Uncertainty measures				
VIX	0.971*	0.953	0.982	0.960
MOVE	0.992	0.932	1.023	0.970
OVX	0.938*	0.843	0.989	0.909
CSDR	1.002	0.984	1.011	1.016
CSDRsic	0.981*	0.953	0.996	0.971
FDISP	0.992	0.972	1.003	0.987
CEgdp	0.982	0.867	1.044	0.972
LLv	0.982	0.819	1.068	0.977
LLh	0.963	0.758	1.072	0.927
EPU+	1.025	0.852	1.117	1.076
EPU	1.003	0.913	1.051	1.034
MPU	0.982	0.986	0.980	0.992
JLNm	0.954	0.700	1.090	0.925
JLNf	0.970*	0.906	1.005	0.950*
JLNr	1.039	0.677	1.231	0.915
	Panel B: Industrial production			
MTL of HQ	2.064	6.378	1.548	1.465
NFCI	0.901*	0.742	0.980	0.842*
FRED-MD	1.015	0.855	1.094	0.918
Factor models				
$r = 1$	0.953	0.748	1.054	0.895
$r = 2$	0.939	0.739	1.037	0.894
$r = 3$	0.919	0.672	1.041	0.841*
Uncertainty measures				
VIX	0.936*	0.850	0.978	0.902*
MOVE	1.024	1.018	1.027	1.024
OVX	0.961*	0.933	0.975	0.958
CSDR	0.974	0.892	1.014	0.959
CSDRsic	0.978*	0.942	0.995	0.967*
FDISP	0.971**	0.945	0.983	0.948**
CEgdp	1.009	0.935	1.045	0.994
LLv	1.010	0.960	1.034	0.996
LLh	1.017	0.966	1.043	1.011
EPU+	1.012	0.907	1.064	1.000
EPU	0.989	0.892	1.037	0.966
MPU	0.955**	0.963	0.951	0.937*
JLNm	0.960	0.641	1.118	0.862*
JLNf	0.957	0.796	1.036	0.939
JLNr	1.089	0.569	1.345	0.899

The table presents the relative mean tick loss (RMTL) with the historical quantile (HQ) as benchmark for various evaluation periods for a forecast horizon of 3 months and quantile level of $\alpha = 0.2$. Further, the table presents the mean tick loss (MSPE) of the historical quantile. ***,**, and * denote significance of a one-sided Diebold-Mariano test at the 1%, 5% and 10%, respectively. See Table A.1 for an explanation of the abbreviations.

Table D.2: Relative mean tick loss by evaluation period, $h = 3$ and $\alpha = 0.2$ (continued)

Evaluation period	2000M1–2021M12	Recessions	Expansions	2000M1–2019M12
	Panel C: Nonfarm payroll employment			
MTL of HQ	1.157	2.750	0.966	0.313
NFCI	0.943**	0.848	0.975	0.787**
FRED-MD	0.968	0.858	1.005	0.916
Factor models				
$r = 1$	0.915**	0.835	0.943	0.810**
$r = 2$	0.918**	0.837	0.946	0.804**
$r = 3$	0.890**	0.800	0.921	0.764**
Uncertainty measures				
VIX	0.943**	0.884	0.963	0.868**
MOVE	0.981	0.958	0.989	0.900***
OVX	0.963**	0.946	0.969	0.930**
CSDR	0.980**	0.947	0.991	0.939**
CSDRsic	0.981**	0.969	0.985	0.940**
FDISP	0.996	0.974	1.004	0.967*
CEgdp	0.983	0.971	0.988	0.960**
LLv	0.993	0.981	0.997	1.019
LLh	0.983	0.951	0.994	0.981
EPU+	0.977	0.904	1.001	1.010
EPU	0.975	0.907	0.998	1.024
MPU	0.975***	0.961	0.980	0.940**
JLNm	0.925*	0.790	0.971	0.841*
JLNf	0.955**	0.863	0.986	0.877**
JLNr	0.917	0.778	0.965	0.940
	Panel D: Manufacturing and trade sales			
MTL of HQ	1.884	6.873	1.287	1.492
NFCI	0.877**	0.655	1.019	0.811**
FRED-MD	1.020	0.731	1.204	0.951
Factor models				
$r = 1$	0.926	0.654	1.100	0.867*
$r = 2$	0.941	0.659	1.122	0.885
$r = 3$	0.937	0.597	1.154	0.852*
Uncertainty measures				
VIX	0.962	0.835	1.044	0.941
MOVE	0.986	0.876	1.057	0.964
OVX	0.978	0.827	1.074	0.952
CSDR	0.988	0.855	1.074	0.968
CSDRsic	0.976	0.883	1.035	0.948
FDISP	1.012	0.970	1.038	1.008
CEgdp	0.991	0.785	1.123	0.966
LLv	1.006	0.853	1.104	0.991
LLh	0.997	0.788	1.131	0.965
EPU+	1.016	0.716	1.209	1.003
EPU	1.021	0.779	1.176	1.009
MPU	0.995	0.948	1.025	0.977
JLNm	0.895	0.501	1.147	0.793**
JLNf	0.940	0.769	1.050	0.928
JLNr	1.018	0.502	1.349	0.846*

The table presents the relative mean tick loss (RMTL) with the historical quantile (HQ) as benchmark for various evaluation periods for a forecast horizon of 3 months and quantile level of $\alpha = 0.2$. Further, the table presents the mean tick loss (MSPE) of the historical quantile. ***,**, and * denote significance of a one-sided Diebold-Mariano test at the 1%, 5% and 10%, respectively. See Table A.1 for an explanation of the abbreviations.

Table D.2: Relative mean tick loss by evaluation period, $h = 3$ and $\alpha = 0.2$ (continued)

Evaluation period	2000M1–2021M12	Recessions	Expansions	2000M1–2019M12
Panel E: Personal income excluding transfer receipts				
MTL of HQ	1.062	2.659	0.863	0.765
NFCI	0.897	0.752	0.953	0.841
FRED-MD	0.953	0.745	1.033	0.948
Factor models				
$r = 1$	0.956	0.790	1.019	0.890
$r = 2$	1.011	0.839	1.077	0.941
$r = 3$	1.014	0.746	1.117	0.935
Uncertainty measures				
VIX	0.971	0.934	0.985	0.967
MOVE	0.978	1.003	0.969	0.971
OVX	0.928	0.886	0.944	0.930
CSDR	0.971	0.963	0.974	0.938
CSDRsic	0.980	0.979	0.981	0.952
FDISP	0.994	0.949	1.011	0.982
CEgdp	1.010	0.887	1.057	0.963
LLv	1.027	0.844	1.097	0.973
LLh	0.984	0.839	1.040	0.929
EPU+	1.075	0.911	1.137	1.080
EPU	1.022	0.938	1.053	1.010
MPU	0.988	0.993	0.986	0.985
JLNm	0.950	0.708	1.043	0.874
JLNf	0.972	0.915	0.994	0.960
JLNr	1.051	0.648	1.206	0.900

The table presents the relative mean tick loss (RMTL) with the historical quantile (HQ) as benchmark for various evaluation periods for a forecast horizon of 3 months and quantile level of $\alpha = 0.2$. Further, the table presents the mean tick loss (MSPE) of the historical quantile. ***,**, and * denote significance of a one-sided Diebold-Mariano test at the 1%, 5% and 10%, respectively. See Table A.1 for an explanation of the abbreviations.

Table D.3: Relative mean tick loss by quantile, $h = 3$

α	0.1	0.2	0.5	0.8
	Panel A: Coincident economic index			
MTL of HQ	0.946	1.105	1.210	0.995
NFCI	0.912*	0.896**	0.942**	0.986
FRED-MD	1.096	1.019	0.975	1.079
Factor models				
$r = 1$	0.902	0.933	1.010	1.022
$r = 2$	0.925	0.963	1.013	1.016
$r = 3$	0.977	0.947	0.962	1.009
Uncertainty measures				
VIX	0.943	0.971*	0.998	1.006
MOVE	0.993	0.992	1.001	1.001
OVX	0.873**	0.938*	1.002	1.006
CSDR	0.980	1.002	1.006	1.002
CSDRsic	0.957	0.981*	1.007	0.998
FDISP	0.970	0.992	1.002	1.003
CEgdp	0.952	0.982	1.023	0.990**
LLv	0.970	0.982	1.013	1.003
LLh	0.937	0.963	1.008	1.004
EPU+	1.009	1.025	1.002	1.009
EPU	1.013	1.003	0.992	1.011
MPU	0.970	0.982	0.993	1.001
JLNm	0.935	0.954	1.021	1.034
JLNf	0.945	0.970*	0.999	0.999
JLNr	0.984	1.039	1.031	1.030
	Panel B: Industrial production			
MTL of HQ	1.595	2.064	2.477	1.848
NFCI	0.870*	0.901*	0.987	1.007
FRED-MD	1.041	1.015	1.006	1.107
Factor models				
$r = 1$	0.900	0.953	1.013	1.008
$r = 2$	0.883*	0.939	1.011	1.006
$r = 3$	0.875	0.919	1.004	1.072
Uncertainty measures				
VIX	0.890**	0.936*	0.996	1.017
MOVE	1.005	1.024	1.007	1.001
OVX	0.924**	0.961*	0.990	1.002
CSDR	0.960*	0.974	1.004	1.006
CSDRsic	0.971	0.978*	1.009	1.000
FDISP	0.966*	0.971**	0.993	0.999
CEgdp	1.051	1.009	1.000	0.977**
LLv	1.056	1.010	1.002	0.983***
LLh	1.045	1.017	1.003	0.982***
EPU+	1.048	1.012	1.011	1.012
EPU	1.033	0.989	1.001	1.030
MPU	0.980	0.955**	0.976***	0.989**
JLNm	0.868	0.960	1.052	1.011
JLNf	0.921**	0.957	0.987	1.008
JLNr	0.975	1.089	1.100	1.053

The table presents the relative mean tick loss (RMTL) with the historical quantile (HQ) as benchmark for various quantile levels for the out-of-sample period 1999M12+h–2021M12, with a forecast horizon of 3 months. Further, the table presents the mean tick loss (MSPE) of the historical quantile. ***,**, and * denote significance of a one-sided Diebold-Mariano test at the 1%, 5% and 10%, respectively. See Table A.1 for an explanation of the abbreviations.

Table D.3: Relative mean tick loss by quantile level, $h = 3$ (continued)

α	0.1	0.2	0.5	0.8
Panel C: Nonfarm payroll employment				
MTL of HQ	1.102	1.157	1.072	0.819
NFCI	0.937**	0.943**	0.981	1.011
FRED-MD	0.982	0.968	0.986*	0.981
Factor models				
$r = 1$	0.882***	0.915**	1.041	1.034
$r = 2$	0.891***	0.918**	1.023	1.047
$r = 3$	0.849**	0.890**	1.008	0.988
Uncertainty measures				
VIX	0.915***	0.943**	1.001	1.012
MOVE	0.973	0.981	0.991	0.992*
OVX	0.938**	0.963**	0.995	1.005
CSDR	0.962***	0.980**	1.005	0.999
CSDRsic	0.964***	0.981**	1.007	0.992**
FDISP	0.991	0.996	1.003	0.998
CEgdp	0.972	0.983	1.026	0.978***
LLv	0.966	0.993	1.006	1.019
LLh	0.952*	0.983	1.005	1.016
EPU+	0.922*	0.977	1.013	1.022
EPU	0.914*	0.975	1.008	1.013
MPU	0.966***	0.975***	0.997	0.996*
JLNm	0.857**	0.925*	1.040	1.032
JLNf	0.923***	0.955**	1.004	1.011
JLNr	0.736*	0.917	0.979	0.960
Panel D: Manufacturing and trade sales				
MTL of HQ	1.404	1.884	2.239	1.646
NFCI	0.841**	0.877**	0.958	0.985
FRED-MD	0.986	1.020	1.083	1.229
Factor models				
$r = 1$	0.904	0.926	1.033	1.026
$r = 2$	0.903	0.941	1.025	1.029
$r = 3$	0.890	0.937	1.022	1.071
Uncertainty measures				
VIX	0.891**	0.962	1.020	1.006
MOVE	0.930	0.986	1.014	1.009
OVX	0.934	0.978	1.011	0.996
CSDR	0.932**	0.988	1.024	1.011
CSDRsic	0.945*	0.976	1.019	1.007
FDISP	0.984	1.012	1.004	1.003
CEgdp	0.954	0.991	1.023	1.011
LLv	0.995	1.006	1.019	1.004
LLh	0.964	0.997	1.019	1.006
EPU+	1.042	1.016	1.036	1.021
EPU	1.036	1.021	1.029	1.023
MPU	1.023	0.995	1.008	0.982
JLNm	0.843*	0.895	1.017	1.066
JLNf	0.868***	0.940	0.999	1.005
JLNr	0.947	1.018	1.083	1.115

The table presents the relative mean tick loss (RMTL) with the historical quantile (HQ) as benchmark for various quantile levels for the out-of-sample period 1999M12+h–2021M12, with a forecast horizon of 3 months. Further, the table presents the mean tick loss (MSPE) of the historical quantile. ***,**, and * denote significance of a one-sided Diebold-Mariano test at the 1%, 5% and 10%, respectively. See Table A.1 for an explanation of the abbreviations.

Table D.3: Relative mean tick loss by quantile level, $h = 3$ (continued)

α	0.1	0.2	0.5	0.8
Panel E: Personal income excluding transfer receipts				
MTL of HQ	0.830	1.062	1.264	0.907
NFCI	0.884**	0.897**	0.925**	0.971*
FRED-MD	0.931	0.953	0.996	1.177
Factor models				
$r = 1$	0.924	0.956	0.983	1.038
$r = 2$	0.959	1.011	1.043	1.096
$r = 3$	0.907	1.014	1.043	1.107
Uncertainty measures				
VIX	0.966	0.971	0.987	0.994
MOVE	0.978	0.978	0.994	0.983*
OVX	0.884**	0.928*	0.989*	0.998
CSDR	0.955*	0.971	1.006	1.001
CSDRsic	0.944*	0.980	1.007	0.999
FDISP	0.973*	0.994	1.006	1.012
CEgdp	0.994	1.010	1.018	1.045
LLv	0.986	1.027	1.012	1.064
LLh	0.953	0.984	0.982	1.037
EPU+	1.065	1.075	1.018	1.046
EPU	1.086	1.022	1.017	1.053
MPU	0.991	0.988	1.007	1.002
JLNm	0.881	0.950	1.033	1.077
JLNf	0.966	0.972	0.984**	0.986
JLNr	0.915	1.051	1.112	1.133

The table presents the relative mean tick loss (RMTL) with the historical quantile (HQ) as benchmark for various quantile levels for the out-of-sample period 1999M12+h–2021M12, with a forecast horizon of 3 months. Further, the table presents the mean tick loss (MSPE) of the historical quantile. ***,**, and * denote significance of a one-sided Diebold-Mariano test at the 1%, 5% and 10%, respectively. See Table A.1 for an explanation of the abbreviations.

Appendix E Hit rates

Table E.1: Hit rates by horizon

Horizon (months)	1	3	6	12	24
Panel A: Coincident economic index					
HQ	0.182	0.195	0.197	0.237	0.336 ‡
NFCI	0.186	0.179	0.189	0.304 ‡	0.390 † ‡
Factor models					
$r = 1$	0.182	0.168	0.174	0.229	0.349
$r = 2$	0.182	0.187	0.216	0.320 † ‡	0.344 † ‡
$r = 3$	0.163	0.134 † ‡	0.178	0.273	0.311 ‡
Uncertainty measures					
VIX	0.201	0.244	0.209	0.257	0.332 ‡
MOVE	0.227	0.294 † ‡	0.309 †	0.332 † ‡	0.357 ‡
OVX	0.193	0.210	0.201	0.245	0.328
CSDR	0.193	0.218	0.243	0.277	0.365 †
CSDRsic	0.208	0.206	0.247	0.273	0.415 † ‡
FDISP	0.208	0.237	0.228	0.273	0.357
CEgdp	0.189	0.210	0.255	0.289	0.340
LLv	0.152 †	0.160	0.182	0.237	0.332
LLh	0.163	0.172	0.197	0.237 ‡	0.344
EPU+	0.159	0.164 ‡	0.178	0.253 ‡	0.320 ‡
EPU	0.167	0.187	0.174	0.241	0.315 ‡
MPU	0.189	0.199	0.232	0.265	0.328 ‡
JLNm	0.171	0.160	0.154	0.245 ‡	0.307 ‡
JLNf	0.197	0.210	0.197	0.273	0.299 ‡
JLNr	0.159	0.137 †	0.143	0.154	0.178
Panel B: Industrial production					
HQ	0.246	0.233	0.247 ‡	0.324 ‡	0.361 ‡
NFCI	0.235	0.237	0.247	0.391 † ‡	0.432 † ‡
Factor models					
$r = 1$	0.235	0.191	0.247	0.324	0.357
$r = 2$	0.246	0.233	0.286	0.403 † ‡	0.456 † ‡
$r = 3$	0.212	0.176	0.232	0.395 † ‡	0.481 † ‡
Uncertainty measures					
VIX	0.239	0.263	0.278	0.344 † ‡	0.411 † ‡
MOVE	0.277 † ‡	0.282 † ‡	0.324 † ‡	0.387 † ‡	0.332
OVX	0.258 †	0.229	0.251	0.304 ‡	0.373 † ‡
CSDR	0.246	0.233	0.301 †	0.360 † ‡	0.407 † ‡
CSDRsic	0.258 †	0.260	0.282	0.372 † ‡	0.465 † ‡
FDISP	0.269 † ‡	0.233	0.259	0.312	0.436 † ‡
CEgdp	0.254 † ‡	0.225	0.270	0.336 ‡	0.365 † ‡
LLv	0.231	0.225	0.255 ‡	0.320	0.365
LLh	0.239	0.233	0.251 ‡	0.304 ‡	0.365 †
EPU+	0.227	0.214 ‡	0.243 ‡	0.336 † ‡	0.340 ‡
EPU	0.197	0.187	0.247	0.308 ‡	0.353 ‡
MPU	0.258 †	0.256	0.247	0.348 †	0.340 ‡
JLNm	0.208	0.187	0.178	0.245	0.340
JLNf	0.239	0.237	0.270	0.320	0.378 †
JLNr	0.171	0.118 † ‡	0.143	0.166	0.266

The table presents hit rates for various forecasting horizons, for the out-of-sample period 1999M12+h–2021M12 and quantile $\alpha = 0.2$. The † denotes rejection of the null hypothesis of correct unconditional coverage, and ‡ denotes rejection of the null hypothesis of correct coverage conditional on an intercept and the quantile estimates q_t, all at a 5% significance level, based on the DQ test with HAC standard errors. See Table A.1 for an explanation of the abbreviations.

Table E.1: Hit rates by horizon (continued)

Horizon (months)	1	3	6	12	24
		Panel C: Employment			
HQ	0.136 † ‡	0.179	0.201	0.281	0.344 ‡
NFCI	0.114 † ‡	0.164	0.193	0.293	0.423 † ‡
Factor models					
$r = 1$	0.114 † ‡	0.172	0.185	0.269	0.365 †
$r = 2$	0.117 † ‡	0.179	0.216	0.308	0.378 † ‡
$r = 3$	0.125 † ‡	0.164 ‡	0.197	0.300	0.398 † ‡
Uncertainty measures					
VIX	0.117 † ‡	0.168 ‡	0.216	0.245	0.390 † ‡
MOVE	0.152 †	0.229	0.255	0.364 † ‡	0.361 ‡
OVX	0.136 † ‡	0.183 ‡	0.201	0.261	0.357 ‡
CSDR	0.133 † ‡	0.187	0.232	0.293	0.390 †
CSDRsic	0.155	0.199	0.239	0.285	0.440 † ‡
FDISP	0.144 †	0.191	0.216	0.296	0.378 †
CEgdp	0.133 † ‡	0.187	0.201	0.285	0.344 ‡
LLv	0.117 † ‡	0.168	0.209	0.285	0.344 ‡
LLh	0.114 † ‡	0.157	0.201	0.277 ‡	0.353
EPU+	0.102 † ‡	0.149	0.220	0.277 ‡	0.336 ‡
EPU	0.110 † ‡	0.157	0.205	0.285 ‡	0.340 ‡
MPU	0.136 † ‡	0.176	0.220	0.285	0.349 ‡
JLNm	0.106 † ‡	0.130 † ‡	0.135	0.194	0.278
JLNf	0.117 † ‡	0.183	0.232	0.253	0.365 †
JLNr	0.106 † ‡	0.122 † ‡	0.143	0.194	0.274
		Panel D: Manufacturing and trade sales			
HQ	0.193	0.263	0.270 ‡	0.296	0.261 ‡
NFCI	0.178	0.279 † ‡	0.321 † ‡	0.403 † ‡	0.419 † ‡
Factor models					
$r = 1$	0.163	0.241	0.259 ‡	0.300	0.282
$r = 2$	0.174	0.252 ‡	0.317 † ‡	0.379 † ‡	0.419 † ‡
$r = 3$	0.182	0.214	0.290 † ‡	0.340 † ‡	0.349 † ‡
Uncertainty measures					
VIX	0.189	0.271	0.293	0.289	0.320
MOVE	0.239	0.321 † ‡	0.394 † ‡	0.328 † ‡	0.295
OVX	0.178	0.267 ‡	0.263	0.273	0.282 ‡
CSDR	0.182 ‡	0.290 † ‡	0.336 † ‡	0.340 † ‡	0.344
CSDRsic	0.201	0.298 † ‡	0.344 † ‡	0.376 † ‡	0.336
FDISP	0.201	0.260	0.286	0.289	0.311
CEgdp	0.208	0.267	0.290 ‡	0.316 ‡	0.291 ‡
LLv	0.171	0.244	0.266 ‡	0.308	0.278
LLh	0.182	0.241 ‡	0.259 ‡	0.281	0.295
EPU+	0.171	0.195 ‡	0.259 ‡	0.332 † ‡	0.278
EPU	0.186	0.199 ‡	0.259	0.308	0.307 ‡
MPU	0.212	0.271 †	0.301 † ‡	0.296	0.266 ‡
JLNm	0.140 † ‡	0.191	0.224	0.269 ‡	0.324 ‡
JLNf	0.212	0.260	0.278	0.281	0.291 ‡
JLNr	0.140 †	0.141 †	0.174	0.166	0.170

The table presents hit rates for various forecasting horizons, for the out-of-sample period 1999M12+h–2021M12 and quantile $\alpha = 0.2$. The † denotes rejection of the null hypothesis of correct unconditional coverage, and ‡ denotes rejection of the null hypothesis of correct coverage conditional on an intercept and the quantile estimates q_t, all at a 5% significance level, based on the DQ test with HAC standard errors. See Table A.1 for an explanation of the abbreviations.

Table E.1: Hit rates by horizon (continued)

Horizon (months)	1	3	6	12	24
Panel E: Personal income excluding transfer receipts					
HQ	0.155	0.188	0.213	0.279	0.298
NFCI	0.167	0.207	0.225	0.304	0.400 † ‡
Factor models					
$r = 1$	0.155	0.223	0.245	0.300	0.328
$r = 2$	0.143 †	0.203	0.269	0.344 † ‡	0.340 † ‡
$r = 3$	0.120 † ‡	0.164 ‡	0.210	0.259 ‡	0.328 ‡
Uncertainty measures					
VIX	0.167	0.227	0.253 ‡	0.316	0.302
MOVE	0.205	0.254	0.312 †	0.389 † ‡	0.345
OVX	0.155 ‡	0.207	0.237	0.304	0.306
CSDR	0.178	0.215	0.249	0.328 †	0.323
CSDRsic	0.178	0.242	0.249	0.316	0.353 †
FDISP	0.163	0.211	0.237	0.304	0.353 † ‡
CEgdp	0.186	0.219	0.257	0.332 † ‡	0.379 † ‡
LLv	0.159	0.180	0.221	0.300 ‡	0.289 ‡
LLh	0.167	0.180	0.245	0.287 ‡	0.319
EPU+	0.136 † ‡	0.172 ‡	0.217 ‡	0.275 ‡	0.277 ‡
EPU	0.143 † ‡	0.164	0.225	0.255	0.281 ‡
MPU	0.171	0.223	0.245	0.300	0.306 ‡
JLNm	0.140 † ‡	0.184	0.202	0.259 ‡	0.298 ‡
JLNf	0.167	0.219	0.221	0.304	0.302
JLNr	0.116 † ‡	0.141 †	0.154	0.198	0.204 ‡

The table presents hit rates for various forecasting horizons, for the out-of-sample period 1999M12+h–2021M12 and quantile $\alpha = 0.2$. The † denotes rejection of the null hypothesis of correct unconditional coverage, and ‡ denotes rejection of the null hypothesis of correct coverage conditional on an intercept and the quantile estimates q_t, all at a 5% significance level, based on the DQ test with HAC standard errors. See Table A.1 for an explanation of the abbreviations.