TI 2022-069/III
Tinbergen Institute Discussion Paper

Does economic uncertainty predict real activity in real-time?

Revision: September 2023

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

4 September 2023

Abstract

We assess the predictive ability of 15 economic uncertainty measures in a real-time out-of-sample forecasting exercise for 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 accurate forecasts are obtained using the National Financial Conditions Index (NFCI), the uncertainty factor models are better 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`. Faculty of Economics and Business, University of Amsterdam, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands.

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. Bloom (2009) sparked a new line of research, on empirically measuring economic uncertainty and assessing its relationship with real macroeconomic variables such as output and employment, 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 dispersion in forecast errors (Jurado et al., 2015).

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 quality. 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 almost exclusively based on in-sample analysis. It is important for the validity of these findings to test whether the relationship continues to hold out-of-sample. 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. Policy decisions often rely on forecasts of economic activity. More accurate forecasts could therefore lead to more informed decision making.

In this paper we address both open issues identified above. First, we collect an extensive 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.

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 vintages of the uncertainty measures should be available, in case they are subject to revisions. For the uncertainty measures of Jurado et al. (2015) we reconstruct real-time versions thereof, aiming to make our analysis as realistic as possible. 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 fairly strong common component among the uncertainty measures. 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 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 loads most heavily on news based and consumer survey based uncertainty measures. We therefore interpret this factor as media/consumer uncertainty. This second factor remains elevated after recessions, reflecting that media and consumers need more time to become confident about the recovery than reflected by economic 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 (CEI), 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 monthly data vintages to take into account that publications of macroeconomic variables are revised multiple times after their initial 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 the uncertainty measures mostly have predictive ability for the lower quantiles for CEI (and its components') growth rates. This mirrors the asymmetric relationship between GDP 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 strong predictor. When forecasting employment though, a factor model with uncertainty factors performs better at forecasting horizons shorter than 12 months. Interestingly, Bloom (2009) finds that employment responds negatively to uncertainty shocks, and uses this to build a labor-capital model. Moreover, we find that uncertainty factor models hold predictive content complementary to NFCI for other target variables as well. From the individual uncertainty measures financial volatility perform best. The performance of individual media and news based measures is disappointing. The Jurado et al. (2015) measures – from the volatility of forecast errors on a large set of macroeconomic and financial variables – are one of the best performing measures if the final vintage is used. Their predictive ability is substantially worse when using a real-time version that we construct for this exercise. 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 on the relationship between economic uncertainty and real macroeconomic variables by conducting a real-time quantile forecasting exercise. Second, we show 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.

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 GDP growth 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. We distinguish our analysis by performing a real-time forecasting exercise for different, monthly measures of real economic activity, and using a more extensive set of uncertainty measures.

While research on the predictive ability of uncertainty is limited, there is an extensive literature on forecasting economic output using measures of financial conditions and risk. We relate to this literature and draw inspiration from a few specific papers. First, Adrian et al. (2019) allow for asymmetry across the density based on quantile forecasts and find that especially the left tail of GDP growth is affected by financial conditions. Second, most similar to our paper in spirit, Giglio et al. (2016) conduct a quantile forecasting exercise for a set of systemic risk measures. They find that a single common factor improves forecast accuracy, and that predictive power for the mean is limited. The main difference with our paper – other than using uncertainty measures instead of systemic risk measures – is 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.

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. Full sample quantile regressions are presented in Section 5, followed by the forecasting results 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 could 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. 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 are only available at a lower frequency, or because they need to be computed ex post. 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 to estimate 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 uncertainty and Bloomberg announcements deviate more from expectations. The most prominent measures in this category are the indexes from Baker et al. (2016), based on newspaper article counts, from which we select general economic policy uncertainty (EPU and EPU+) and monetary policy uncertainty (MPU).

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

Fourth, outcomes of polls or surveys taken among consumers, professional forecasters or firms gauging their expectations for the coming period. This 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 five 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.

Only the EPU and JLN measures are subject to revisions. For the EPU measures, vintages are available from 2013M6 (EPU+) or 2019M10 (EPU and MPU). Observations before that period are from the first vintage. More details are provided in Appendix B.1. The JLN measures are not published in real-time. These measures are very popular and quite different from the other measures though. Therefore, as an exception, we construct our own monthly vintages using the methodology of Jurado et al. (2015), see Appendix B.2. The vintages are available monthly from 1999M8 onwards.

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 the 15 selected uncertainty measures. They are largely similar in that all measures peak around the time of recessions. Nevertheless, the average correlation between the uncertainty measures is quite modest at 40.8%. 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 63.0%. 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 76.8% and 73.8%, respectively. Across categories, measures based on stock market data are similar. For example, the correlation between VIX and JLNf is 83.3%.

The only outlier is FDISP. Other uncertainty measures are correlated no more than 23# with FDISP, and some are even negatively correlated: −9.9% (with LLv) and −10.2% (with LLh). Figure A.1f confirms idiosyncratic pattern for FDISP, which could be due its regional focus or because business surveys capture a unique part of uncertainty.

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 63.9% 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 almost 50% 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 uncertainty among the public remains elevated. 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 the NBER's Business Cycle Dating Committee usually 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 mostly FDISP and MPU, but there is no clear link between them. The fourth factor loads heavily on MOVE, and the fifth explains most of FDISP.

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 MOVE. Afterwards there are more pronounced

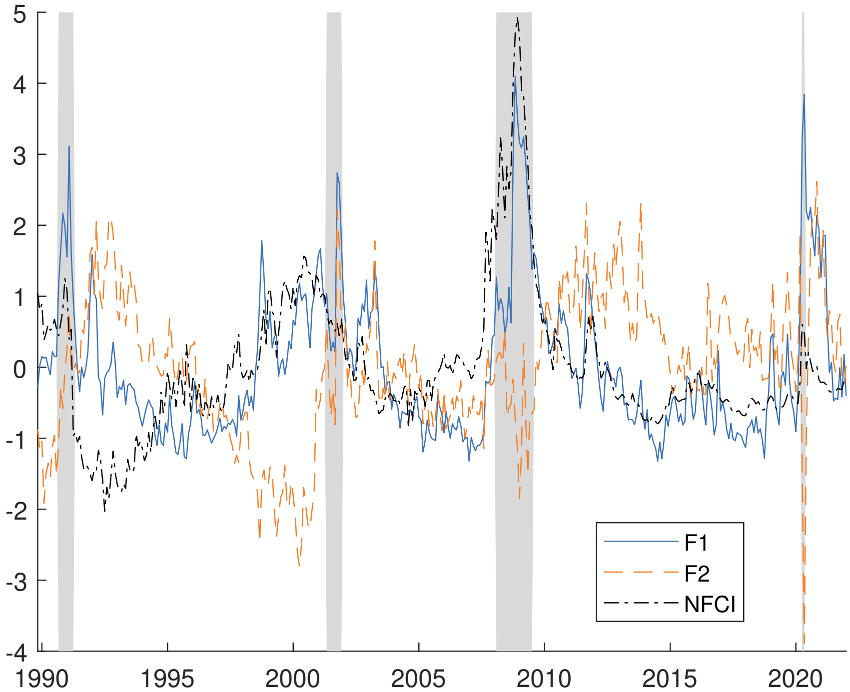
[2]The second factor is multiplied by -1. This does not matter for how much of the variance is explained or for forecasting, but makes it more intuitive to explain our interpretation of the factor.

Table 1: Factor loadings and marginal R^2

r	1	2	3	4	5
VIX	0.292	−0.206	0.104	0.018	−0.044
MOVE	0.120	−0.384	−0.038	0.533	0.189
OVX	0.245	−0.155	0.236	−0.248	−0.060
CSDR	0.277	−0.292	−0.193	−0.008	−0.282
CSDRsic	0.249	−0.344	−0.228	0.105	−0.224
FDISP	0.052	−0.254	0.540	−0.018	0.653
CEgdp	0.265	0.096	−0.345	0.130	0.318
LLv	0.252	0.376	−0.128	0.223	0.161
LLh	0.267	0.296	−0.200	0.314	0.194
EPU+	0.274	0.360	0.216	0.014	−0.107
EPU	0.265	0.327	0.311	−0.093	−0.230
MPU	0.206	0.006	0.467	0.447	−0.300
JLNm	0.315	0.036	−0.083	−0.364	0.227
JLNf	0.308	−0.204	−0.021	−0.177	−0.118
JLNr	0.334	−0.030	−0.087	−0.324	0.134
R^2	0.478	0.161	0.078	0.070	0.055

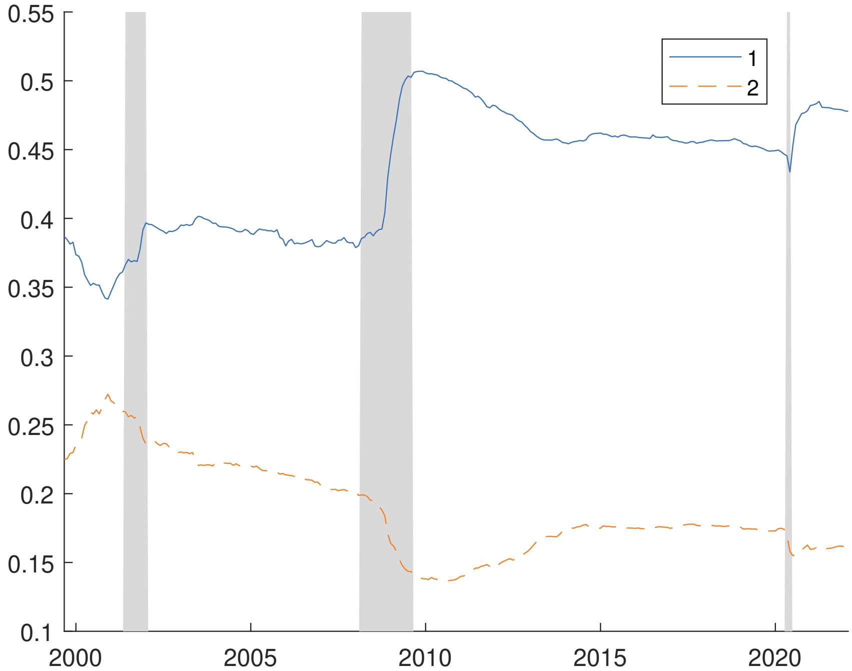
The table presents the factor loadings (top panel) and the marginal R^2, the fraction of total variance explained by the r-th principal component, for the first five principal components for the sample 1989M10–2021M12, based on the final vintage of the EPU and JLN measures (2022M1). The second factor is rotated (loadings multiplied by -1) for interpretation purposes. See Table A.1 for an explanation of the abbreviations.

Figure 1: Uncertainty factors and NFCI



The figure presents the time series of first releases 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 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.

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 C.1. The first factor is nearly identical (99% correlation), while the correlation is also high (89.0%) for the second factor between the different samples. For more details see Section C.

4 Methodology

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 D.

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

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

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

4.2 Quantile forecasts

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]

[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/`.

[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.

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 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 $Q_\alpha(y)$ 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 $\rho_\alpha(x) = \big(\alpha - \mathbf{1}(x \leq 0)\big)x$ is the tick loss function, and we specify q as a linear

[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

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

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 $\boldsymbol{w}_t^v = z_{i,t}^v$ is the i-th uncertainty measure at time t from vintage v. For the measures other than the EPU and JLN measures, $z_{i,t}^v = z_{i,t}$ since they are not subject to revisions.

Second, we consider the factors extracted from the uncertainty measures using PCA, $\boldsymbol{w}_t^v = \boldsymbol{f}_t^v$, the vector of k uncertainty factors at time t from vintage v. The factors are constructed each month in real-time based on the latest vintage data at the point of forecasting. 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

[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).

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. We use the 'unofficial' real-time version constructed by Amburgey and McCracken (2022) available, which coincides with the official Fed vintages from May 2011 onwards. It is available on McCracken's website.

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 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.

[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 (but published in real-time from the 2015M01 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 E.

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 $y_{t+h}^{h,\text{PBM}}$ 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.

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.

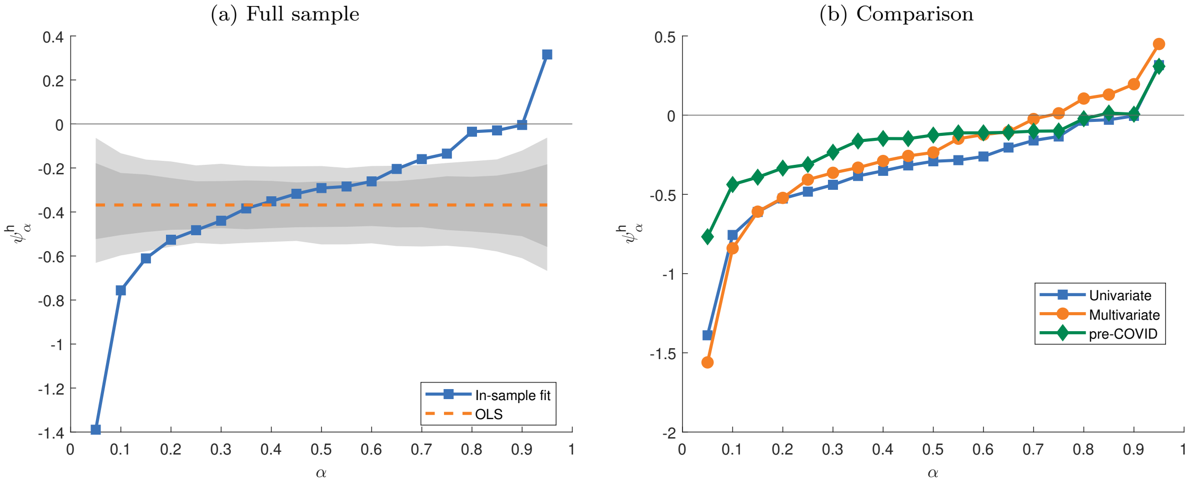
The parameter estimates in Figure 3 show that there is substantial evidence that

[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).

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 its constituents. 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 predictive ability of uncertainty for quantile forecasts of real activity, 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 more accurate quantile forecasts than the historical quantile benchmark.

We first examine the performance in forecasting the CEI. The RMTLs 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 about half of the 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 0.963 for quantile 0.1 and all but 2 measures yield an RMTL below 1, while it is on average 1.005 at quantile 0.5 with only 3 out of 15 measures outperforming the benchmark.

A few uncertainty measures stand out in terms of predictive power: OVX and VIX. 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.

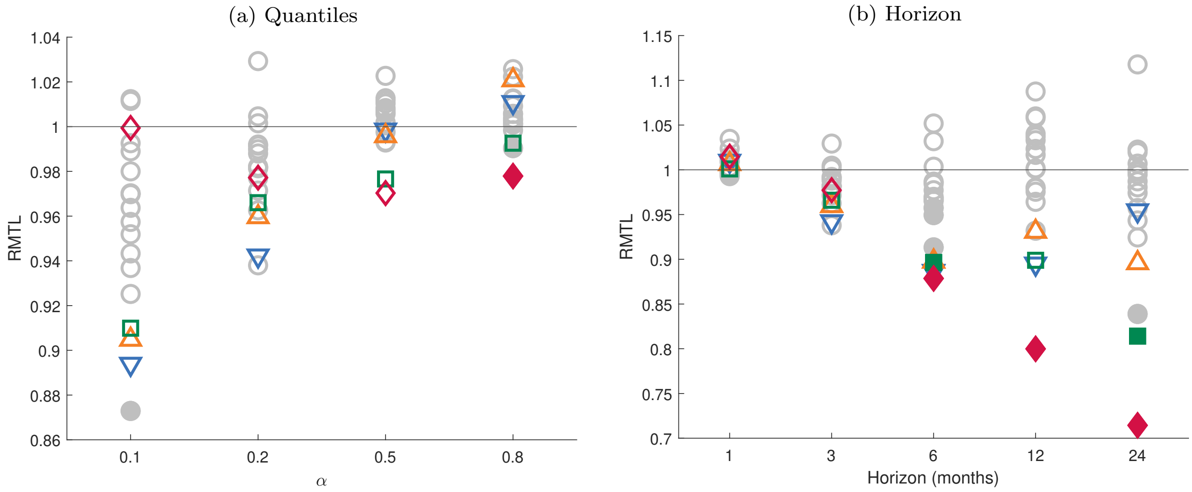
Second, the VIX yields good forecasts at medium horizons, with p-values under 10% between 3 and 12 month horizons at quantile 0.2. The gains in terms of tick loss are between 3% (3 month horizon) and 9% (6 month horizon) compared to the historical quantile.

Models with other individual 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.

6.2 Uncertainty factors

Turning to the uncertainty factor models, Figure 4 shows that they generally perform on par with or better than the best individual uncertainty measures. In particular at the medium horizon, the uncertainty factor models perform well, with gains in tick loss of

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. The RMTL values are presented in tables in Appendix E.

11.3% and 10.5% compared to the benchmark at the 6 and 12 month horizon for the 0.2 quantile for the one factor model. The forecasts are (close to) significantly better than the benchmark with p-values of 3.3% (6 month horizon) and 6.2% (12 month horizon). Importantly, the gains when using a factor are more consistent across horizons and (lower) quantiles compared to individual uncertainty measures.[9]

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

[9]In Section E.1, as an alternative aggregation method, we inspect the forecasting performance when averaging uncertainty measures within the categories from Section 2. In short, the results confirm the findings from the individual uncertainty measures: conditional volatility, cross-sectional dispersion, and forecast errors yield the smallest RMTL. Which category holds most predictive information depends on the horizon and target variable. Forecasts from the factor models are more accurate or on par with the category averages, and are more consistent across horizons and target variables.

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 models with individual uncertainty measures as well as 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.[10] 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 almost doubles from 0.566 to 1.105. Other models are impacted even more, as the relative tick loss also worsens compared to the pre-COVID period. Again for the 3 month horizon and the 0.2 quantile, the RMTL increases for 10 out of 15 uncertainty measures when including 2020 and 2021 in the sample period. Still, they at least somewhat capture the downturn as the average increase in RMTL over all uncertainty measures is only 2.2 percentage points. The same holds for the uncertainty factor models, where the RMTL increases with 0.5 to 3.0 percentage points. Ultimately though, the ordering of the models' performance and the significance levels of the forecasting performance are not affected much.

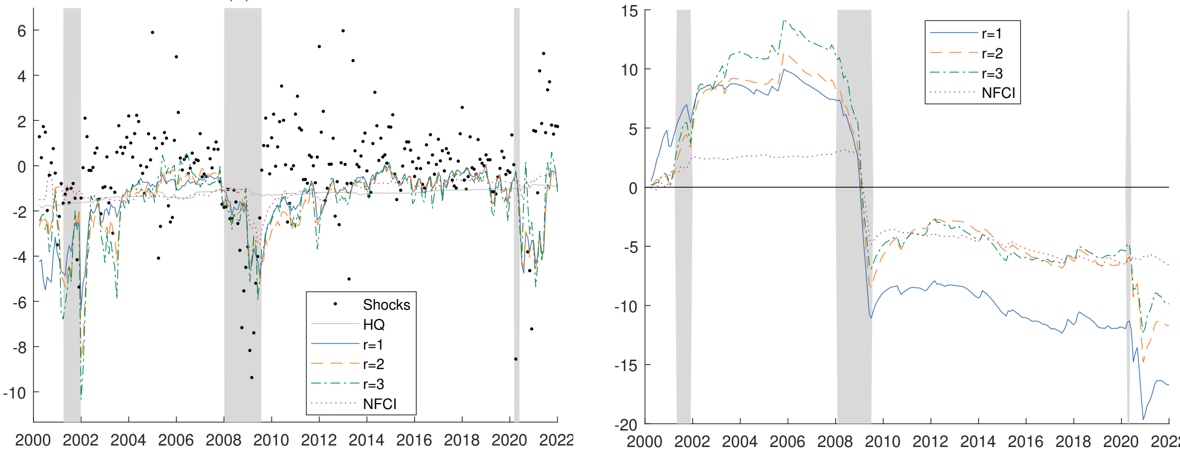
6.4 Forecasting the components of the coincident index

Forecasting results for the components of the coincident index (industrial production, employment, manufacturing and trade sales, and personal income) are largely consistent with those for the CEI. 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.

[10] 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).

Figure 5: Tick loss for coincident index over time

(a) Forecasts (b) Cumulative tick loss



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.

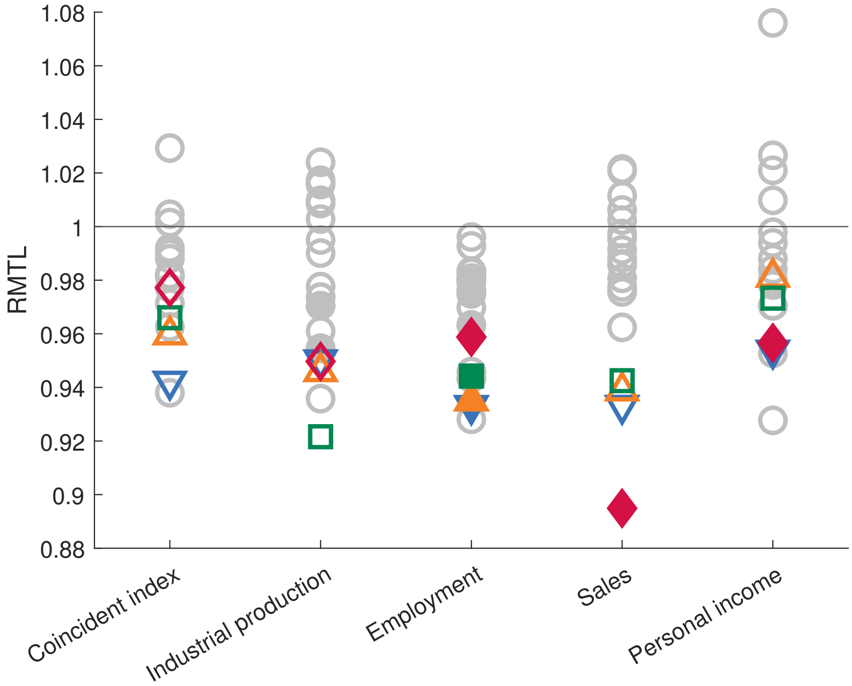
The forecasting results are most convincing for employment. At the 3 month horizon, for quantile 0.2 all uncertainty measures beat the historical quantile, and the uncertainty factors render the best performance with improvements of 5.6% to 6.7% in mean tick loss compared to the benchmark. Moreover, 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), and MPU. 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 short to medium forecasting horizon up to 12 months.

6.5 Relevance of real-time data

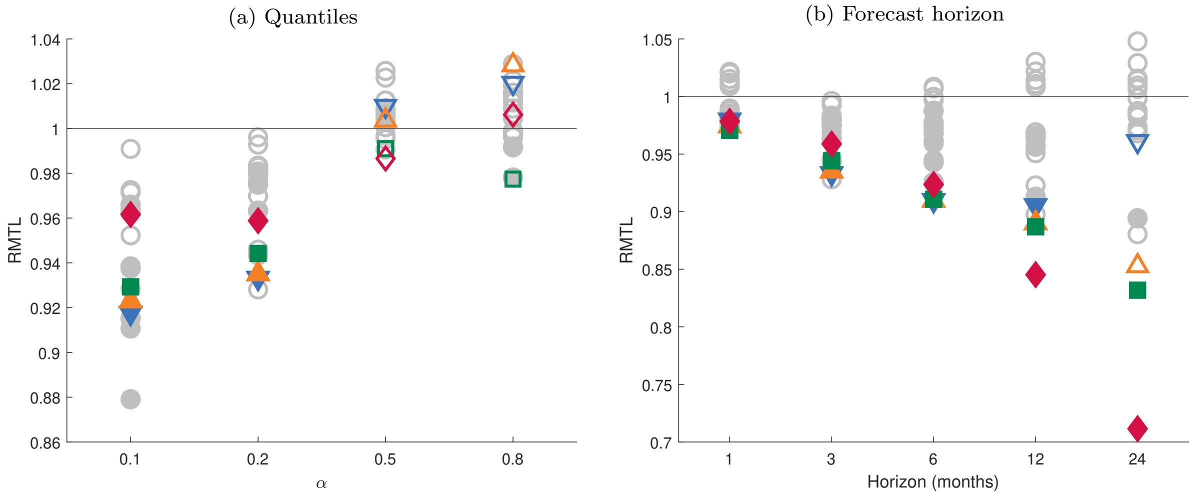
What are the consequences of ignoring revisions in the uncertainty measures and use the latest available observations? We address this question by repeating the forecasting exercise using the last vintage of the EPU and JLN measures (data up to 2021M12

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 E in tables.

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 E in tables.

as published on Ludvigson's website).[11] When using these observations instead of the real-time data, the RMTL for the models with the EPU measures still perform poorly. By contrast, the models with the final vintage JLN measures improve drastically, in line with findings by Rogers and Xu (2019). The RMTL decreases across all target variables from a 3 month or longer horizon. The relative gains are larger at longer horizons, with gains in the 3.4 to 16.3 percentage point range at the 6 and 12 month horizon for JLNm and JLNf. The most dramatic decrease in RMTL is 34.8 percentage points when forecasting employment at the 24 month horizon with the last vintage JLNr instead of the real-time version. All JLN measures would be among the best individual uncertainty measures.

Using final vintage data also improves the forecasting performance of the uncertainty factor models, in particular for employment. The RMTL for the two factor model decreases by 1.4 to 5.9 percentage points from the 3 to the 24 months horizon. When using real-time data, the results are more modest. A forecaster should therefore be cautious in interpreting forecasting results based on final vintage data.

6.6 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 F. At shorter horizons, they are slightly below the expected level of 0.2, but the null of correct coverage is not rejected for most models. As the forecast horizon increases, the hit rates increase and match expectations best at the 3 and 6 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

[11] We retain the real-time data for the target variables.

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.

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	0.208 ‡	0.241	0.216	0.257	0.278
Factor models					
$r = 1$	0.186	0.191	0.201	0.281	0.365 † ‡
$r = 2$	0.182	0.195	0.236	0.332 † ‡	0.386 † ‡
$r = 3$	0.186	0.183	0.209	0.316 † ‡	0.361 † ‡
Panel B: Employment					
HQ	0.136 † ‡	0.179	0.201	0.281	0.344 ‡
NFCI	0.114 † ‡	0.168	0.154	0.225	0.291
Factor models					
$r = 1$	0.117 † ‡	0.191	0.205	0.285	0.373 †
$r = 2$	0.125 † ‡	0.202	0.220	0.273	0.390 † ‡
$r = 3$	0.140 †	0.210	0.209	0.289	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.

7 Economic uncertainty and financial conditions

Adrian et al. (2019) show that the NFCI has predictive power for the left tail of US GDP growth. 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.

The correlations between the uncertainty measures and NFCI are all positive.[12] As expected, it is quite strongly correlated with the financial measures (65.7% with VIX, 57.1% with CSDR, and 56.4% with JLNf). The strongest correlation pre-COVID is actually with the forecast error based measures JLNm (79.1%) and JLNr (76.2%), but the response to COVID-19 was so different that the correlations drop to 53.6% and

[12]Correlations in this section are based on first release data, from Amburgey and McCracken (2022).

58.9% when including 2020 and 2021. 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 correlations with consumer survey based and news-based measures are more modest and in the range of 8.7% to 16.7%.

Figure 1 shows that the NFCI closely resembles the first uncertainty 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 60.7%, and even 80.3% 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 −41.4% (−28.9% 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 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 achieves the smallest tick loss in a substantial number of cases. When forecasting employment, the uncertainty factor models do edge out the NFCI.[13] At the 3 month horizon for the 0.2 quantile, the uncertainty factor models have a 1.5% (three factors) to 2.7% (one factor) lower tick loss compared to the NFCI – according to the DM tests there is no significant improvement though. There is some weak evidence at the 0.1 quantile with p-values of 5.8% (1 month horizon) and 7.2% (3 month horizon) that the factor models yields statistically better forecasts than the NFCI.

[13]Using last vintage date, the uncertainty factor models actually outperform NFCI at the 3, 6 and 24 month horizon when forecasting employment, see Section 6.5.

Even if NFCI outperforms the uncertainty measures and factor models in many cases, 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) quantile encompassing test.[14] The table shows whether none, both or only one of the variables should be included. In the latter case, one encompasses the excluded one. 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, Table 3 shows that at the 0.2 quantile it can be valuable to add uncertainty factors when forecasting at short to medium horizons up to 12 months. For employment at the 3 month horizon, there is some evidence that the uncertainty factor model encompasses NFCI. Second, the NFCI encompasses the uncertainty factor model in many cases, especially at horizons of 6 and 12 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. Diks et al. (2011) show that using a different scoring rule can substantially affect the conclusions though. Third, the uncertainty factor models comparatively do better further into the tail. At the 0.1 quantile it is more clear that the uncertainty factor model tends to encompass NFCI at shorter horizons up to 6 months, and vice versa at longer horizons.

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 (also 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

[14]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).

Table 3: Preferred model from encompassing test

	Horizon (months)				
Target variable	1	3	6	12	24
	$\alpha = 0.1$				
Coincident index	Uncertainty	Uncertainty	Both	None	None
Industrial production	Uncertainty	Both	NFCI	NFCI	NFCI
Employment	Uncertainty	Both	Uncertainty	Both	None
Manufacturing and trade sales	None	NFCI	NFCI	Both	Both
Personal income	Uncertainty	None	Uncertainty	Both	Both
	$\alpha = 0.2$				
Coincident index	Both	Uncertainty	Both	NFCI	None
Industrial production	NFCI	Uncertainty	Both	NFCI	None
Employment	Both	Uncertainty	Both	NFCI	NFCI
Manufacturing and trade sales	NFCI	NFCI	NFCI	NFCI	Both
Personal income	Uncertainty	NFCI	NFCI	NFCI	Both

The table presents what model is preferred, based on the Giacomini and Komunjer (2005) encompassing test at a 5% significance level, for forecasts of quantiles $\alpha = 0.1$ and $\alpha = 0.2$. The competing models are the factor model with the first two uncertainty factors and the model with NFCI.

best model at the longer horizons of 12 and 24 months, except for employment at the 24 month horizon.

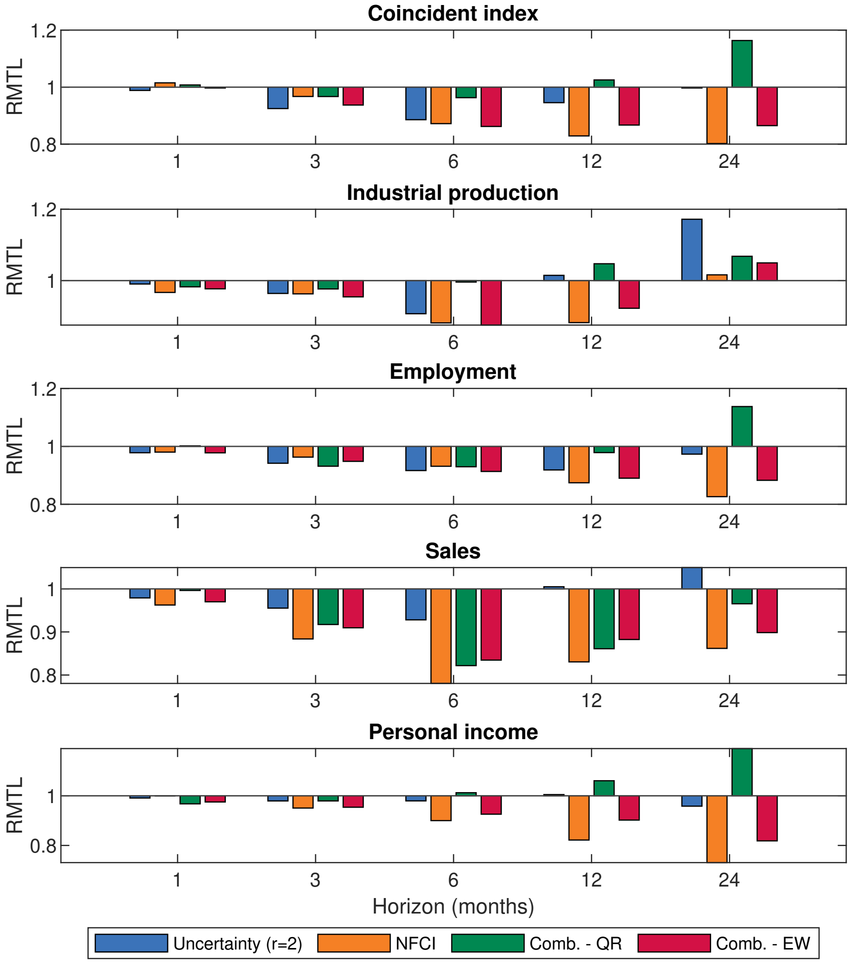
At shorter horizons up to 6 months, combining forecasts often yields a smaller RMTL. This is almost only when considering an equally weighted combination of the uncertainty factor model and NFCI model forecasts. The forecast combination based on quantile regression is actually the worst model in many cases. On average, the RMTL at the 3 and 6 month horizon is 1.4 and 6.3 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 seems to hold some relevant predictive information that is not captured by financial conditions.

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 financial conditions or macroeconomic uncertainty. Caldara et al. (2016) and Ludvigson et al.

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.

(2021) also differentiate between macroeconomic and financial uncertainty. Caldara et al. (2016) report a negative effect on economic activity from both sources. Ludvigson et al. (2021) 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.

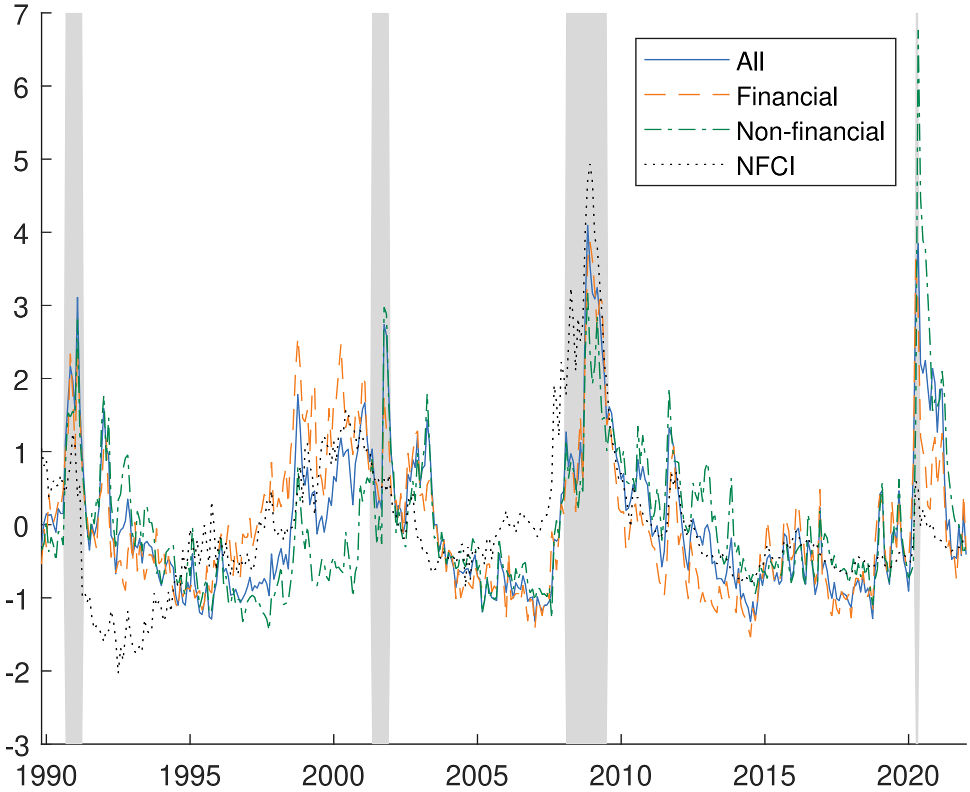
At first sight the first factors from both subsets of uncertainty measures seem to be very similar, see Figure 9; with correlations of 87.7% (financial) and 86.0% (non-financial) with the first factor obtained from the the full set. The correlation between the factors of the subsets is only 55.1% though. 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: 66.2%, which is higher than the factor based on the full set of measures (60.7%) or the non-financial based measures (34.8%).

The second factor from the full set is captured less strongly by the factors from the subsets. The largest (absolute) correlation is 73.5%, with the second factor from the non-financial subset, and 40.1% for the financial subset.

Since the NFCI outperforms the uncertainty factor models in many cases, we expect that the factors based on the financial-based uncertainty measures outperform the other factors models as well. 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. At horizons of 12 months and more though, the NFCI yields smaller RMTL. This could be due to that the NFCI is constructed from indicators of three categories: risk, credit, and leverage indicators. By contrast, the financial uncertainty measures are associated with the risk category only.

Interestingly, the second factor contains important predictive power when considering the non-financial set, in particular when forecasting employment. The second factor loads heavily on FDISP. Individually, FDISP doesn't decrease the MTL by much compared to the benchmark, up to 3.1% for a 12 month horizon or less. Though the gains are significant

Figure 9: First factor from subsets of measures



The figure presents the time series of first releases 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 gray bars are recessions as determined by NBER's Business Cycle Dating Committee. All series are standardized.

at all but the 3 month horizon. The non-financial factor models yield good forecasts for employment also in the first ten years of the sample, from January 2001 to December 2009, so including the financial crisis. The non-financial factor models yield the smallest tick loss with an RMTL of 0.820 (one factor) and 0.815 (two factors) at the 3 month horizon and 0.2 quantile. So it seems that most relevant information comes from the financial uncertainty measures, but not all predictive power.

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.

Table 4: RMTL for factor models on subsets

Horizon (months)	1	3	6	12	24
Panel A: Coincident economic index					
Financial information based measures					
$r = 1$	0.999	0.960**	0.906***	0.924**	0.931
$r = 2$	1.001	0.960**	0.923**	0.951*	0.933
Non-financial information based measures					
$r = 1$	1.022	0.966	0.962	0.990	0.992
$r = 2$	1.021	0.967	0.921	0.957	0.911
Panel B: Industrial production					
Financial information based measures					
$r = 1$	0.981*	0.943*	0.897**	0.942	1.022
$r = 2$	0.969*	0.945*	0.911*	0.942	1.028
Non-financial information based measures					
$r = 1$	0.982	0.989	0.976	1.009	0.992
$r = 2$	0.971*	0.976	0.896	0.909	0.902
Panel C: Nonfarm payroll employment					
Financial information based measures					
$r = 1$	0.972***	0.956***	0.930***	0.917***	0.965
$r = 2$	0.968***	0.954***	0.933***	0.930**	0.989
Non-financial information based measures					
$r = 1$	1.000	0.959	0.947	0.969	0.992
$r = 2$	0.983*	0.930**	0.908***	0.867**	0.797**
Panel D: Manufacturing and trade sales					
Financial information based measures					
$r = 1$	0.986*	0.946	0.946	0.946	1.045
$r = 2$	0.984	0.962	0.972	0.983	1.070
Non-financial information based measures					
$r = 1$	1.006	0.958	0.980	1.030	1.005
$r = 2$	1.003	0.936	0.886	0.902	0.911
Panel E: Personal income excluding transfer receipts					
Financial information based measures					
$r = 1$	0.962***	0.954*	0.988	0.997	0.955
$r = 2$	0.950***	0.962	0.981	1.007	0.990
Non-financial information based measures					
$r = 1$	0.977	1.011	1.015	1.028	0.982
$r = 2$	1.002	0.995	1.014	1.038	0.886

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.

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 and OVX are recommended individual measures, but using the factors is preferred for more consistent gains. The predictive content of economic uncertainty is relevant for professional forecasters and policy makers to keep an eye on, in particular when interested in the labor market. Moreover, at shorter horizons, the uncertainty measures seem to hold predictive content in addition to financial conditions as captured by the NFCI.

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.

Caldara, D., Fuentes-Albero, C., Gilchrist, S., and Zakrajšek, E. (2016). The macroeconomic impact of financial and uncertainty shocks. *European Economic Review*, 88:185–207.

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.

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.

Gürkaynak, R. S., Sack, B., and Wright, J. H. (2007). The us treasury yield curve: 1961 to the present. *Journal of Monetary Economics*, 54(8):2291–2304.

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.

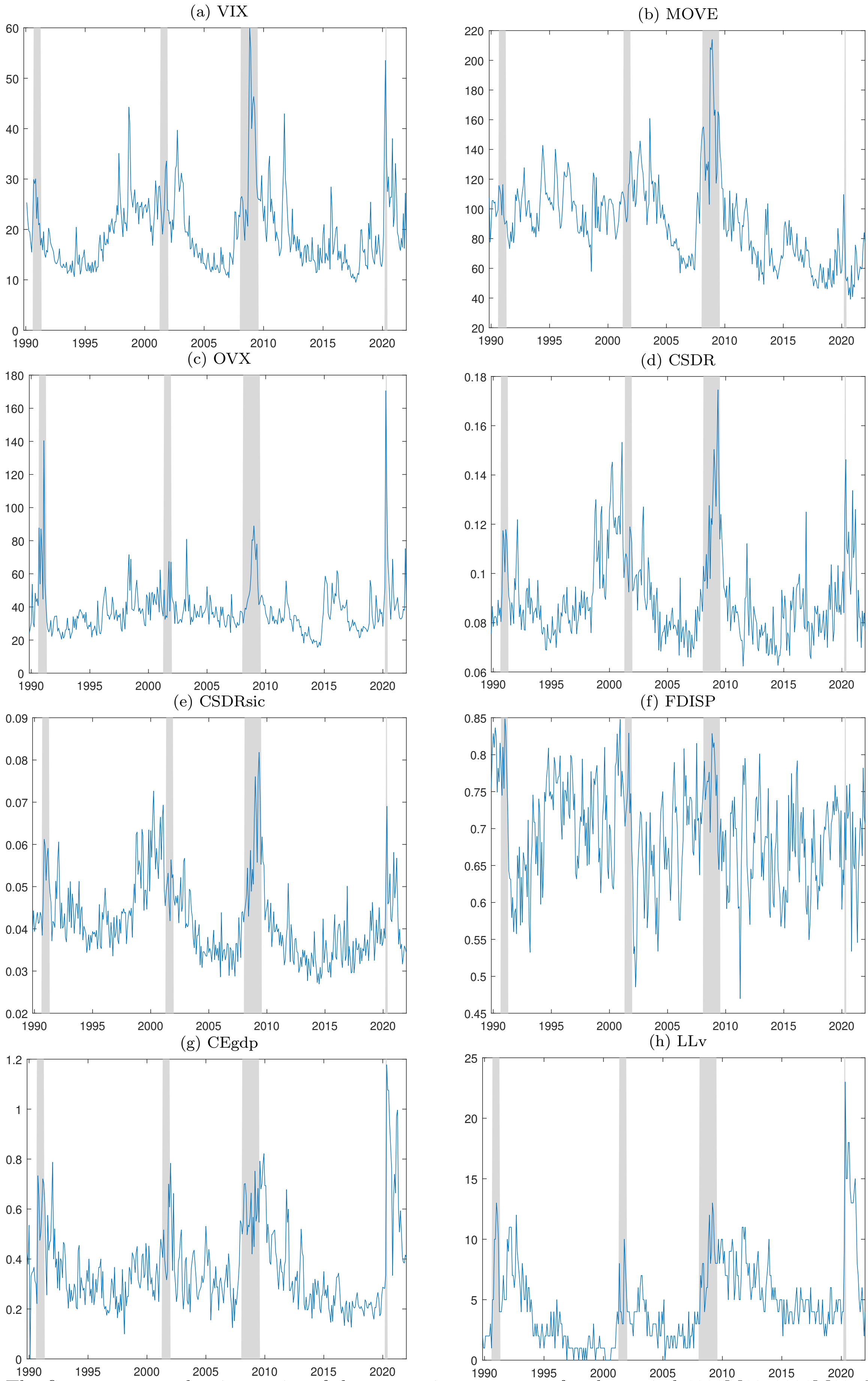
Appendix A Uncertainty measures

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)	ALFRED	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)	ALFRED	C	1900M1	1985M1
12	MPU	Monetary Policy Uncertainty; category of EPU, counts of articles additionally containing monetary policy related keywords	Baker et al. (2016)	ALFRED	C	1985M1	
13	JLNm	Macroeconomic variables' forecast error variance based on large factor model, horizon=12	Jurado et al. (2015)	FRED-MD, CRSP, Kenneth French's website, and Federal Reserve	E	1960M7	
14	JLNf	Financial variables' forecast error variance based on large factor model, horizon=12	Jurado et al. (2015)	FRED-MD, CRSP, Kenneth French's website, and Federal Reserve	E	1960M7	
15	JLNr	Real activity variables' forecast error variance based on large factor model, horizon=12	Jurado et al. (2015)	FRED-MD, CRSP, Kenneth French's website, and Federal Reserve	E	1960M7	

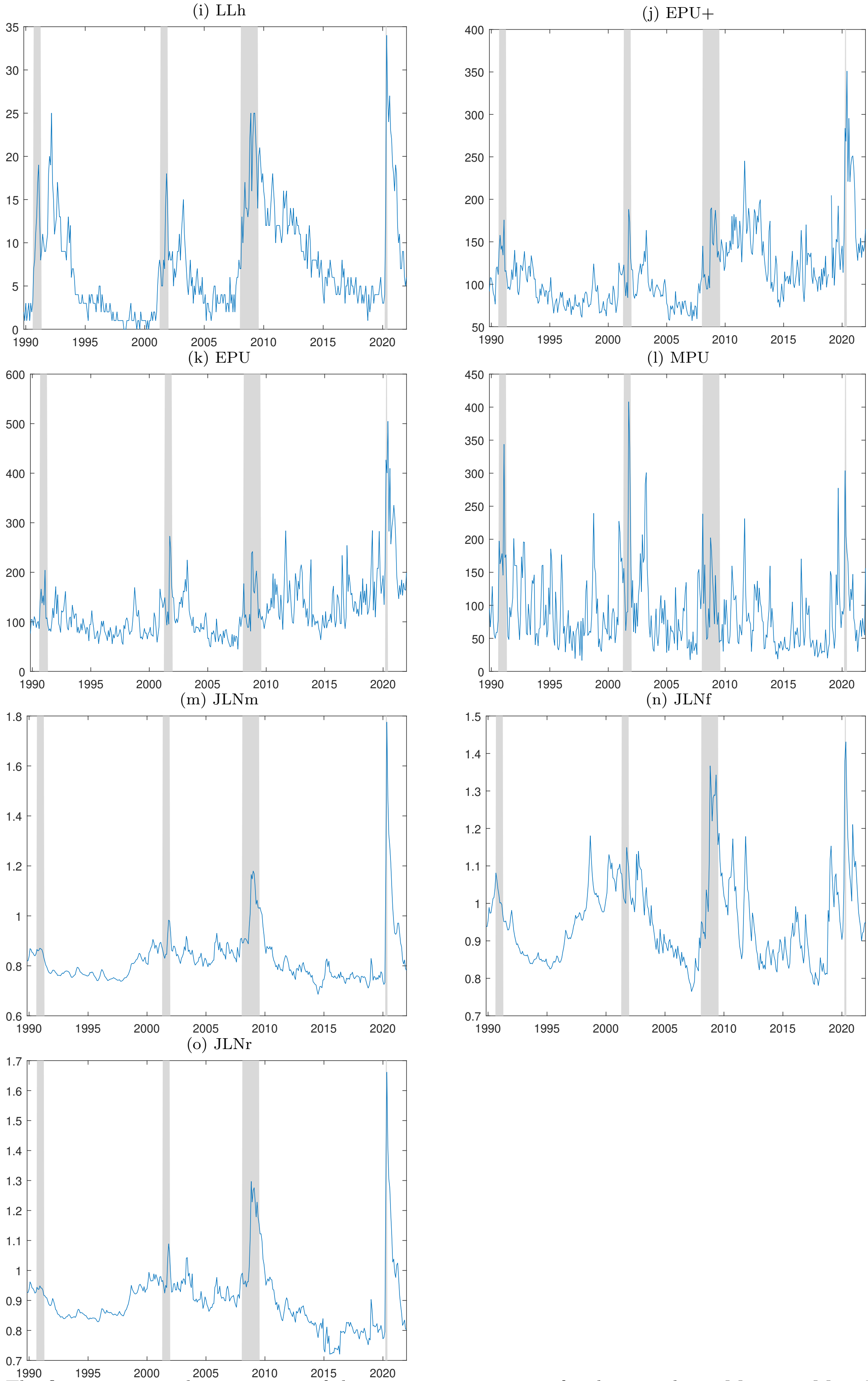
The table presents the uncertainty measures, their source, reference, uncertainty type (conditional volatility [A], cross-sectional dispersion [B], news [C], surveys [D], forecast errors [E]), start of the sample and when the data merge starts, where relevant.

Figure A.1: Uncertainty measure time series



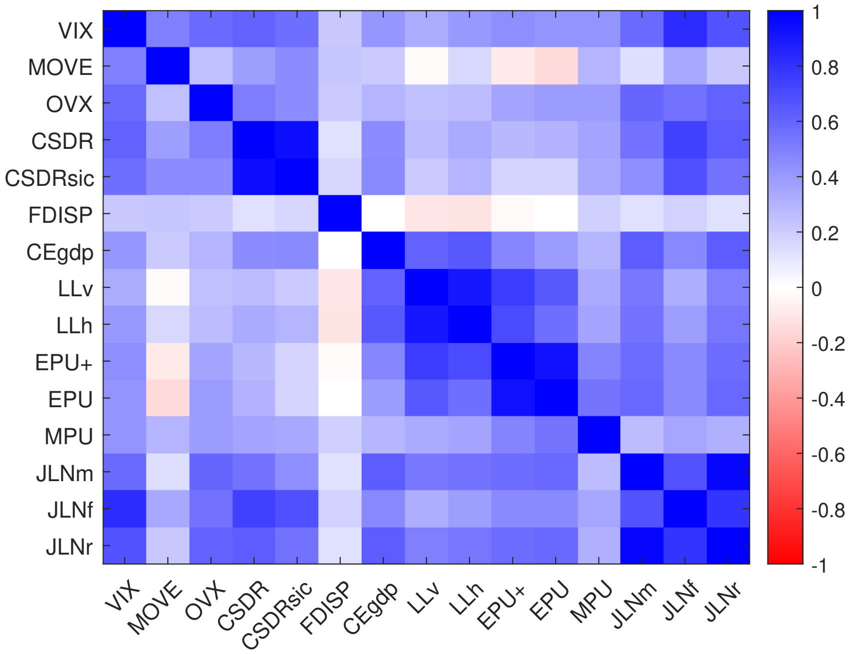
The figure presents the time series of the uncertainty measures for the period 1989M10–2021M12. First releases for the EPU and JLN measures.

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



The figure presents the time series of the uncertainty measures for the period 1989M10–2021M12. First releases for the EPU and JLN measures.

Figure A.2: Correlation matrix



The figure presents the correlation matrix of the uncertainty measures for the period 1989M10–2021M12. Last vintage for EPU and JLN measures (2022M1).

Appendix B Real-time data: uncertainty measures

Two types of uncertainty measures are subject to revisions: the newspaper article based measures of Baker et al. (2016), and the Jurado et al. (2015) (JLN) measures.

B.1 Economic policy uncertainty

The newspaper based uncertainty measures of Baker et al. (2016) are subject to revisions because there is a delay in posting all the articles online for some newspapers. The first vintages for these uncertainty measures are 2013M6 (EPU+) and 2019M10 (EPU and MPU), available in ALFRED. Observations before that period are from the first vintage.

Two data issues need to be addressed. First, the vintages 2018M12 and 2019M1 of EPU+ are missing. This is solved by imputing the x-th release data for those months with an AR(p) model with up to 6 lags and the number of lags p selected by BIC. We only use the observations that were available at that time, so only data prior to the 2018M12 vintage. Second, the reporting lag of MPU varies across the vintages from 1 to 2 months. Most of the time it is 2 months, which is why we include it with a publication lag of 2 months.

B.2 JLN measures

The JLN measures depend on the macroeconomic variables that are subject to revisions, so the measures themselves are subject to revisions. Sydney Ludvigson publishes the JLN measures semi-annually, not monthly.[15] Therefore, we construct a monthly real-time version, using the methodology of Jurado et al. (2015).[16]

The first vintage is 1999M8, equal to the first available FRED-MD vintage. We compute the vintages each month up to 2022M1, such that we have a total of 270 vintages. The sample starts in 1960M1, in line with Jurado et al. (2015). The JLN measures are computed for horizons up to $h = 12$ months. JLN publish the 1, 3, and 12 month version. We use the 3 month version, but using the 12 month version does not affect our

[15] Ludvigson's website: `https://www.sydneyludvigson.com/macro-and-financial-uncertainty-indexes`

[16] Thanks to Jurado, Ludvigson and Ng for sharing their code on Ludvigson's website.

conclusions. Data sources are the FRED-MD, CRSP and Kenneth French's website. The Cochrane-Piazzesi factor is constructed using the Gürkaynak et al. (2007) nominal yield curve data, available at the Federal Reserve website: `https://www.federalreserve.gov/data/nominal-yield-curve.htm`.

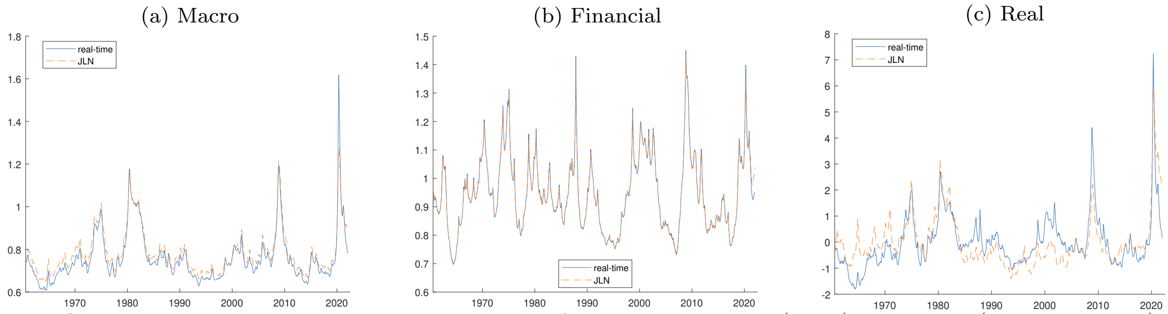
JLN use a balanced panel. However, in the monthly real-time setting, the latest month is not available due to differences in reporting lag of some of the macro series. Sydney Ludvigson publishes semi-annual updates with a longer lag, which avoids this issue. The methods used to construct the JLN measures are robust to missing values (at least at the start and end of the sample). So instead of a longer reporting lag or removing the variable, we use an imbalanced panel. When there are observations missing in the middle of the sample, we use an AR(p) model within that vintage to impute its values. This is only for three observations: one in the series COMPAPFFx (2020M4) and two in CP3Mx (2020M4 and 2020M5).

Two variables are adjusted or removed to handle outliers. First, the oil price is in log differences instead of the twice log differenced series. The FRED-MD suggests transformations to ensure stationarity for each of its macro series, and recommends twice log differencing the oil price. However, using log returns makes more sense in terms of interpretation and we don't find evidence of non-stationarity. Additionally, using the second log differenced series leads to very high (individual) uncertainties, dominating the JLN measures at multiple vintages. Second, the variable NONBORRES is excluded from the vintages 2008M3 and 2008M4. In those vintages, the predicted individual uncertainty for NONBORRES is extremely high. The reason is an outlier in the underlying data: in February 2008 the percentage change (transformation recommended by FRED-MD) is very high because the previous value is relatively close to 0 (-800 in January compared to -16,300 in February 2008). Finally, eight times a series is excluded because it has too few available observations in that vintage.[17]

Figure B.1 shows plots our real-time version (vintage 2021M7) and the JLN measures

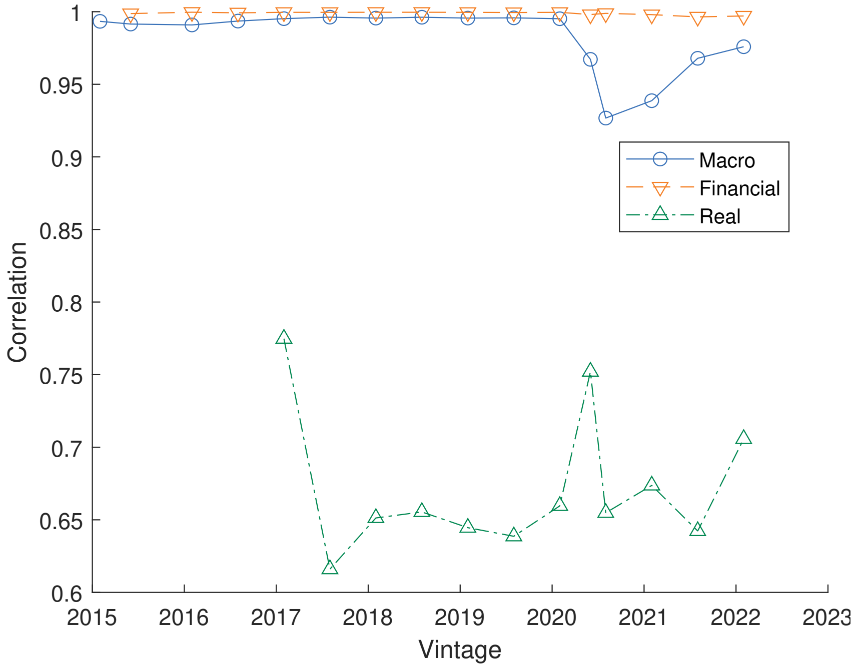
[17] RPI from vintage 1999M11; DSERRG3M086SBEA, DNDGRG3M086SBEA, DDURRG3M086SBEA, PCEPI, W875RX1, and RPI from vintage 2003M12; CMRMTSPLx from vintage 2015M8.

Figure B.1: Real-time JLN uncertainty measures



The figures are our real-time vintage 2021M7 of the Jurado et al. (2015) measures (solid blue line) and the one published on Ludvigson's website (dashed orange line) for the $h = 3$ month horizon version. The 'real' version are standardized for ease of comparison.

Figure B.2: Correlation between real-time version and Jurado et al. (2015)



The figure presents the correlation between the real-time version and the one published on Ludvigson's website for the available vintages there, for the $h = 3$ month horizon version. The sample runs from 1960M1 until the end of the vintage.

published on Ludvigson website for the $h = 3$ month horizon. From the time series and the correlations between the real-time version and the one published by JLN in Figure B.2, it is clear that our version is very close the original. The 'financial' version is exactly equal in 7 out of 15 vintages and the correlation is at least 99.6%. The correlations are also very high for the 'macro' version with at least 92.7%, and up to 99.6%. The similarity of the 'real' version is a bit less strong, but still quite high as the correlation fluctuates between 61.6% and 77.5% depending on the vintage.

The main reason for the lower correlation for the real version is a difference in the underlying data. As of June 2016, the series NAPMPI, NAPMEI, NAPM, NAPMNOI, NAPMSDI, NAPMII, and NAPMPRI are no longer included in the FRED-MD vintages. These series have been manually updated by JLN from the original source (the Institute for Supply Management Report on Business) since its removal. We do not have access to the historical data and cannot append it to the FRED-MD. To see the impact, we can compare the correlation of the last published version of the JLN 'real' measure (data up to 2022M6) with last vintage of the real-time version that included NAPM (2016M5) and the correlation with the vintage on month later, for the same sample (1960M1−−2016M4). The correlation decreases from 87.4% to 66.7%, due to the exclusion of the NAPM series.

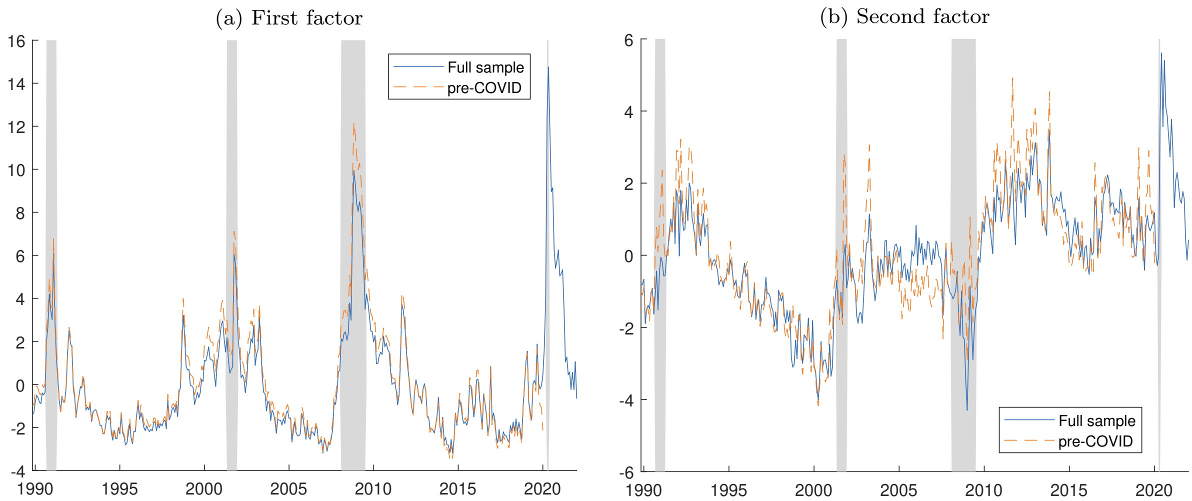
Overall, our real-time version is close to the original by Jurado et al. (2015), and the 'real' version could be improved further by including the NAPM series.

Rogers and Xu (2019) also construct a real-time version of the JLN measures. Our version differs in a number of ways. First, Rogers and Xu remove a substantial number of the FRED-MD set, see notes 17 and 18 in their paper. They use 120 out of 132 of the macroeconomic variables of the FRED-MD from the 2004M1 vintage onwards. The main reason is that some series, e.g. the NAPM series, are excluded from the FRED-MD at some point. They opt to remove them from preceding vintages as well. Instead, we take the perspective that the real-time forecaster could not have predicted this change and keep them in the earlier vintages. Second, Rogers and Xu exclude four variables in the factor estimation: 'MZMSL', 'DTCOLNVHFNM', 'DTCTHFNM', 'INVEST'. Third, Rogers and Xu (2019) strictly work with a balanced panel, starting in 1978M6. Our version allows for missing observations and starts in 1960M1.

Appendix C Uncertainty factors and COVID-19

Figure C.1 shows the first two uncertainty factors for the full sample and when excluding 2020–2021. The first and second factor are clearly identified both in both samples, and follow the same general pattern. The first factor is nearly identical (99% correlation), while the correlation is also high (89.0%) for the second factor between the different samples. The difference between the samples is due to MOVE deviating more from the average uncertainty measure during the COVID-19 period. The correlations of MOVE with other uncertainty measures before 2020 are on average 36.9%. When 2020 and 2021 are included, the correlations decrease on average by 16.3 percentage points and even become negative in three cases, down to -14.4% with EPU. This results in a weaker loading on the first factor and an increase in the second factor's loading for MOVE. In turn, the loadings of various news and consumer based measures (LLh, LLv and EPU) in the second factor decrease.

Figure C.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, vintage 2022M1 (solid blue line), and the first two factor estimated on the pre-COVID sample, 1989M10–2019M12, vintage 2020M1 (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 D 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.[18] 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

[18] 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 E Relative mean tick loss

Table E.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	1.015	0.977	0.879***	0.800**	0.714**
FRED-MD	1.011	1.069	1.098	1.003	0.922
Factor models					
$r = 1$	1.010	0.942*	0.887**	0.895*	0.955
$r = 2$	1.007	0.960	0.898**	0.931	0.896
$r = 3$	1.001	0.966	0.896**	0.899*	0.814**
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+	1.009	1.029	1.052	1.088	1.023
EPU	1.015	1.005	0.986	1.033	1.020
MPU	1.002	0.982	0.949***	0.980	0.998
JLNm	1.035	0.990	0.984	1.037	1.007
JLNf	0.998	0.988	0.968	0.980	1.000
JLNr	1.015	0.977	0.977	1.060	1.118
Panel B: Industrial production					
MTL of HQ	3.080	2.064	1.882	1.590	1.356
NFCI	0.964**	0.950*	0.865*	0.835	0.818
FRED-MD	1.001	1.069	1.031	1.038	1.050
Factor models					
$r = 1$	0.973*	0.950	0.878*	0.978	1.025
$r = 2$	0.973*	0.946	0.886*	0.939	0.950
$r = 3$	0.958**	0.922*	0.837**	0.915	0.896
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+	0.994	1.016	1.044	1.026	1.006
EPU	0.981	0.990	0.995	1.008	1.051
MPU	0.970**	0.955**	0.969	0.994	1.020
JLNm	0.995	1.003	0.981	1.018	1.006
JLNf	0.990	0.974	0.972	0.985	0.999
JLNr	0.981	0.995	0.985	1.110	1.146

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 E.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	0.978***	0.959**	0.924**	0.845**	0.712**
FRED-MD	0.939***	0.997	1.005	0.971	0.973
Factor models					
$r = 1$	0.980**	0.933***	0.910***	0.905**	0.961
$r = 2$	0.974***	0.935***	0.910***	0.890***	0.853*
$r = 3$	0.971***	0.944**	0.911***	0.887***	0.832**
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	0.979	1.008	1.030	1.016
EPU	1.012	0.976	0.997	1.008	1.029
MPU	0.981***	0.975***	0.968***	0.969*	1.007
JLNm	1.009	0.928*	0.926*	0.898*	0.880
JLNf	0.983**	0.970*	0.959*	0.923*	0.968
JLNr	0.985	0.946*	0.945	0.951	0.999
Panel D: Manufacturing and trade sales					
MTL of HQ	3.419	1.884	1.420	1.224	1.065
NFCI	0.965***	0.895**	0.800**	0.759**	0.758**
FRED-MD	1.035	1.031	1.040	1.075	1.059
Factor models					
$r = 1$	0.985	0.933	0.925	0.958	1.043
$r = 2$	0.992	0.939	0.916	0.893	0.925
$r = 3$	1.001	0.943	0.895	0.922	0.895
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+	1.024	1.021	1.029	1.069	0.962
EPU	1.028	1.022	1.018	1.032	0.968
MPU	1.016	0.995	0.998	1.016	0.995
JLNm	0.995	0.986	0.950	0.987	1.041
JLNf	1.000	0.981	0.982	0.998	1.121
JLNr	0.983	1.002	0.937	1.014	1.124

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 E.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	1.004	0.957**	0.903**	0.848*	0.680**
FRED-MD	0.982	0.957	1.001	1.006	0.960
Factor models					
$r = 1$	0.950**	0.954	0.972	0.992	0.935
$r = 2$	0.997	0.981	0.986	1.016	0.888
$r = 3$	0.990	0.973	0.975	0.934	0.822**
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+	1.036	1.076	1.073	1.041	1.102
EPU	1.032	1.021	1.067	1.012	1.048
MPU	0.994	0.988	0.997	0.994	1.019
JLNm	0.972	0.970	1.047	1.002	0.960
JLNf	0.989	0.998	1.068	1.040	1.006
JLNr	0.956**	0.953	1.053	1.087	1.067

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 E.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	0.977	0.939	0.998	0.952
FRED-MD	1.069	0.847	1.187	1.061
Factor models				
$r = 1$	0.942*	0.832**	1.001	0.912
$r = 2$	0.960	0.860**	1.013	0.952
$r = 3$	0.966	0.842**	1.032	0.961
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+	1.029	0.854	1.123	1.082
EPU	1.005	0.914	1.053	1.034
MPU	0.982	0.986	0.980	0.992
JLNm	0.990	0.863	1.057	0.990
JLNf	0.988	0.930	1.019	0.985
JLNr	0.977	0.907	1.015	0.981
	Panel B: Industrial production			
MTL of HQ	2.064	6.378	1.548	1.465
NFCI	0.950*	0.824**	1.012	0.917*
FRED-MD	1.069	0.956	1.124	1.084
Factor models				
$r = 1$	0.950	0.829**	1.010	0.907*
$r = 2$	0.946	0.822**	1.007	0.907*
$r = 3$	0.922*	0.823**	0.970	0.900*
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+	1.016	0.908	1.069	1.004
EPU	0.990	0.895	1.037	0.967
MPU	0.955**	0.963	0.951	0.937*
JLNm	1.003	0.838	1.084	0.955
JLNf	0.974	0.859	1.030	0.952
JLNr	0.995	0.784	1.099	0.975

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. 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. See Table A.1 for an explanation of the abbreviations.

Table E.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	0.959**	0.877**	0.987**	0.842**
FRED-MD	0.997	0.913	1.025	0.933
Factor models				
$r = 1$	0.933***	0.884**	0.950**	0.829**
$r = 2$	0.935***	0.893**	0.949**	0.826***
$r = 3$	0.944**	0.895**	0.961*	0.862**
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+	0.979	0.906	1.004	1.016
EPU	0.976	0.909	1.000	1.024
MPU	0.975***	0.961	0.980	0.940**
JLNm	0.928*	0.884	0.943	0.883*
JLNf	0.970*	0.907	0.991	0.918
JLNr	0.946*	0.877	0.969	0.893
Panel D: Manufacturing and trade sales				
MTL of HQ	1.884	6.873	1.287	1.492
NFCI	0.895**	0.692***	1.025	0.842**
FRED-MD	1.031	0.700	1.242	0.968
Factor models				
$r = 1$	0.933	0.727***	1.065	0.875*
$r = 2$	0.939	0.724***	1.076	0.885*
$r = 3$	0.943	0.703***	1.096	0.877*
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+	1.021	0.724	1.211	1.009
EPU	1.022	0.782	1.175	1.010
MPU	0.995	0.948	1.026	0.977
JLNm	0.986	0.702	1.167	0.913
JLNf	0.981	0.852	1.063	0.962
JLNr	1.002	0.768	1.152	0.966

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. 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. See Table A.1 for an explanation of the abbreviations.

Table E.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	0.957**	0.934	0.965**	0.931**
FRED-MD	0.957	0.824	1.008	0.957
Factor models				
$r = 1$	0.954	0.857*	0.991	0.910*
$r = 2$	0.981	0.907	1.010	0.944
$r = 3$	0.973	0.868*	1.014	0.954
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+	1.076	0.916	1.137	1.083
EPU	1.021	0.941	1.052	1.010
MPU	0.988	0.993	0.986	0.985
JLNm	0.970	0.815	1.030	0.887*
JLNf	0.998	0.980	1.005	0.982
JLNr	0.953	0.883	0.980	0.886*

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. 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. See Table A.1 for an explanation of the abbreviations.

Table E.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	0.999	0.977	0.970*	0.978**
FRED-MD	1.058	1.069	1.023	1.033
Factor models				
$r = 1$	0.894*	0.942*	0.999	1.011
$r = 2$	0.905	0.960	0.996	1.021
$r = 3$	0.910	0.966	0.977	0.993
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+	1.012	1.029	1.006	1.009
EPU	1.012	1.005	0.993	1.012
MPU	0.970	0.982	0.993	1.001
JLNm	0.989	0.990	1.012	1.026
JLNf	0.925	0.988	1.001	1.013
JLNr	0.964	0.977	1.010	1.022
		Panel B: Industrial production		
MTL of HQ	1.595	2.064	2.477	1.848
NFCI	0.925*	0.950*	0.991	1.005
FRED-MD	1.139	1.069	1.043	1.024
Factor models				
$r = 1$	0.914*	0.950	1.004	1.007
$r = 2$	0.897*	0.946	1.002	1.007
$r = 3$	0.894*	0.922*	0.956**	0.966
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+	1.055	1.016	1.014	1.012
EPU	1.034	0.990	1.001	1.030
MPU	0.980	0.955**	0.976***	0.989**
JLNm	0.962	1.003	1.034	0.995
JLNf	0.948	0.974	1.004	1.021
JLNr	0.970	0.995	1.019	0.983

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 E.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	0.962**	0.959**	0.987	1.006
FRED-MD	1.011	0.997	1.020	1.021
Factor models				
$r = 1$	0.918***	0.933***	1.010	1.020
$r = 2$	0.923***	0.935***	1.004	1.028
$r = 3$	0.929***	0.944**	0.991	0.977*
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+	0.928*	0.979	1.013	1.022
EPU	0.915*	0.976	1.008	1.014
MPU	0.966***	0.975***	0.997	0.996*
JLNm	0.879**	0.928*	1.023	1.029
JLNf	0.938***	0.970*	1.013	1.009
JLNr	0.911**	0.946*	1.009	1.016
Panel D: Manufacturing and trade sales				
MTL of HQ	1.404	1.884	2.239	1.646
NFCI	0.850**	0.895**	0.966	0.974
FRED-MD	1.075	1.031	1.108	1.128
Factor models				
$r = 1$	0.900*	0.933	1.021	1.011
$r = 2$	0.920	0.939	1.026	1.012
$r = 3$	0.956	0.943	1.012	0.989
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+	1.052	1.021	1.038	1.018
EPU	1.041	1.022	1.028	1.023
MPU	1.023	0.995	1.008	0.982
JLNm	0.961	0.986	1.044	1.069
JLNf	0.935*	0.981	1.011	1.014
JLNr	0.933	1.002	1.036	1.022

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 E.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	0.950**	0.957**	0.992	0.989**
FRED-MD	0.923	0.957	0.975	1.006
Factor models				
$r = 1$	0.944	0.954	0.974	1.022
$r = 2$	0.953	0.981	0.983	1.033
$r = 3$	0.935	0.973	0.978	1.017
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+	1.070	1.076	1.022	1.049
EPU	1.085	1.021	1.017	1.054
MPU	0.991	0.988	1.007	1.002
JLNm	0.958	0.970	1.008	1.082
JLNf	0.979	0.998	0.988	1.002
JLNr	0.952	0.953	0.991	1.034

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.

E.1 Forecast results: averaging within category

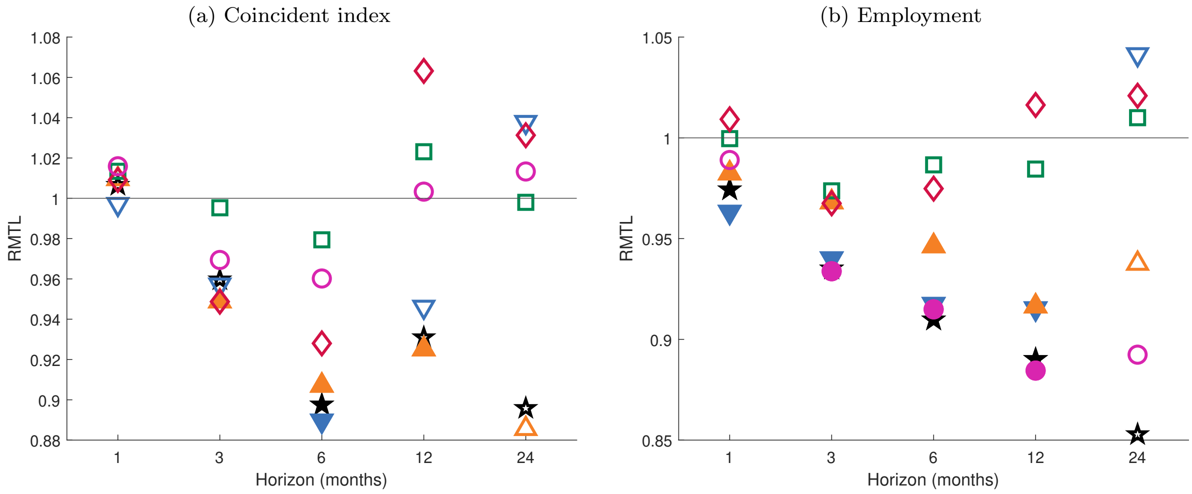
Applying principal components is one possible way of aggregating the uncertainty measures. An obvious alternative would be to use the categories as identified in Section 2 to aggregate the uncertainty measures. This could also provide more insight into how each category contributes to the forecasting results. We aggregate by taking the average of the (standardized) uncertainty measures within each category.

Focusing on forecasting the coincident index and employment, the results in Figure E.1 show that there are clear differences across the categories. These are in line with what we found with the individual uncertainty measures. Indeed, the news and survey averages yield comparatively worse forecasts than the other categories.

Further, the results show that it is worthwhile to combine information, and most predictive information is in the three categories conditional volatility, cross-sectional dispersion, and forecast errors. Conditional volatility performs well at the shorter horizon, up to 6 months. Cross-sectional dispersion is among the best models when forecasting the coincident index. The forecast error average leads to the lowest tick loss across these averaged uncertainty measures when forecasting employment for all horizons except at the 1 month horizon.

Figure E.1 shows that RMTL from the two uncertainty factor model is smaller or on par with those of the category averages. The factor model combines all uncertainty measures, which yields more robust predictive performance over the forecast horizons for CEI and employment.

Figure E.1: Tick loss for averages within category



The figure presents the relative mean tick loss (RMTL), with the historical quantile as benchmark, from forecasting the coincident economic index (left figure) and employment (right figure) across forecasting horizons in months. The quantile is $\alpha = 0.2$. Black stars are RMTL from the two uncertainty factor model. Other RMTLs are from the average of the measures within the categories conditional volatility (blue down-pointing triangle), cross-sectional dispersion (orange up-pointing triangles), news (green squares), surveys (red diamonds), and forecast errors (magenta circles). Filled symbols indicate significance of the one-sided DM test against the historical quantile at the 5% significance level.

Appendix F Hit rates

Table F.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	0.208 ‡	0.241	0.216	0.257	0.278
FRED-MD	0.208 ‡	0.233 ‡	0.270 ‡	0.265 ‡	0.361 † ‡
Factor models					
$r = 1$	0.186	0.191	0.201	0.281	0.365 † ‡
$r = 2$	0.182	0.195	0.236	0.332 † ‡	0.386 † ‡
$r = 3$	0.186	0.183	0.209	0.316 † ‡	0.361 † ‡
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	0.168 ‡	0.174	0.253 ‡	0.320 ‡
EPU	0.167	0.187	0.178	0.241	0.315 ‡
MPU	0.189	0.199	0.232	0.265	0.328 ‡
JLNm	0.193 ‡	0.214	0.174 ‡	0.245 ‡	0.361 † ‡
JLNf	0.193	0.199	0.205	0.269 ‡	0.274 ‡
JLNr	0.189 ‡	0.229	0.243	0.285 ‡	0.465 † ‡
		Panel B: Industrial production			
HQ	0.246	0.233	0.247 ‡	0.324 ‡	0.361 ‡
NFCI	0.231	0.218	0.205	0.277	0.390 †
FRED-MD	0.284 † ‡	0.294 † ‡	0.274 ‡	0.391 † ‡	0.448 † ‡
Factor models					
$r = 1$	0.246	0.241	0.255	0.344 † ‡	0.411 † ‡
$r = 2$	0.242	0.252	0.305 †	0.391 † ‡	0.490 † ‡
$r = 3$	0.250 †	0.233	0.274	0.399 † ‡	0.473 † ‡
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+	0.231	0.210 ‡	0.247 ‡	0.336 † ‡	0.353 ‡
EPU	0.193	0.187 ‡	0.247	0.308 ‡	0.353 ‡
MPU	0.258 †	0.256	0.247	0.348 †	0.340 ‡
JLNm	0.227	0.214	0.216	0.285	0.415 † ‡
JLNf	0.239	0.244	0.293	0.348 †	0.373 †
JLNr	0.239	0.225	0.266	0.304 ‡	0.436 † ‡

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 F.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 † ‡	0.168	0.154	0.225	0.291
FRED-MD	0.140 †	0.199	0.216	0.277	0.390 † ‡
Factor models					
$r = 1$	0.117 † ‡	0.191	0.205	0.285	0.373 †
$r = 2$	0.125 † ‡	0.202	0.220	0.273	0.390 † ‡
$r = 3$	0.140 †	0.210	0.209	0.289	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	0.224	0.277 ‡	0.332 ‡
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 † ‡	0.164	0.162	0.241	0.373 † ‡
JLNf	0.117 † ‡	0.191	0.209	0.237	0.349
JLNr	0.133 † ‡	0.202	0.243	0.296	0.432 † ‡
			Panel D: Manufacturing and trade sales		
HQ	0.193	0.263	0.270 ‡	0.296	0.261 ‡
NFCI	0.186	0.267	0.270	0.265	0.365 † ‡
FRED-MD	0.227 ‡	0.286 † ‡	0.290 † ‡	0.300 ‡	0.357 † ‡
Factor models					
$r = 1$	0.178	0.248	0.309 † ‡	0.328 †	0.336
$r = 2$	0.186	0.256	0.348 † ‡	0.348 † ‡	0.427 † ‡
$r = 3$	0.178	0.244	0.305 † ‡	0.360 † ‡	0.328 ‡
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+	0.167	0.191 ‡	0.255 ‡	0.332 † ‡	0.270
EPU	0.186	0.199 ‡	0.259	0.308	0.307 ‡
MPU	0.212	0.271 †	0.301 † ‡	0.296	0.266 ‡
JLNm	0.171	0.233 ‡	0.297 † ‡	0.281 ‡	0.477 † ‡
JLNf	0.201	0.263	0.309 † ‡	0.273	0.261 ‡
JLNr	0.197	0.302 † ‡	0.348 † ‡	0.356 † ‡	0.465 † ‡

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 F.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	0.186	0.219	0.241	0.316 †	0.272
FRED-MD	0.205	0.234	0.245 ‡	0.296	0.357 † ‡
Factor models					
$r = 1$	0.178	0.227	0.253	0.340 †	0.366 † ‡
$r = 2$	0.155	0.223	0.269	0.364 † ‡	0.396 † ‡
$r = 3$	0.143 † ‡	0.227	0.281	0.377 † ‡	0.298
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 ‡	0.210 ‡	0.271 ‡	0.285 ‡
EPU	0.143 † ‡	0.164	0.225	0.255	0.281 ‡
MPU	0.171	0.223	0.245	0.300	0.306 ‡
JLNm	0.167	0.199	0.245	0.316 ‡	0.404 † ‡
JLNf	0.167	0.219	0.225 ‡	0.287	0.251 ‡
JLNr	0.209	0.227	0.269 ‡	0.344 † ‡	0.400 † ‡

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.