Schick, Manuel

Working Paper

Real-time nowcasting Growth-at-Risk using the Survey of Professional Forecasters

AWI Discussion Paper Series, No. 750

Provided in Cooperation with:
Alfred Weber Institute, Department of Economics, University of Heidelberg

Suggested Citation: Schick, Manuel (2024) : Real-time nowcasting Growth-at-Risk using the Survey of Professional Forecasters, AWI Discussion Paper Series, No. 750, University of Heidelberg, Department of Economics, Heidelberg, https://doi.org/10.11588/heidok.00034995

This Version is available at:
https://hdl.handle.net/10419/301190

HEIDELBERG UNIVERSITY
DEPARTMENT OF ECONOMICS

UNIVERSITÄT
HEIDELBERG
ZUKUNFT
SEIT 1386

Real-time Nowcasting Growth-at-Risk using the Survey of Professional Forecasters

Manuel Schick

AWI DISCUSSION PAPER SERIES NO. 750
June 2024

Real-time Nowcasting Growth-at-Risk using the Survey of Professional Forecasters*

Manuel Schick†

June 10, 2024

Abstract

This paper investigates nowcasting Growth-at-Risk (GaR) using consensus forecasts from the Survey of Professional Forecasters (SPF) in the US. Incorporating SPF consensus forecasts into the conditional mean of an AR-GARCH type model significantly enhances nowcasting accuracy for GaR and the conditional density of GDP growth. While there is strong time variation in both the lower and upper quantiles of the GDP growth distribution, integrating skewness and fat tails into the model does not improve forecasting accuracy. By accounting for changes in the conditional mean of the GDP growth distribution over time, these findings highlight the value of SPF consensus projections for GaR nowcasting.

Keywords: Growth-at-Risk, GARCH, Survey of Professional Forecasters
JEL Classification: C22, C52, C53

*I would like to thank Sarah Arndt, Christian Conrad, Timo Dimitriadis, Thomas Eife, Alexander Glas, Fabian Krueger, Faek Menla Ali, Anne Opschoor, and Julius Schoelkopf, as well as seminar and conference participants at the International Conference on Computational and Financial Econometrics (December 2022), the HKMetrics Research Network Workshop at the Karlsruhe Institute of Technology (July 2023), and the University of Heidelberg (May 2024) for helpful comments and suggestions.

†Department of Economics, Heidelberg University, Email: manuel.schick@awi.uni-heidelberg.de.

1 Introduction

In monitoring the economic outlook, the focus of researchers and policymakers has recently shifted toward measuring economic downside risks. In particular, Growth-at-Risk (GaR), popularized by Adrian et al. (2019), typically represents the 5% or 10% quantile of the conditional GDP growth distribution as a measure of tail risk. This shift has prompted major institutions, including the International Monetary Fund (IMF), and central banks like the Federal Reserve Bank of New York and the European Central Bank (ECB), to publish GaR forecasts for major economies.

This paper proposes nowcasting quarterly GaR in real-time by leveraging the consensus forecasts from the US Survey of Professional Forecasters (SPF). Specifically, I use the median point predictions of the SPF as a predictor for the conditional mean of quarterly GDP growth. To assess the uncertainty around consensus forecasts, I employ a Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model. The GARCH model, incorporating SPF consensus projections, provides timely nowcasts of GaR as well as the entire conditional density of GDP growth. A realistic out-of-sample evaluation demonstrates that incorporating the current economic outlook through SPF consensus forecasts significantly improves predictive performance for GaR and the conditional density of GDP growth.

So far, a substantial part of the literature employs Quantile Regressions (QR) to assess economic downside risk, indicating that lower quantiles of GDP growth vary with financial conditions (Giglio et al., 2016; Adrian et al., 2019, 2022; Ferrara et al., 2022). Conditional on financial indicators, the predictive GDP growth distribution exhibits time-varying lower quantiles, while the center and the upper part of the distribution are relatively constant over time. This has led to a surge in research findings that GDP growth is negatively skewed during economic downturns and vice versa (see, for example, Adrian et al., 2019; Delle Monache et al., 2023).

In the context of predicting GaR, Brownlees and Souza (2021) provide evidence that a standard volatility model, such as the GARCH(1,1) model, tends to exhibit better out-of-sample predictive performance than the QR approach. Moreover, their findings suggest that the conditional GDP growth distribution is not significantly skewed and is moderately fat-tailed. Similarly, Carriero et al. (2022), including a large set of macroeconomic indicators in quantile regression models, also find that skewness is not a robust feature of GDP growth. Overall, these results highlight that both asymmetry and time-variation in higher moments crucially depend on the choice of the model and the conditioning information.

What these models have in common is that the practical application for real-time economic monitoring faces two main challenges. The first challenge stems from the delayed releases of macroeconomic indicators, often referred to as the *ragged-edge* problem (Wallis,

1986). For instance, the first estimate of US GDP is typically released at the end of the first month of the subsequent quarter, while the nowcast from the US SPF becomes available mid-quarter. Additionally, macroeconomic indicators are available at different frequencies. While GDP is measured quarterly, some indicators like the Chicago Fed's National Financial Conditions Index (NFCI) are reported weekly, and others are even reported daily. This discrepancy in reporting frequencies necessitates the use of models capable of handling mixed-frequency data (see, for example, Ghysels et al., 2004, 2007; Giannone et al., 2008; Andreou et al., 2013). In the context of GaR, the timely incorporation of new data releases and revisions is crucial. Carriero et al. (2022) demonstrate that exploiting a large set of macroeconomic indicators in real-time improves the nowcast accuracy throughout the current quarter. Similarly, Ferrara et al. (2022) utilize daily financial conditions indicators in real-time for daily GaR predictions.

To circumvent the problem of delayed releases and mixed-frequency data, I exploit the consensus point predictions of the SPF. Rather than relying on a large set of macroeconomic indicators with varying release frequencies, the SPF offers a comprehensive assessment about the current economic outlook. The SPF is a quarterly survey that asks a panel of professional forecasters to provide point predictions of quarterly GDP for the current quarter and up to four quarters ahead (Clements et al., 2023). In particular, at short horizons of up to one quarter ahead, the SPF consensus point forecasts often outperform standard time-series models (Stark et al., 2010; Faust and Wright, 2013). This superior performance can likely be attributed to the aggregation of forecasts from multiple experts who efficiently incorporate a large pool of information and quickly adapt to major changes in the economic environment (Ang et al., 2007).

Regarding the prediction of conditional quantiles of GDP growth, the SPF is relatively unexplored. Adams et al. (2021) use financial conditions to explain fluctuations in the uncertainty around SPF consensus forecasts. Based on the conditional forecast error distribution, they construct GaR nowcasts, finding that incorporating financial conditions leads to significant out-of-sample improvements compared to GaR forecasts constructed solely from past forecast errors. The SPF also elicits density forecasts, which have been studied in the context of Growth-at-Risk by Ganics et al. (2020) and of Inflation-at-Risk by Andrade et al. (2014).

This paper contributes to the literature in several ways. First, it bridges the gap between the existing evidence on the short-horizon accuracy of the SPF consensus forecasts and the predictive performance of the GARCH model for nowcasting GaR. By integrating the SPF consensus forecasts into the GARCH model, this paper provides a tool for timely assessments of economic downside risks. Second, it relates to the nowcasting literature by leveraging the SPF in real-time, utilizing the timely information captured by the SPF

consensus forecasts. This approach incorporates the current economic outlook without necessitating the inclusion of potentially numerous macroeconomic indicators with varying frequencies, delayed releases, and later revisions. Thirdly, this paper contributes to the literature on assessing the uncertainty around consensus forecasts. Instead of eliciting entire probability distributions, this paper derives conditional densities by exploiting the forecast errors of the survey consensus forecasts (Reifschneider and Tulip, 2019; Clark et al., 2020; Adams et al., 2021).

The real-time out-of-sample evaluation reveals that the consensus forecasts of the SPF offer valuable information for GaR prediction. When conditioned on the SPF's median nowcasts, the GARCH model significantly outperforms quantile regressions with NFCI in predicting lower quantiles of the GDP growth distribution. Moreover, the NFCI provides limited additional predictive gains once the SPF nowcasts become available mid-quarter. Furthermore, the SPF's consensus projections are not only informative about downside risk but also about the upper quantiles of the GDP growth distribution. By accounting for changes in the conditional mean of the GARCH model, the SPF enhances forecasts of the entire GDP growth distribution. In addition, the out-of-sample analysis indicates that incorporating skewness and fat tails does not lead to an improved forecasting performance. While this observation does not necessarily imply symmetry in the true conditional distribution of GDP growth, it suggests that employing a more flexible distribution than the normal distribution lacks justification from a short-horizon forecasting perspective.

The remainder of the paper is organized as follows. Section 2 provides a review of the related literature. Section 3 details the data and the associated release calendar. The forecasting models are explained in Section 4, followed by an explanation of the evaluation metrics in Section 5. Section 6 discusses the findings of the out-of-sample analysis and provides robustness checks. Finally, Section 7 concludes the paper.

2 Related Literature

Growth-at-Risk as a risk-measure gained prominence following its adoption by institutions like the IMF (Prasad et al., 2019), making it a focal point for research on economic tail risks modeled through predictive quantile regressions (see, for example, De Nicolò and Lucchetta, 2017; Adams et al., 2021; Amburgey and McCracken, 2023). In particular, due to the influential paper by Adrian et al. (2019), much of the empirical literature focuses on the negative correlation of financial conditions and economic downside risks. They find that lower quantiles of the conditional GDP growth distribution significantly vary with financial conditions whereas the upper part of the distribution is relatively stable over time. This time-varying asymmetry led to the conclusion that GDP growth

exhibits negative skewness during recessions and is closer to being symmetric or even positively skewed during expansions (Delle Monache et al., 2023). Similarly, Giglio et al. (2016) provide evidence that systemic risk measures have an asymmetric effect on the conditional distribution of macroeconomic outcomes. However, rather than interpreting this as evidence in favor of time-varying skewness, they argue that systemic risk measures are more informative about the lower tail than the central tendency or the upper tail of macroeconomic indicators.

The existing literature also suggests that the predictive distribution of GDP depends on the conditioning information and the methods employed. For example, Carriero et al. (2022) employ Bayesian quantile regressions with a rich set of macroeconomic indicators, revealing that GDP growth exhibits only a slight degree of asymmetry. In a complementary approach, Brownlees and Souza (2021) utilize a GARCH model to analyze GDP growth, estimating the innovation distribution nonparametrically. Their findings suggest that GDP growth is not significantly skewed and displays moderately fat tails. Furthermore, the finding of Adrian et al. (2019) that only the lower tail of GDP growth is time-varying could also be explained by a symmetric distribution that exhibits simultaneous changes in the mean and variance (Carriero et al., 2020). This suggests that predicting GaR entails modeling the conditional mean and variance, along with possibly kurtosis, rather than focusing on time-varying skewness (Fagiolo et al., 2008; Figueres and Jarociński, 2020; Plagborg-Møller et al., 2020).

There is also a more critical perspective on the predictive ability of financial conditions for growth vulnerability. For instance, Reichlin et al. (2020) find little additional predictive power of financial conditions for the lower tail of GDP when real economic indicators are incorporated. Similarly, Plagborg-Møller et al. (2020) demonstrate that predictive quantile regressions using both financial and economic variables only provide limited information about the conditional mean of GDP growth at very short horizons, with imprecise estimates for higher moments and longer horizons. The evidence suggests that aggregate financial indexes exhibit strong co-movements with real variables that are related to GDP growth. This would imply that movements in the NFCI are endogenous to economic conditions and explain why the marginal predictive information of financial conditions beyond macroeconomic variables is modest (Plagborg-Møller et al., 2020). Nevertheless, given the strong correlation of the NFCI and real variables relevant to GDP growth, the NFCI's inclusion in forecasting models may still offer valuable insights into Growth-at-Risk dynamics.

This paper also directly draws on the work of Brownlees and Souza (2021), who apply the GARCH model to a country-panel for GaR forecasting. Their study shows that the purely backward-looking univariate GARCH model tends to outperform the quantile

regression approach that includes the NFCI in predicting the lower tail of GDP growth. Additionally, they find that the density forecasts generated by the GARCH model are more accurate. While their findings raise further questions about the effectiveness of using the NFCI as an advanced risk measure, it is worth noting that they focus on forecast horizons of at least one quarter ahead, in contrast to nowcasting current quarter GDP growth. However, their results underscore the advantage of modeling the volatility of GDP growth through dynamics captured in the GARCH model, which avoids the need to specify predictors for tail risks that may not exhibit robust predictive ability over time. By remaining agnostic about the source of economic downside risks, the GARCH model can yield better forecasts even in the case of misspecification (Brownlees and Souza, 2021).

One limitation is that applying a univariate GARCH model for monitoring GaR only exploits past information through lags of GDP. By incorporating the SPF into the conditional mean of the GARCH model, this paper is directly related to the literature on constructing predictive densities by assessing the uncertainty around survey consensus forecasts. For example, Reifschneider and Tulip (2019) discuss how the Federal Open Market Committee (FOMC) of the US Federal Reserve uses historical forecast errors of the *Summary of Economic Projections* to gauge the FOMC's forecasting uncertainty. Based on the root mean squared forecast errors estimated on a rolling window, they construct unconditional confidence bands around consensus projections of the FOMC. While many central banks report fan charts based on the size of past forecast errors, these approaches assume that forecast uncertainty is constant or evolves slowly. However, Clark et al. (2020) find that historical forecast errors exhibit notable time variation. Specifically for the US SPF, they show that modeling the forecast error variance through stochastic volatility enhances density forecasts. Similarly, Adams et al. (2021) use quantile regressions incorporating financial conditions to explain variations in the uncertainty around SPF point projections, demonstrating that conditioning on the NFCI results in significantly improved predictive densities around SPF point forecasts.

By leveraging the SPF median forecasts to nowcast GaR, this paper draws on the existing evidence on the forecasting accuracy of SPF consensus projections. While studies have noted biases in individual survey forecasts (see, for example, Zarnowitz, 1985; Elliott et al., 2008), consensus forecasts of professional forecasters offer accurate macroeconomic predictions (Ang et al., 2007; Giannone et al., 2008; Bańbura et al., 2013, 2021), and often outperform standard time-series models, particularly at shorter horizons (Stark et al., 2010; Faust and Wright, 2013). For instance, Stark et al. (2010) find that the real-time SPF median projections for current quarter GDP outperform model estimates. However, they also highlight the rapid deterioration of forecast accuracy for horizons beyond one quarter.

While the SPF also provides density forecasts through histograms, which have been explored in the context of macroeconomic tail risks (Andrade et al., 2014; Ganics et al., 2020), this paper opts against utilizing them for several reasons. One significant drawback is that these density forecasts pertain to current and next-year GDP growth, making them fixed-event forecasts. Ganics et al. (2020) propose a density combination approach to obtain fixed-horizon forecasts that are well calibrated. However, their method relies on a continuous distribution, which needs to be constructed from the bin probabilities. Inferring aggregate predictive densities from individual survey histograms presents challenges, with common approaches involving fitting a parametric distribution to the empirical CDF implied by the bin probabilities and aggregating over individual forecaster densities, not necessarily in that order. Moreover, difficulties arise from changes in bin widths over time and unbounded intervals at the tails. In addition, forecasters contribute to challenges by rounding their forecasts and occasionally assigning zero probability to some bins (Bassetti et al., 2023).

Despite efforts, judgmental density forecasts are often found to be misspecified and poorly calibrated (Rossi and Sekhposyan, 2019). Krüger and Plett (2024) find that methods based on the forecast error distribution provide better interval estimates than those implied by SPF density forecasts. In addition, subjective density means are less accurate and updated less frequently than point predictions from survey panelists (Clements, 2010, 2014). Lastly, subjective distributions may not accurately reflect GDP growth but rather indicate survey panelists' confidence in their forecasts. Bańbura et al. (2021) observe overconfidence among SPF respondents and document that integrating point predictions improves model-based density forecasts, while second moments compromise predictive accuracy.

Lastly, this paper relates to the large body of literature on real-time nowcasting, which focuses on updating GDP forecasts as new data becomes available. This approach, formalized by Giannone et al. (2008) and surveyed by Bańbura et al. (2013), highlights the importance of incorporating data releases and revisions in real-time to improve the accuracy of GDP growth forecasts. Similarly, incorporating high-frequency financial data in real-time improves out-of-sample nowcasts of GDP growth both in terms of the central tendency (Andreou et al., 2013) and tail risks (Ferrara et al., 2022). In the context of GaR, Carriero et al. (2022) demonstrate that exploiting a large set of macroeconomic indicators in real-time improves the nowcast accuracy throughout the current quarter.

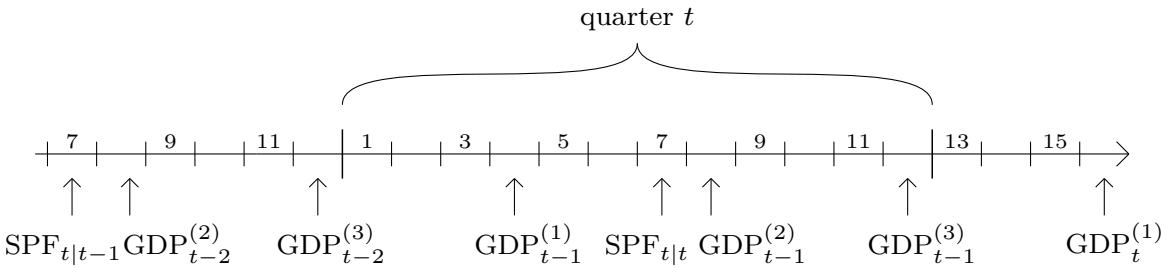
For a realistic out-of-sample nowcasting evaluation, real-time vintage data on GDP is readily available, and surveys like the SPF provide real-time forecasts (Clements et al., 2023). However, the first vintage of NFCI data only became available in 2011. Consequently, a substantial amount of evidence is based on the final NFCI vintage observed at the end

of the evaluation period (see, for example, Adrian et al., 2019; Reichlin et al., 2020; Brownlees and Souza, 2021; Carriero et al., 2022). To avoid look-ahead bias in the out-of-sample analysis, I rely on unofficial real-time weekly vintages of the NFCI constructed by Amburgey and McCracken (2023).

3 Real-time data and predictors

This paper focuses on real-time nowcasting of Growth-at-Risk in quarter t, necessitating the use of real-time data to avoid look-ahead bias. Given the delayed publication of GDP data, Figure 1 provides an overview of the releases and revisions of quarterly GDP.

Figure 1: Release calender



Notes: This graph sketches the release and revision dates of GDP together with the release dates of the SPF forecasts. The numbers on the time line indicate the weeks in each quarter.

In the fourth week of quarter t, the Bureau of Economic Analysis (BEA) releases the first estimate of the previous quarter's GDP, denoted by $\text{GDP}^{(1)}_{t-1}$. Consequently, during the early weeks of quarter t, the GDP of $t-1$ remains unavailable. Approximately one and two months after the initial release, the BEA issues the first and second major revisions, leading to the second ($\text{GDP}^{(2)}_{t-1}$) and third ($\text{GDP}^{(3)}_{t-1}$) estimate of GDP, respectively. Analogously, the initial estimate for current-quarter GDP, $\text{GDP}^{(1)}_{t}$, is published at the end of the first month of quarter $t+1$.

The release calendar depicted in Figure 1 also shows release dates of the SPF. The SPF is a quarterly survey of macroeconomic forecasts for the US. Starting in the fourth quarter of 1968, the survey was initially conducted by the American Statistical Association and the National Bureau of Economic Research. Since 1990, the Federal Reserve Bank of Philadelphia (FED) has been running the survey that consists of roughly 40 professional forecasters (Clements et al., 2023). For GDP, based on the median SPF projections, the FED provides forecasts for the annualized percentage points quarter-over-quarter growth rates. This includes nowcasts for the current quarter and up to forecasts of one-year ahead quarterly growth rates. Usually, the survey is released mid-quarter, i.e., around the sixth to seventh week of a quarter.

Table 1 summarizes the latest available information in real-time on a weekly basis, omitting even numbered weeks to save space. To reconstruct the information set available in each week of quarter t, I take monthly vintages of real GDP from the Real-Time Data Set for Macroeconomists of the FED and match them with the historical release dates available on ALFRED.[1] For quarterly GDP, in week w of quarter t, let $y_{s,w}$ denote the annualized quarterly log growth rates for quarters $s < t$ according to

$$y_{s,w} = y_{s|t,w} = 400 \cdot \log(\text{GDP}_{s|t,w} - \text{GDP}_{s-1|t,w}), \tag{1}$$

where $\text{GDP}_{s|t,w}$ denotes real GDP in quarter s obtained from the monthly vintage available in week w of quarter t. Due to delayed releases of GDP, in the first three weeks of quarter t, since GDP in $t-1$ is not available, growth rates are only obtained for $s = 1, \ldots, t-2$. In the following weeks, $w = 4, \ldots, 15$, a forecaster has information on GDP for $s = 1, \ldots, t-1$, with a first revision, $\text{GDP}^{(2)}_{t-1}$, between the 7th and the 9th week.

Table 1: Real-time information available for predicting GDP growth in quarter t

week (w)	Variables
1	$\text{GDP}^{(3)}_{t-2} = \text{GDP}_{t-2\|t,12}$, $\text{SPF}_{t\|t-1}$, $\text{NFCI}_{t-1,12\|t,1}$
3	$\text{GDP}^{(3)}_{t-2} = \text{GDP}_{t-2\|t,12}$, $\text{SPF}_{t\|t-1}$, $\text{NFCI}_{t,2\|t,3}$
5	$\text{GDP}^{(1)}_{t-1} = \text{GDP}_{t-1\|t,4}$, $\text{SPF}_{t\|t-1}$, $\text{NFCI}_{t,4\|t,5}$
7	$\text{GDP}^{(1)}_{t-1} = \text{GDP}_{t-1\|t,4}$, $\text{SPF}_{t\|t}$, $\text{NFCI}_{t,6\|t,7}$
9	$\text{GDP}^{(2)}_{t-1} = \text{GDP}_{t-1\|t,8}$, $\text{SPF}_{t\|t}$, $\text{NFCI}_{t,8\|t,9}$
11	$\text{GDP}^{(2)}_{t-1} = \text{GDP}_{t-1\|t,8}$, $\text{SPF}_{t\|t}$, $\text{NFCI}_{t,10\|t,11}$
13	$\text{GDP}^{(3)}_{t-1} = \text{GDP}_{t-1\|t,12}$, $\text{SPF}_{t\|t}$, $\text{NFCI}_{t,12\|t+1,1}$
15	$\text{GDP}^{(3)}_{t-1} = \text{GDP}_{t-1\|t,12}$, $\text{SPF}_{t\|t}$, $\text{NFCI}_{t+1,2\|t+1,3}$

Notes: This table demonstrates the latest realizations of GDP, SPF, and the NFCI available for odd numbered weeks throughout quarter t and the first three weeks of quarter $t+1$.

Concerning the SPF, at any given moment and across forecast horizons, the survey projections are real-time forecasts. To reconstruct the information flow provided by the SPF thus only requires knowledge about the timing of the survey. Since the FED is in charge of the SPF, release dates are available on the web page of the FRED, permitting an exact alignment with the available information set.[2] In the first five to six weeks of

[1] The BEA reported GNP until the end of 1991. Thus, for a substantial part of the sample, GDP refers to GNP.

[2] For the first quarter in 1990, the FED is uncertain about the release date. The release date is thus

each quarter, only the one-step ahead projections, $\text{SPF}_{t|t-1}$, are available. Consequently, one can only exploit the nowcasts, $\text{SPF}_{t|t}$, in the second half of a quarter. To facilitate notation, let $SPF_{s,w}$ denote the SPF median projections for $s = 1, \ldots, t$, available in week $w = 1, \ldots, 15$ of quarter t according to

$$SPF_{t,w} = \begin{cases} \text{SPF}_{t|t-1} & \text{if } w \leq 6 \\ \text{SPF}_{t|t} & \text{if } w > 6 \end{cases}. \tag{2}$$

Lastly, Table 1 displays data releases of the Chicago Fed's NFCI. The NFCI is a weekly estimate of financial conditions in the US, incorporating a comprehensive set of financial indicators including money markets, debt and equity markets, and the banking system. Its construction is described in detail in Brave and Butters (2018). The index is standardized to have a mean of zero and a standard deviation of one, with positive readings historically associated with tighter than average financial conditions, and vice versa.

Official real-time vintages of the NFCI are obtainable on ALFRED. As the first vintage became available in 2011, many studies employ the final vintage of the NFCI. However, pseudo out-of-sample analyses ignore the filtering uncertainty that primarily affects the real-time NFCI at sample endpoints (Brownlees and Souza, 2021). Moreover, the NFCI incorporates factors directly dependent on GDP, such as the corporate debt-to-GDP ratio, which can introduce look-ahead bias when used in pseudo out-of-sample analyses. Therefore, I rely on unofficial real-time weekly vintages of the NFCI constructed by Amburgey and McCracken (2023). These unofficial vintages begin in 1988 and include NFCI data dating back to 1973.

The NFCI is typically released on Wednesdays, incorporating data up to the preceding Friday, resulting in a one-week publication delay. Considering the quarterly frequency of GDP data, a simple aggregation scheme is to align weekly NFCI data in week w at a quarterly frequency such that

$$NFCI_{s,w} = \text{NFCI}_{s,w-1|t,w}, \tag{3}$$

where $\text{NFCI}_{s,w-1|t,w}$ is the NFCI in week $w-1$ of quarter s, available from the latest NFCI vintage in week w of quarter t. For instance, in the first week of quarter t, $NFCI_{s,1} = \text{NFCI}_{s-1,12|t,1}$, whereas in the second week, $NFCI_{s,2} = \text{NFCI}_{s,1|t,2}$, and so on.

chosen to be in the seventh week of the first quarter of 1990. However, this only affects the first quarter of the out-of-sample evaluation.

4 Forecasting models

The aim is to nowcast Growth-at-Risk for quarter t in real-time. Therefore, GaR is re-estimated on a weekly frequency, taking into account the latest information available at the end of each week. The nowcasting exercise starts in the first fully observed week of the quarter t and ends in the third week of the subsequent quarter $t+1$. Thus, for each quarter t through the evaluation period, there are 15 weeks considered. The release of the first estimate of GDP is usually in the fourth week of quarter $t+1$. Therefore, the real-time data and the associated nowcasts are updated up to the latest releases and revisions just before the Bureau of Economic Analysis (BEA) publishes the first estimate of GDP for quarter t. This approach to timing is identical to the one of Carriero et al. (2022), ensuring that the forecast horizons are comparable across quarters.

4.1 GARCH-type models

For GaR prediction, Brownlees and Souza (2021) find that the GARCH(1,1) model and quantile regressions exhibit similar out-of-sample performance. Therefore, I consider AR-GARCH-type models to nowcast quantiles of the GDP growth distribution. The conditional mean and variance of $y_{t,w}$ are modeled as

$$\mu_{t|t,w} = \begin{cases} \phi_{0,w} + \phi_{1,w} y_{t-2,w} + x^{\mu}_{t,w}\delta_w & \text{if } w \leq 3 \\ \phi_{0,w} + \phi_{1,w} y_{t-1,w} + x^{\mu}_{t,w}\delta_w & \text{if } w > 3 \end{cases} \tag{4}$$

$$\sigma^2_{t|t,w} = \omega_w + \alpha_w \varepsilon^2_{t-1,w} + \beta_w \sigma^2_{t-1|t-1,w} + \exp(x^{\sigma}_{t,w}\gamma_w), \tag{5}$$

where $\varepsilon_{t,w} = y_{t,w} - \mu_{t|t,w}$, $\omega_w > 0$, $\alpha_w > 0$, $\beta_w > 0$, and $\alpha_w + \beta_w < 1$. Thus, all GARCH-type specifications considered are of the form AR(1)-$x^{\mu}_{t,w}$-GARCH(1,1)-$x^{\sigma}_{t,w}$, where $x^{\mu}_{t,w}$ and $x^{\sigma}_{t,w}$ represent vectors of additional explanatory variables for the conditional mean and the conditional variance, respectively. The elements in $x^{\sigma}_{t,w}$ are standardized to have mean zero and variance one. Assuming normality, the conditional distribution of GDP growth in week w is

$$y_{t|t,w} \sim \mathcal{N}(\mu_{t|t,w}, \sigma^2_{t|t,w}). \tag{6}$$

Equation (4) resembles a linear regression of $y_{t,w}$ on a constant, one lag of GDP growth, and additional predictors in $x^{\mu}_{t,w}$. Similarly, equation (5) corresponds to a GARCH(1,1) specification, allowing for potential predictors in $x^{\sigma}_{t,w}$. To ensure a strictly positive conditional variance, explanatory variables for $\sigma^2_{t|t,w}$ are incorporated through the exponential function, $\exp(\cdot)$.

In the baseline specification (henceforth AR-GARCH), both $x^{\mu}_{t,w}$ and $x^{\sigma}_{t,w}$ are omitted, resulting in a purely univariate time series model with an AR(1) specification for the conditional mean. When incorporating real-time median SPF projections in equation (4), i.e., $x^{\mu}_{t,w} = SPF_{t,w}$, the model is denoted as SPF-GARCH.[3] Regarding the NFCI, it is uncertain whether including it in the conditional mean or variance would refine GaR nowcasts. Thus, the out-of-sample predictive efficacy of the NFCI is evaluated both in the conditional mean and variance.

The GARCH model parameters in equations (4) and (5) are estimated by quasi-maximum likelihood. Mean predictions, $\hat{\mu}_{t|t,w}$, are readily available from estimates on the parameters in equation (4). Conditional quantiles and density forecasts can be obtained from estimates on the conditional volatility, $\hat{\sigma}^2_{t|t,w}$, in equation (5) and assuming normality in equation (6). In week w of quarter t, the nowcast of the conditional τ-quantile, where $\tau \in (0,1)$, is then given by

$$\hat{Q}^{\text{GARCH}}_{t|t,w}(\tau) = \hat{\mu}_{t|t,w} + \sqrt{\hat{\sigma}^2_{t|t,w}} F^{-1}(\tau), \tag{7}$$

where $F^{-1}(\cdot)$ is the inverse cumulative distribution function of $\mathcal{N}(0,1)$. Due to the small sample sizes in the out-of-sample analysis, assuming a standard normal distribution for the innovations aims to achieve a parsimonious parameterization. A more flexible distribution that features skewness and fatness is studied in the robustness section, revealing that the normality assumption is justified from a forecasting perspective.[4]

4.2 Quantile regressions

As popularized by Adrian et al. (2019), among others, in the GaR literature, quantile regressions of Koenker and Bassett (1978) are a commonly used method to predict conditional quantiles of the GDP growth distribution. At time T, for the conditional τ-quantile, the vector of quantile regression parameters β^{τ}_{w} is estimated by minimizing the asymmetric absolute loss function

$$\hat{\beta}^{\tau}_{w} = \underset{\beta^{\tau}_{w} \in \mathbb{R}^k}{\operatorname{argmin}} \sum_{t=1}^{T-j} (y_{t,w} - x_{t,w}\beta^{\tau}_{w})(\tau - \mathbb{1}_{(y_{t,w} - x_{t,w}\beta^{\tau}_{w} < 0)}), \tag{8}$$

where $y_{t,w}$ represents GDP growth rates available in week w of quarter T, $x_{t,w}$ is a vector of predictors, $\mathbb{1}_{(\cdot)}$ is the indicator function, and j depends on the release date of previous

[3] Choosing to include the median instead of the mean SPF projections is arbitrary but more robust against extreme individual forecasts. However, in the robustness section 6.5.1, employing the mean SPF forecasts yields nearly identical out-of-sample results for GaR and density forecasts.

[4] Instead of assuming normality, employing the empirical quantiles of the estimated standardized GARCH residuals results in almost identical predictive performance.

quarters' GDP, taking values in $\{1, 2\}$. Thus, in weeks $w = 1, \ldots, 3$, the sum runs from $t = 1$ to $T - 2$, and in weeks $w = 4, \ldots, 15$, it runs from $t = 1$ to $T - 1$. The index w emphasizes that the quantile regressions are re-estimated each week. Then, the predicted τ-quantile of GDP growth in quarter t conditional on data up to week w reads

$$\hat{Q}^{\mathrm{QR}}_{t|t,w}(\tau) \equiv \hat{Q}^{\mathrm{QR}}_{y_{t,w}}(\tau|x_{t,w}) = x_{t,w}\hat{\beta}^{\tau}_{w}. \tag{9}$$

Thus, for each quarter t through the evaluation period, there are 15 nowcasts $\hat{Q}^{\tau}_{t|t,w}$, for $w = 1, \ldots, 15$.

In the baseline specification, the autoregressive quantile regression (QR-AR) only includes a constant and one lag of $y_{t,w}$, i.e., $x_{t,w} = (1, y_{t-j,w})$. If $x_{t,w}$ also includes the $NFCI_{t,w}$ ($SPF_{t,w}$), this specification is referred to as QR-NFCI (QR-SPF).

5 Forecast evaluation

I define the "true" outcomes of GDP growth following the approach of Carriero et al. (2022) and others. This involves using GDP growth rates available one quarter after the initial release of GDP. Thus, actual GDP growth is defined as annualized quarterly log growth rates according to

$$y_t = 400 \cdot (\log(\mathrm{GDP}_{t|t+2,1}) - \log(\mathrm{GDP}_{t-1|t+2,1})), \tag{10}$$

where $\mathrm{GDP}_{t|t+2,1}$ denotes the latest revised version of GDP for quarter t available in the first week of quarter $t + 2$, typically corresponding to the third monthly release, $\mathrm{GDP}^{(3)}_t$. Therefore, identical to Carriero et al. (2022), the actual GDP growth rates are given by the second releases in the quarterly real-time vintages available at the FED database.

For evaluating conditional quantile nowcasts, the quantile score (QS) is a strictly consistent scoring rule (Gneiting, 2011). The asymmetric piecewise linear loss function for the τ-quantile in week w of quarter t is computed as

$$\mathrm{QS}^{\tau}_{t,w} = (\mathbb{1}_{(\hat{Q}_{t|t,w}(\tau) \geq y_t)} - \tau)(\hat{Q}_{t|t,w}(\tau) - y_t), \tag{11}$$

where y_t denotes actual GDP. To evaluate predictive distributional forecasts, the Continuous Ranked Probability Score (CRPS) is strictly consistent and is approximated by

$$\mathrm{CRPS}_{t,w} = \frac{2}{J-1} \sum_{j=1}^{J-1} \hat{Q}_{t|t,w}(\tau_j), \tag{12}$$

where $\tau_j = j/J$ and $J = 20$ (Gneiting and Raftery, 2007).

Statistical inference on the out-of-sample predictive performance requires comparing many models against each other. Instead of testing each competitor model against a benchmark model, say QR-AR, I report the model confidence set (MCS) of Hansen et al. (2011) based on the Diebold and Mariano (1995) and West (1996) test to account for the problem of multiple comparisons. This approach has the advantage of jointly testing the null hypothesis of equal predictive performance of the considered models. Let $\mathcal{M}_0$ denote the set of all competing models indexed by $k = 1, \dots, K$. Then, the relative performance of two models, k and l, is measured by the loss differential, $d^w_{kl,t} \equiv L^k_{t,w} - L^l_{t,w}$, at week w through quarters t, for all $k, l \in \mathcal{M}_0$. Here, the loss functions considered are either $\text{QS}^\tau_{t,w}$ or $\text{CRPS}_{t,w}$. The null hypothesis is $H_0 : \mathbb{E}(d^w_{kl,t}) = 0$ for all $k, l \in \mathcal{M}$, where $\mathcal{M} \subset \mathcal{M}_0$. The $1 - \alpha\%$ MCS is defined as $\mathcal{M}^\star \equiv \{k \in \mathcal{M}_0 : \mathbb{E}(d^w_{kl,t}) \leq 0 \text{ for all } l \in \mathcal{M}_0\}$. The test statistic, $T_\mathcal{M} = \max_{k,l \in \mathcal{M}} |t^w_{kl}|$, is based on the average loss difference

$$t^w_{kl} = \frac{\overline{d}^w_{kl}}{se(\overline{d}^w_{kl})} \text{ for all } k, l \in \mathcal{M}, \tag{13}$$

where $\overline{d}^w_{kl}$ and $se(\overline{d}^w_{kl})$ are the mean and the standard error of $\overline{d}^w_{kl,t}$. Thus, the test sequentially eliminates the worst-performing model from $\mathcal{M}_0$ if $T_\mathcal{M}$ exceeds the critical value, indicating a rejection of equal performance. Since the asymptotic distribution of $T_\mathcal{M}$ is nonstandard, it is simulated by block-bootstrapping, where I use 15.000 replications. As barely any autocorrelations of loss differentials are significant after the second lag, the block length chosen is 3. However, the results are not sensitive to these settings.[5]

One critical aspect is the limiting distribution of $T_\mathcal{M}$ if models in $\mathcal{M}_0$ are nested. Specifically, for this distribution to exist, $d^w_{kl,t}$ must be stationary with strictly positive variance. Thus, a common approach is to estimate the models on a rolling window, ensuring a well-behaved asymptotic distribution with the variance of $d^w_{kl,t}$ bounded away from zero (Giacomini and White, 2006). However, the use of quarterly data results in considerable estimation uncertainty due to small sample sizes. Consequently, I employ expanding window estimation to reduce estimation uncertainty and thus increase power of the MCS. Additionally, robustness checks using rolling window estimation show that the results remain qualitatively unchanged.

In determining the evaluation period, a trade-off exists between selecting a later starting point for higher estimation accuracy and choosing an earlier starting point to increase the test procedure's power. Given the SPF's inception in 1968, the evaluation period spans from the first quarter of 1990 to the fourth quarter of 2019, resulting in 85 in-sample observations at the beginning of the evaluation period. Excluding the onset of the

[5]The MCS procedure is implemented in the MFE Toolbox of Sheppard (2009), available at `https://www.kevinsheppard.com/code/matlab/mfe-toolbox/`.

COVID-19 pandemic yields 120 out-of-sample observations to assess the models' prediction accuracy. Additionally, to evaluate the predictive performance during the pandemic, the analysis extends to the subsequent two years. For the evaluation, I report the 90% MCS, conducting tests of equal predictive performance at the 10% significance level.

6 Empirical results

This section presents the empirical results on the accuracy of the real-time nowcasts of Growth-at-Risk. In the following, I exclude the COVID-19 pandemic to ensure that one extreme outlier is not driving the results. Later, I will extend the sample and moreover study two periods of economic turmoil separately. Excluding the pandemic, the evaluation starts in the first quarter of 1990 and ends in the fourth quarter of 2019, resulting in 120 out-of-sample observations.

Before turning to quantile- and density forecasts, I provide motivating evidence that supports exploiting the information content provided by the SPF median forecasters and updating data in real-time. Regarding the SPF forecasting performance, Table 2 shows summary statistics of the forecast errors of the SPF median projections. The average one-step ahead forecast error of the SPF is close zero, whereas the nowcasts seem to be too optimistic on average. However, for both forecast horizons, the forecast errors are not significantly different from zero, suggesting that unbiasedness of the SPF median forecasts cannot be rejected. In addition, the root mean squared forecast error (RMSFE) of the nowcasts is significantly smaller than of the one-step ahead forecasts (p-value = 0.06). Thus, the forecast accuracy of the SPF increases for shorter horizons.

The lower panel reports estimates for the Mincer and Zarnowitz (1969) (MZ) regression, i.e., a regression of actual GDP growth on the SPF projections, resulting in coefficient estimates on the constant, ρ_0, and the SPF median forecasts, ρ_1. Jointly testing $\rho_0 = 0$ and $\rho_1 = 1$ reveals that unbiasedness and efficiency of the SPF median forecasts cannot be rejected. In particular, for the nowcasts, the constant is close to zero and the slope coefficient is close to unity. Moreover, the R^2 increases from 0.28 to 0.49. Thus, the SPF consensus point predictions have desirable properties, especially for short horizons.

Next, to motivate utilizing the real-time information flow as GDP and SPF data are released, Table 3 presents the RMSFE for different forecast origins within a quarter, starting from the first fully observed week. Specifically, I consider nowcasts of the conditional mean of GDP growth based on different AR-GARCH type specifications, with the univariate AR-GARCH model solely relying on past releases of GDP growth. Gray areas in the table indicate the 90% MCS.

First, the RMSFE of the purely autoregressive AR-GARCH model tends to become

Table 2: SPF unbiasedness and efficiency

SPF error	One-step ahead		Nowcasts	
mean	-0.035	(0.200)	0.220	(0.144)
RMSFE	1.850		1.557	
MZ regression	**One-step ahead**		**Nowcasts**	
ρ_0	-0.658	(0.816)	0.027	(0.340)
ρ_1	1.252	(0.292)	1.087	(0.132)
R^2	0.280		0.492	
p-value	0.680		0.220	

Notes: The upper panel reports the mean forecast error and root mean squared forecast error (RMSFE) of the SPF. The lower panel shows estimates of the Mincer-Zarnowiz (MZ) regressions. Newey-West standard errors are reported in parentheses. The reported p-values concern the joint null hypothesis $H_0 : \rho_0 = 0, \rho_1 = 1$. The evaluation period is 1990:Q1 to 2019:Q4, consisting of 120 observations.

Table 3: Out-of-sample evaluation of mean forecasts

	week							
RMSFE	1	3	5	7	9	11	13	15
AR-GARCH	2.21	2.19	2.02	2.02	1.99	1.99	2.01	2.01
$+x^{\mu}_{t,w} = SPF_{t,w}$	**1.96**	**1.95**	**1.89**	**1.65**	**1.60**	**1.60**	**1.60**	**1.60**
$+x^{\mu}_{t,w} = NFCI_{t,w}$	2.18	2.20	1.95	1.99	1.98	1.99	1.95	2.01

Notes: This table reports the root mean squared forecast errors (RMSFE) over the out-of-sample evaluation period. The columns indicate the week of a quarter at which the nowcast is formed. Gray areas represent the 90% model confidence set based on the mean squared error loss and bold letters are the lowest RMSFE within each column. All models are estimated on an expanding window. The evaluation period is 1990:Q1 to 2019:Q4, consisting of 120 observations.

smaller with decreasing forecast horizon, demonstrating that incorporating the latest releases and revisions of GDP data leads to improvements in nowcasting GDP growth. The reduction in the average forecast errors is most prominent between the third and fifth weeks, associated with the publication of the first estimate of the previous quarter's GDP. Later revisions published between weeks 7 and 9 and between weeks 11 and 13 appear less informative.

Second, adding the SPF to the conditional mean of the AR-GARCH mostly significantly reduces out-of-sample forecast errors. In the first four weeks, the SPF-GARCH model exploits the one-step ahead median SPF forecast. Apart from a few exceptions, the release of the SPF nowcast occurs between the fifth and seventh weeks. The results reveal that the publication of the SPF median nowcast strongly reduces the RMSFE. Thus, accounting for the latest releases of both GDP and SPF reduces forecast errors. Interestingly, the

AR-GARCH model, including the SPF as an additional regressor, results in larger losses than the SPF median projections shown in Table 2. The increased estimation uncertainty of the GARCH model translates into slightly less accurate forecasts of GDP growth, however, these differences are not significant.

Lastly, adding the NFCI into the conditional mean does not improve forecast accuracy, indicating that financial conditions are not informative about the center of the conditional GDP growth distribution.

6.1 SPF versus NFCI

To predict Growth-at-Risk as well as the conditional density of GDP growth, the models considered are GARCH and quantile regressions. In order to evaluate downside and upside risk, I report the average 10% and 90% quantile scores. The predictive density is assessed by the CRPS. Table 4 reports average losses of the quantile scores and the CRPS, along with the 90% MCS for different forecast origins within a quarter, starting in the first fully observed week. The upper panel shows results for 10% quantile nowcasts of GDP growth. Both univariate specifications, AR-GARCH and QR-AR, exhibit increasing nowcast accuracy as GDP releases and revisions become available during the quarter. Notably, there are two distinct declines in average losses observed between weeks 3 to 5 and between weeks 7 to 9. The former coincides with the release of the first estimate of GDP for the previous quarter, while the latter is attributable to its first revision. Moreover, the AR-GARCH model performs better than the QR-AR approach, in line with Brownlees and Souza (2021).

Incorporating the NFCI into quantile regressions underscores its substantial predictive power for downside risk. The results demonstrate that quantile regressions utilizing the NFCI as an explanatory variable perform best in the first weeks. While this finding is not new - avoiding look-ahead bias - it confirms previous research (see, for example, Giglio et al., 2016; Adrian et al., 2019; Ferrara et al., 2022).

In quantile regressions, the NFCI serves directly as a predictor for the targeted quantile. In the GARCH model, assuming normality, the NFCI predicts downside risk by explaining changes in the conditional mean or variance. Incorporating the NFCI into either the conditional mean, $x^{\mu}_{t,w}$, or the conditional variance, $x^{\sigma}_{t,w}$, reduces the nowcast errors compared to the basic AR-GARCH model, contrasting the findings by Brownlees and Souza (2021). However, at the beginning of the quarter, the AR-GARCH with NFCI in the conditional mean or variance yields higher quantile losses than the QR-NFCI model. Towards the end of the quarter, the GARCH models incorporating NFCI, whether in the conditional mean or volatility, demonstrate a predictive performance similar to that of the QR-NFCI model.

Table 4: Out-of-sample evaluation using SPF or NFCI

	week							
10% quantile score	1	3	5	7	9	11	13	15
AR-GARCH	0.44	0.43	0.40	0.40	0.37	0.37	0.37	0.37
$+x^{\mu}_{t,w} = NFCI_{t,w}$	0.40	0.38	0.36	0.36	0.35	0.35	0.34	0.34
$+x^{\sigma}_{t,w} = NFCI_{t,w}$	0.41	0.39	0.37	0.36	0.34	0.34	0.34	0.33
$+x^{\mu}_{t,w} = SPF_{t,w}$	0.36	0.35	0.35	**0.28**	**0.29**	**0.29**	**0.29**	**0.29**
QR-AR	0.50	0.51	0.49	0.49	0.47	0.47	0.47	0.47
$+NFCI_{t,w}$	**0.31**	**0.32**	**0.32**	0.32	0.30	0.33	0.32	0.33
$+SPF_{t,w}$	0.44	0.44	0.43	0.35	0.31	0.31	0.31	0.31
90% quantile score	1	3	5	7	9	11	13	15
AR-GARCH	0.40	0.40	0.35	0.35	0.36	0.36	0.36	0.36
$+x^{\mu}_{t,w} = NFCI_{t,w}$	0.41	0.42	0.38	0.38	0.37	0.39	0.37	0.37
$+x^{\sigma}_{t,w} = NFCI_{t,w}$	0.42	0.41	0.38	0.38	0.39	0.39	0.39	0.38
$+x^{\mu}_{t,w} = SPF_{t,w}$	**0.34**	**0.34**	**0.34**	**0.31**	**0.31**	**0.31**	**0.31**	**0.31**
QR-AR	0.55	0.54	0.46	0.46	0.46	0.46	0.47	0.46
$+NFCI_{t,w}$	0.46	0.45	0.39	0.37	0.37	0.38	0.39	0.40
$+SPF_{t,w}$	0.39	0.39	0.38	0.36	0.33	0.33	0.33	0.33
CRPS	1	3	5	7	9	11	13	15
AR-GARCH	1.25	1.24	1.14	1.14	1.13	1.13	1.14	1.14
$+x^{\mu}_{t,w} = NFCI_{t,w}$	1.27	1.28	1.15	1.17	1.16	1.17	1.15	1.17
$+x^{\sigma}_{t,w} = NFCI_{t,w}$	1.28	1.25	1.17	1.17	1.16	1.15	1.16	1.15
$+x^{\mu}_{t,w} = SPF_{t,w}$	**1.12**	**1.12**	**1.09**	**0.99**	**0.96**	**0.96**	**0.96**	**0.96**
QR-AR	1.35	1.34	1.25	1.25	1.24	1.24	1.25	1.25
$+NFCI_{t,w}$	1.22	1.24	1.17	1.17	1.18	1.18	1.18	1.19
$+SPF_{t,w}$	1.17	1.17	1.14	1.03	1.00	1.00	1.00	1.00

Notes: This table reports the average losses of the 10% and the 90% quantile scores in the top and the middle panel. The panel at the bottom shows the average continuously ranked probability score (CRPS) over the out-of-sample evaluation period. The columns indicate the week of a quarter at which the nowcast is formed. Gray areas represent the 90% model confidence set and bold letters are the lowest average losses within each column. All models are estimated on an expanding window. The evaluation period is 1990:Q1 to 2019:Q4, consisting of 120 observations.

The most striking result of Table 4 is the inclusion of the SPF-GARCH model in the MCS. In the first five weeks, the SPF-GARCH model utilizes the one-step ahead median SPF forecasts in the conditional mean. Throughout this period, the QR-NFCI exhibits lower quantile losses. As the SPF nowcasts become available between weeks 5 and 7, in the second half of the quarter, the SPF-GARCH significantly outperforms its competitors, except for the QR-NFCI approach in weeks 7 to 9. The SPF median nowcast in a GARCH framework outperforms the NFCI in both GARCH and QR specifications in predicting GaR. This finding holds only for the GARCH specification. The MCS does not include

the QR-SPF model, although, in the last weeks of the nowcasting exercise, the QR-SPF model exhibits similar average quantile losses. This suggests that quantile regressions do not exploit the information flow stemming from the SPF median forecasts for predicting downside risk as effectively as the GARCH model does.

Examining the 90% quantile losses, the middle panel of Table 4 assesses the predictive performance concerning upside risk. The SPF notably enhances the 90% quantile nowcasts for both the GARCH model and quantile regressions. However, employing median SPF forecasts in the GARCH model is statistically superior compared to quantile regressions in forecasting the upper quantiles of the GDP growth distribution. While the NFCI also decreases the out-of-sample losses in quantile regressions, its effect is less pronounced compared to the SPF. Interestingly, the QR-NFCI model performs worse than the purely backward-looking AR-GARCH model. Moreover, when explaining changes in the mean and variance in the GARCH model, the NFCI appears uninformative, contributing mainly to estimation uncertainty in the baseline AR-GARCH model. This finding suggests that although the NFCI serves as a strong predictor for downside risk, it is not capable of explaining variation in the upper part of the GDP growth distribution, in line with Giglio et al. (2016).

Finally, the bottom panel of Table 4 examines the CRPS, evaluating the predictive performance across the entire distribution. Given the SPF's enhancement of both lower and upper quantile predictions in the GARCH model, it's unsurprising that the SPF-GARCH model significantly outperforms other models in predicting the conditional density of GDP growth. Moreover, the SPF median forecasts improve the predictive performance of quantile regressions. It is important to note that the median SPF projections focus on the center of the distribution. Consequently, using the SPF nowcasts directly as an explanatory variable for the conditional mean calibrates the entire distributional forecast through mean changes. Conversely, in quantile regressions, the weight of the SPF median nowcast is determined by different quantile regression coefficients for each quantile separately. As the SPF-GARCH model yields more precise predictive densities, the former approach proves to be more effective in this context. Moreover, the AR-GARCH results in a sharper predictive density than the QR-NFCI approach, underscoring that the NFCI's predictive ability holds exclusively for lower quantiles of GDP growth.

In summary, the findings highlight that the median SPF nowcasts are informative across all quantiles of GDP growth, with the SPF-GARCH model producing the most precise predictive densities. In the SPF-GARCH model, the conditional mean utilizes the SPF consensus nowcasts to capture variation in the central tendency of GDP growth, while the conditional volatility of the GARCH model helps generate a conditional density around the predicted mean of GDP growth. However, GARCH effects are strong, as indicated by

β in the conditional variance in equation (5), which typically ranges between 0.8 and 0.9 throughout the out-of-sample period. Hence, incorporating the SPF median projections enhances predicting the distribution of GDP growth by providing accurate forecasts of the conditional mean, thus reducing past forecast errors, which alters the dynamics in the conditional volatility of the GARCH model. The GARCH model imposes more structure, while the SPF captures structural changes, which turns out advantageous, particularly with small sample sizes. Additionally, while not in the MCS, quantile regressions incorporating the SPF provide the second-best density nowcasts, demonstrating superior predictive performance in comparison to the QR-NFCI approach.

Figure 2 illustrates the real-time predictive distribution of the SPF-GARCH model in panel (a) and the QR-NFCI approach in panel (b). Both models produce conditional 10% quantile nowcasts that closely track the patterns of quarterly GDP growth (black line). However, the QR-NFCI median nowcasts (blue line) appear to miss economic up- and downswings. In contrast, the SPF-GARCH model shows pronounced dynamics in the conditional mean of its predictive distribution. In addition, the upper quantiles of the QR-NFCI are relatively constant over time, except for the aftermath of the global financial crisis around 2009, in line with Adrian et al. (2019). Notably, most of the variation observed in panel (b) originates from GDP releases when updating the nowcasts within a quarter. The time-variation remains modest for a given week across quarters. In contrast, the predicted 90% quantiles of the SPF-GARCH model are less volatile within a quarter yet exhibit more variation across quarters, closely mirroring the GDP growth evolution. Lastly, as indicated by the significantly smaller CRPS, the density forecasts of the SPF-GARCH model are more closely centered around actual GDP growth than those of the QR-NFCI.

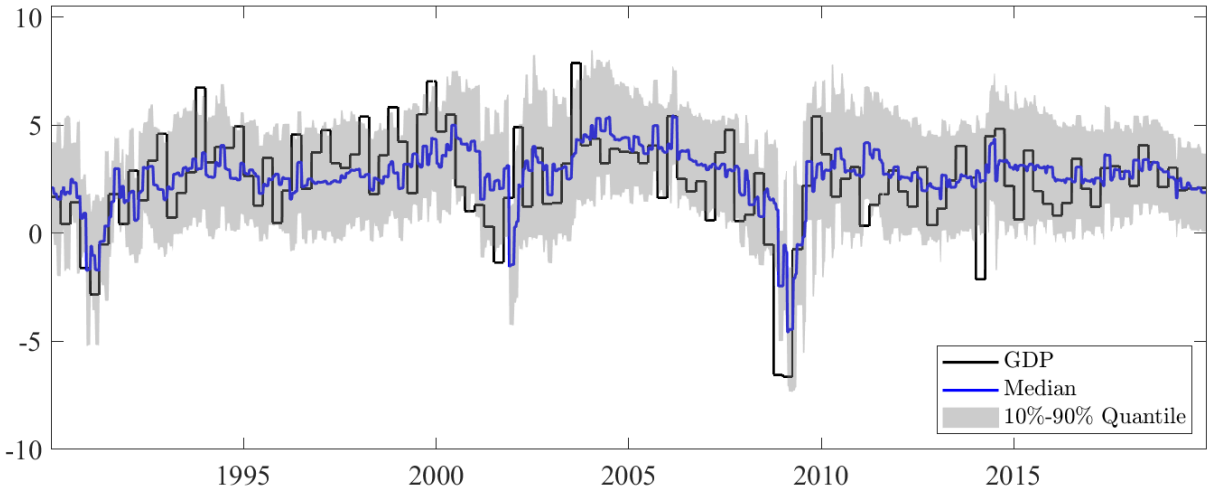
6.2 SPF and NFCI combined

In the previous section, the SPF and the NFCI are considered separately as predictors, demonstrating that the SPF is superior over the entire distribution. The NFCI is informative about lower quantiles but less for the other parts of the distribution. However, in the first weeks of a quarter, the QR-NFCI model achieves the lowest scores at the 10% quantile of GDP growth. Thus, in this section, to assess whether the NFCI adds additional predictive gains to the SPF, both indicators are considered jointly.

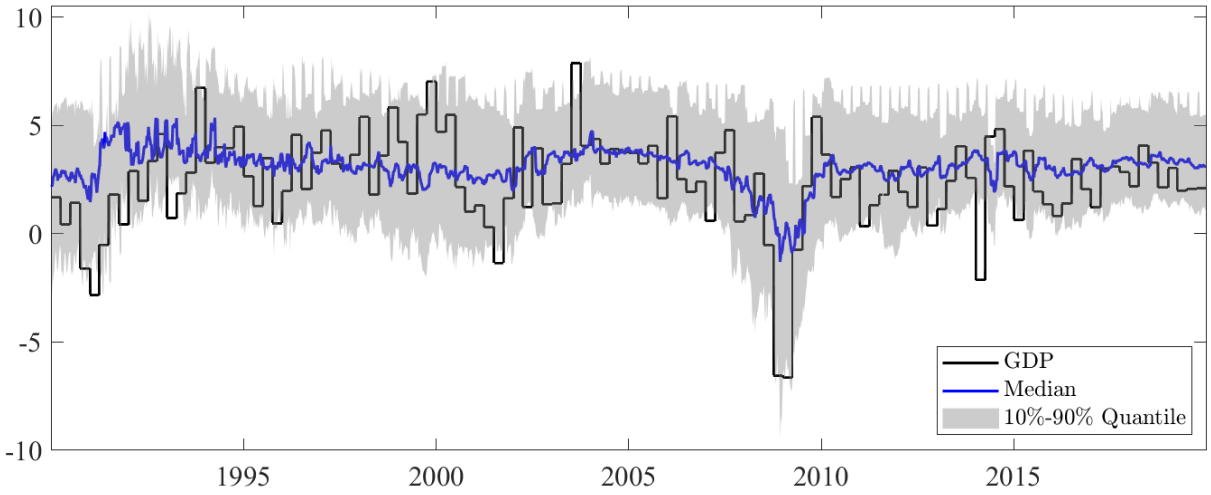
To enhance comparability, I include the SPF-GARCH and the QR-SPF model from Section 6.1 and add the NFCI to these specifications to test for additional predictive gains. In addition, since the quantile regression with NFCI produces the lowest losses at the 10% quantile in the first weeks of a quarter, the QR-NFCI model is also included in this analysis.

Figure 2: Real-time nowcasts

(a) SPF-GARCH



(b) QR-NFCI



Notes: This figure illustrates real-time nowcasts of quarterly US GDP growth over the out-of-sample evaluation period 1990:Q1 to 2019:Q4 for the SPF-GARCH model in panel (a) and the QR-NFCI approach in panel (b). The black line represents actual GDP growth, while the blue line shows the predicted median forecasts. In panel (a), where the SPF-GARCH model assumes a normal distribution, the median corresponds to the mean. The grey area indicates density forecasts ranging from the 10% to the 90% conditional quantiles.

Table 5 presents the out-of-sample results of this comparison. In the upper panel, for the 10% quantile of GDP growth, the QR-NFCI model performs best in the first weeks, while the SPF-GARCH model with NFCI in the conditional mean exhibits the lowest scores towards the end of the quarter. However, all models except the QR-SPF approach are contained in the MCS, indicating that the SPF-GARCH model is not significantly outperformed.

Table 5: Out-of-sample evaluation using NFCI and SPF combined

	week							
10% quantile score	1	3	5	7	9	11	13	15
SPF-GARCH	0.36	0.35	0.35	0.28	0.29	0.29	0.29	0.29
$+x^{\mu}_{t,w} = NFCI_{t,w}$	0.37	0.35	0.34	0.28	0.27	**0.27**	**0.27**	**0.27**
$+x^{\sigma}_{t,w} = NFCI_{t,w}$	0.35	0.34	0.33	0.28	0.28	0.28	0.28	0.28
QR-SPF	0.44	0.44	0.43	0.35	0.31	0.31	0.31	0.31
$+NFCI_{t,w}$	**0.29**	**0.30**	**0.29**	**0.28**	**0.27**	0.27	0.28	0.28
QR-NFCI	0.31	0.32	0.32	0.32	0.30	0.33	0.32	0.33
90% quantile score	1	3	5	7	9	11	13	15
SPF-GARCH	**0.34**	**0.34**	**0.34**	**0.31**	**0.31**	**0.31**	**0.31**	**0.31**
$+x^{\mu}_{t,w} = NFCI_{t,w}$	0.37	0.36	0.34	0.32	0.31	0.31	0.31	0.31
$+x^{\sigma}_{t,w} = NFCI_{t,w}$	0.37	0.36	0.36	0.34	0.32	0.32	0.32	0.31
QR-SPF	0.39	0.39	0.38	0.36	0.33	0.33	0.33	0.33
$+NFCI_{t,w}$	0.41	0.39	0.36	0.35	0.34	0.34	0.34	0.34
QR-NFCI	0.46	0.45	0.39	0.37	0.37	0.38	0.39	0.40
CRPS	1	3	5	7	9	11	13	15
SPF-GARCH	**1.12**	**1.12**	**1.09**	**0.99**	**0.96**	**0.96**	**0.96**	**0.96**
$+x^{\mu}_{t,w} = NFCI_{t,w}$	1.18	1.16	1.09	1.01	0.97	0.97	0.97	0.98
$+x^{\sigma}_{t,w} = NFCI_{t,w}$	1.14	1.14	1.10	1.01	0.98	0.98	0.98	0.98
QR-SPF	1.17	1.17	1.14	1.03	1.00	1.00	1.00	1.00
$+NFCI_{t,w}$	1.17	1.17	1.12	1.02	0.98	0.99	0.98	0.99
QR-NFCI	1.22	1.24	1.17	1.17	1.18	1.18	1.18	1.19

Notes: This table reports the average losses of the 10% and the 90% quantile scores in the top and the middle panel. The panel at the bottom shows the average continuously ranked probability score (CRPS) over the out-of-sample evaluation period. The columns indicate the week of a quarter at which the nowcast is formed. Gray areas represent the 90% model confidence set and bold letters are the lowest average losses within each column. All models are estimated on an expanding window. The evaluation period is 1990:Q1 to 2019:Q4, consisting of 120 observations.

In terms of the 90% quantile, the SPF-GARCH model consistently provides the lowest out-of-sample losses, significantly outperforming the QR-NFCI approach. Moreover, adding the NFCI to the SPF-GARCH model does not result in predictive gains. Similarly, indicated by the CRPS, the SPF-GARCH model produces significantly more precise density forecasts than the QR-NFCI approach. Overall, across the entire distribution, the SPF-GARCH model is among the best performing models, with the addition of the NFCI showing no significant improvement in predictive accuracy.

6.3 Growth-at-Risk during economic turmoils

Up to this point, I excluded the COVID-19 pandemic from the analysis due to its sudden and extreme impact on GDP, which could potentially distort the main findings of the paper. In this section, I extend the evaluation period up to 2022:Q4 to test the robustness of the conclusions drawn in Section 6.1. Given the exceptional nature of the COVID-19 pandemic, both the purely backward-looking AR-GARCH model and the QR-AR model exhibit large forecast errors at the pandemic's onset, leading to a loss in the power of the MCS procedure. As the primary focus remains on the median SPF projections, I exclude the purely backward-looking models and focus on the SPF-GARCH model, as well as the QR-SPF and QR-NFCI approaches.

However, extending the evaluation period masks some insights about the predictive performance of the SPF and the NFCI during economic turmoils. Therefore, I additionally examine two major economic crises separately: the global financial crisis (GFC) in 2007-2008 and the COVID-19 pandemic in 2020-2022.

6.3.1 Including the COVID-19 pandemic

Table 6 presents the out-of-sample results when including the COVID pandemic in the evaluation sample. The main results established earlier remain consistent. In the first weeks, the NFCI is more informative about the 10% quantiles than the SPF one-step ahead forecast. However, once the SPF nowcast becomes available, the SPF-GARCH model is the significantly best-performing approach. Similarly, for the 90% quantile and the predictive density of GDP growth, apart from a few exceptions, the SPF-GARCH model achieves the lowest losses. Again, the GARCH model utilizes the SPF median nowcasts more effectively than quantile regressions.

6.3.2 The global financial crisis

Focusing on Growth-at-Risk during the GFC, the upper panel of Table 7 displays the real-time average 10% quantile losses from 2007:Q4 to 2009:Q3. The table does not contain information about the MCS due to the low statistical power resulting from the small number of observations (eight quarters).

At the beginning of the quarter, the NFCI provides the most accurate GaR nowcasts, suggesting that the NFCI is a strong predictor of economic downside risk during the financial crisis.[6] For instance, in the third week of a quarter, the 10% quantile loss is less than one-third of the loss associated with the SPF-GARCH model. However, in the second

[6]During the GFC, the average quantile losses of the QR-NFCI fluctuate notably throughout quarters, likely reflecting the increased volatility of the real-time NFCI during this period.

Table 6: Out-of-sample evaluation including the COVID pandemic

	week							
10% quantile score	1	3	5	7	9	11	13	15
SPF-GARCH	0.74	0.73	0.73	**0.33**	**0.34**	**0.34**	**0.34**	**0.34**
QR-NFCI	**0.65**	**0.67**	**0.67**	0.67	0.67	0.68	0.67	0.67
QR-SPF	0.76	0.76	0.74	0.39	0.37	0.37	0.37	0.37
90% quantile score	1	3	5	7	9	11	13	15
SPF-GARCH	0.55	0.55	**0.44**	**0.32**	**0.32**	**0.32**	**0.32**	**0.32**
QR-NFCI	0.64	0.65	0.69	0.68	0.67	0.67	0.67	0.69
QR-SPF	**0.54**	**0.54**	0.59	0.38	0.36	0.36	0.36	0.36
CRPS	1	3	5	7	9	11	13	15
SPF-GARCH	1.74	1.74	**1.66**	**1.08**	**1.06**	**1.07**	**1.07**	**1.07**
QR-NFCI	1.82	1.84	1.84	1.84	1.84	1.84	1.84	1.84
QR-SPF	**1.70**	**1.70**	1.69	1.15	1.10	1.11	1.10	1.10

Notes: This table reports the average losses of the 10% and the 90% quantile scores in the top and the middle panel. The panel at the bottom shows the average continuously ranked probability score (CRPS) over the out-of-sample evaluation period. The columns indicate the week of a quarter at which the nowcast is formed. Gray areas represent the 90% model confidence set and bold letters are the lowest average losses within each column. All models are estimated on an expanding window. The evaluation period is 1990:Q1 to 2022:Q4, consisting of 132 observations.

half of the quarter, conditional on the SPF median nowcasts, the SPF-GARCH model and the QR-SPF model match the predictive performance of the QR-NFCI approach.

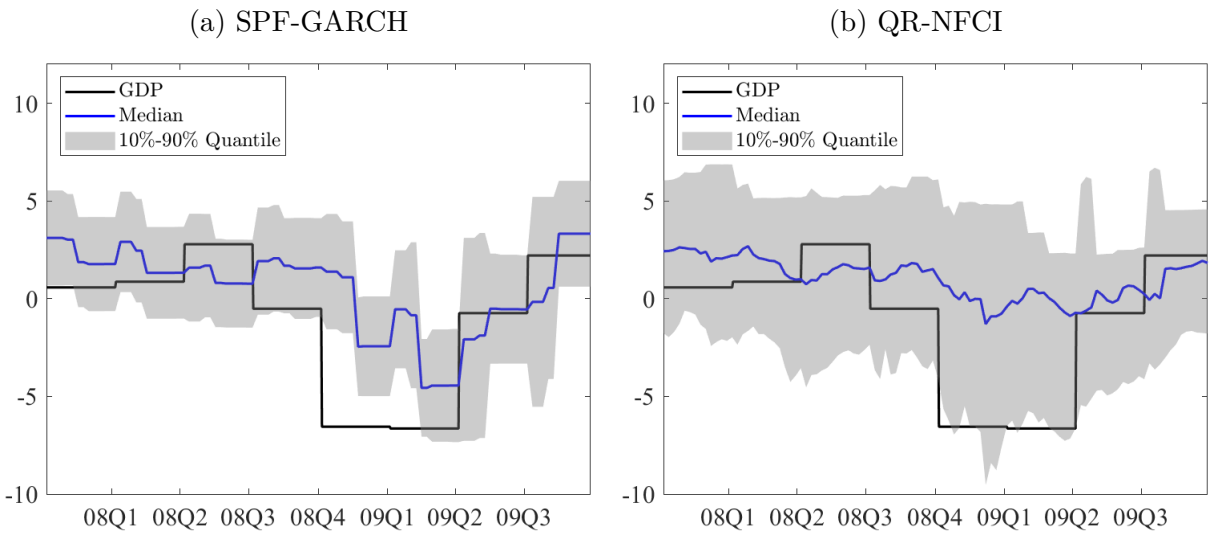
Table 7: Growth-at-Risk during economic turmoils

	week							
Financial Crisis	1	3	5	7	9	11	13	15
SPF-GARCH	1.16	1.16	1.01	0.33	0.34	0.34	0.33	0.33
QR-NFCI	**0.48**	**0.33**	**0.49**	0.32	**0.30**	0.34	**0.30**	0.45
QR-SPF	0.97	0.97	0.81	**0.30**	0.30	**0.30**	0.30	**0.30**
COVID	1	3	5	7	9	11	13	15
SPF-GARCH	4.53	4.53	4.53	**0.85**	**0.86**	**0.86**	**0.87**	**0.87**
QR-NFCI	4.04	4.08	4.14	4.22	4.29	4.23	4.10	4.07
QR-SPF	**3.88**	**3.88**	**3.81**	0.88	0.90	0.90	0.90	0.90

Notes: This table reports the average losses of the 10% quantile score during the global financial crisis and the COVID pandemic. The columns indicate the week of a quarter at which the nowcast is formed. Bold letters indicate the lowest average losses within in each column. All models are estimated on an expanding window. The evaluation period in the top panel is 2007:Q4 to 2009:Q3 i.e., 8 observations, and in the bottom panel is 2020:Q1 to 2022:Q4, i.e., 12 observations.

Figure 3 compares actual GDP against the median forecast of the SPF-GARCH in panel (a) and QR-NFCI in panel (b), respectively. The gray-shaded areas represent the predictive densities in terms of the 10% to 90% quantiles. Preceding the global financial crisis, quantile regressions with NFCI as conditioning information seem to provide overly pessimistic nowcasts of GaR. However, when GDP growth fell sharply in 2008:Q4, the conditional 10% quantiles of the QR-NFCI track the movements of GDP growth well. In contrast, even in the fourth quarter of 2008, the surveys panelists did not anticipate the severity of this event. Consequently, the SPF-GARCH model underestimates GaR in the fourth quarter of 2008. In the first quarter of 2009, the one-step ahead median SPF forecast again resulted in a conditional 10% quantile that was well above the actual growth rate. Eventually, the economic assessment of the median SPF nowcast released in the middle of 2009:Q1 drops to -5.2%, and the associated 10%-90% quantile range covers the actual GDP growth rate of -6.4%.

Figure 3: Real-time nowcasts through the global financial crisis



Notes: This figure depicts real-time nowcasts of quarterly US GDP growth during the global financial crisis from 2007:Q4 to 2009:Q3 for the SPF-GARCH model in panel (a) and the QR-NFCI approach in panel (b). The black line represents actual GDP growth, while the blue line shows the predicted median forecasts. In panel (a), where the SPF-GARCH model assumes a normal distribution, the median corresponds to the mean. The grey area indicates density forecasts ranging from the 10% to the 90% conditional quantiles.

In summary, at the onset of the GFC, the SPF-GARCH model suffers from less frequent updates, whereas the NFCI quickly reacts to new information. It takes the SPF median panelist until the second quarter of subsequently strongly negative growth rates to adjust accordingly. Nevertheless, starting from the middle of 2009:Q1, the SPF anticipates the magnitude and the timing of the financial crisis and the subsequent upswing well. This explains the relatively large 10% quantile losses of the SPF-GARCH in the first weeks of

the quarter, and the mechanism of matching to the performance of the QR-NFCI nowcasts toward the end of the quarter.

6.3.3 The COVID-19 pandemic in isolation

The nature of the shock to the economy represented by the COVID-19 pandemic is quite distinct from that of the global financial crisis. Thus, the bottom panel in Table 7 is concerned with the period 2020:Q1 to 2022:Q4. As becomes evident, the QR-NFCI nowcasts exhibit large 10% quantile losses. As panel (b) in Figure 4 shows, the QR-NFCI density nowcasts completely miss the down- and upswings of GDP growth during the COVID pandemic. Moreover, even extreme quantiles suffer from quantile crossing. Specifically, the median is not strictly between the estimated 10% and 90% quantile and during the fourth quarter of 2020, the 10% quantile lies above the 90% quantile, as indicated by the red-shaded area.[7]

In the first five weeks, based on the SPF median one-step ahead forecast, the SPF-GARCH model also exhibits considerable average 10% quantile losses. By week 7, with the release of the median nowcast, these losses decrease dramatically by over 80%. Panel (a) of Figure 7 illustrates how updating the SPF improves the GaR nowcasting accuracy. In the first quarter of 2020, at the onset of the lockdown measures, the SPF panelists did not adjust their assessment of the economic outlook. However, by mid-2020:Q2, the SPF released a median nowcast of -32.2% GDP growth, close to the actual rate of -32.9%, aligning the predictive density closely with the actual outcome. Moreover, in the following quarter, the SPF median projections anticipated the timing and magnitude of the subsequent upswing.

In conclusion, while the SPF consensus one-step ahead forecast in the second quarter of 2020 resulted in a large forecast error and wide prediction intervals in the subsequent quarters, the SPF consensus nowcasts improve the predictive density by shifting the center of the conditional distribution.[8]

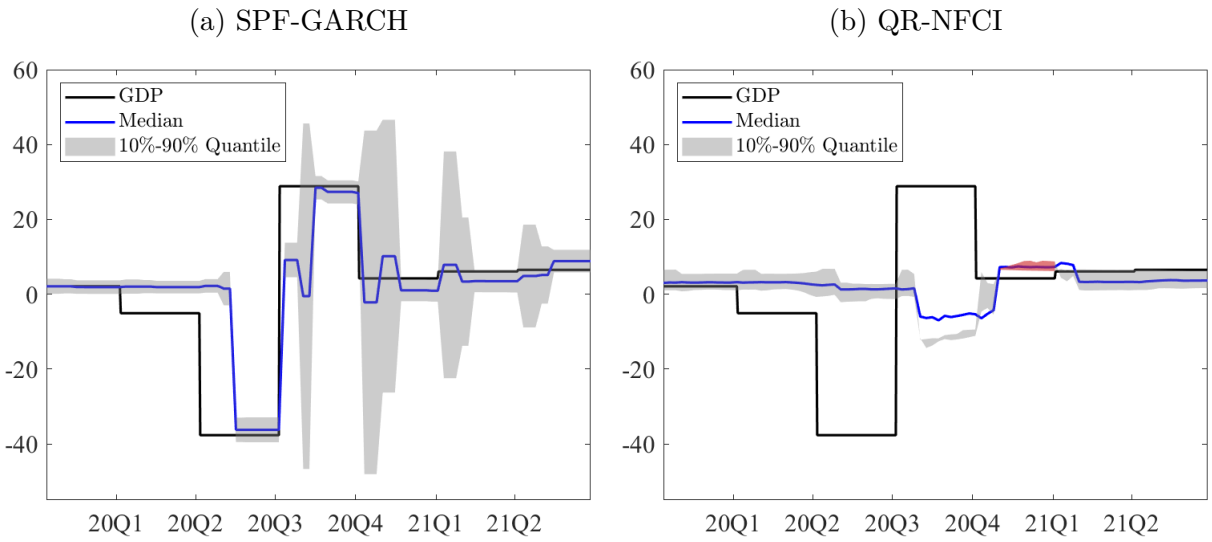
6.4 GDP growth and skewness

The preferred model for nowcasting quantiles and the density of GDP growth is the GARCH model employing SPF nowcasts in the conditional mean. Up to now, to make estimation and forecasting operational, GDP growth is assumed to be normally distributed. However, the evidence provided by Adrian et al. (2019) and Delle Monache et al. (2023),

[7]Prior to the pandemic, there is no instance of the 10% and 90% quantiles crossing. Only in one quarter in the early nineties does the median lie above the 90% quantile.

[8]From a practitioner's perspective, the large conditional variances in the quarters following the COVID-19 pandemic are undesirable. An easy fix is to treat the onset of the pandemic as an extreme outlier and replace the SPF one-step ahead forecast by the median nowcast for 2020:Q2 in the subsequent quarters.

Figure 4: Real-time nowcasts through the COVID-19 pandemic



Notes: This figure depicts real-time nowcasts of quarterly US GDP growth during the COVID-19 pandemic from 2029:Q4 to 2021:Q2 for the SPF-GARCH model in panel (a) and the QR-NFCI approach in panel (b). The black line represents actual GDP growth, while the blue line shows the predicted median forecasts. In panel (a), where the SPF-GARCH model assumes a normal distribution, the median corresponds to the mean. The grey area indicates density forecasts ranging from the 10% to the 90% conditional quantiles. The red area in panel (b) highlights periods where the predicted 10% quantile exceeds the 90% quantile.

among others, suggests that the conditional GDP growth distribution for the US exhibits time-varying skewness and moderately fat tails. They specifically find that US GDP growth is negatively skewed during economic downturns, and vice versa. In contrast, Brownlees and Souza (2021) and Carriero et al. (2022) do not find that GDP growth is significantly skewed.

Owing to the mixed evidence in the literature, I first revisit the out-of-sample predictive performance of the SPF-GARCH model, incorporating skewness and fat tails by assuming that GDP growth follows a skewed t-distribution. Subsequently, I present in-sample evidence on the distribution of GDP growth following the methodology proposed by Adrian et al. (2019).

6.4.1 SPF-GARCH with skew t-distribution

In line with Adrian et al. (2019), the skew t-distribution used in this analysis is the one developed by Azzalini and Capitanio (2003). The probability density function of y, omitting subscripts for notational convenience, is given by

$$f(y;\mu,\sigma,\xi,\nu) = \frac{2}{\sigma} t\left(\frac{y-\mu}{\sigma};\nu\right) T\left(\xi\frac{y-\mu}{\sigma}\sqrt{\frac{\nu+1}{\nu+\frac{(y-\mu)^2}{\sigma^2}}};\nu+1\right), \tag{14}$$

where $t(\cdot)$ and $T(\cdot)$ denote the pdf and the cdf of the Student t-distribution, μ determines the location, and σ is the scale parameter. The shape parameter ξ controls the skewness, and the degrees of freedom parameter ν governs the fatness. During the evaluation period, the conditional skew t-distribution of GDP growth for each week in each quarter is denoted by

$$y_{t|t,w} \sim \text{St}(\mu_{t|t,w}, \sigma_{t|t,w}, \xi_{t|t,w}, \nu_{t|t,w}), \tag{15}$$

where $\mu_{t|t,w}$ and $\sigma_{t|t,w}$ are determined by equations (4) and (5) correspondingly. The log-likelihood of GARCH-type models employing a skew t-distribution follows directly from the density in equation (14). The conditional quantiles are computed using numerical integration of the quantile function.

As special cases, the density in (14) reduces to the Student t-distribution for $\xi = 0$ and to the skew normal distribution for $\xi \neq 0$ and $\nu = \infty$. When $\xi = 0$ and $\nu = \infty$, the distribution simplifies to the normal distribution with mean μ and variance σ^2.

The results of the different SPF-GARCH specifications are presented in Table 8, where 'Skewed t' indicates the average losses associated with the unrestricted skew-t distribution. 'Skewed Normal' and 'Symmetric t' refer to the restricted specifications, featuring skewness or fatness, respectively, and 'Normal' denotes the specification with normally distributed errors studied in Sections 6.1.

The findings suggest limited benefits from departing from the normality assumption, as seen by the similar out-of-sample losses of all specifications. In particular, across the entire distribution of GDP growth, fitting the unrestricted skew t-distribution does not improve the prediction accuracy. Only the skew normal distribution leads to small but significant improvements at the 90% quantile. While this lack of noteworthy improvements doesn't imply symmetry in US GDP growth, the evidence suggests that, at least with the limited sample sizes available, and the short forecasting horizon considered, incorporating both skewness and fatness does not yield out-of-sample forecasting enhancements.

Figure A.1 demonstrates that, conditional on the SPF nowcasts, the skew t-distribution exhibits pronounced variations in the conditional mean in panel (b) and some volatility dynamics in panel (d). Similarly, panels (b) and (d) of Figure A.2 show that skewness is mostly centered around zero, while the kurtosis suggests moderately fat tails. In Figure A.3, panels (a) and (b) depict skewness for the skew normal distribution, while panels (c) and (d) illustrate kurtosis for the Student t-distribution. Once conditioned on the SPF nowcasts, the evidence again suggests a symmetric conditional GDP growth distribution with mild excess kurtosis.

Table 8: Out-of-sample evaluation of SPF-GARCH with skew t-distribution

	week							
10% quantile score	1	3	5	7	9	11	13	15
Normal	0.36	0.35	0.35	0.28	**0.29**	**0.29**	**0.29**	**0.29**
Skewed Normal	0.37	0.37	0.35	0.28	0.29	0.30	0.30	0.30
Symmetric t	**0.35**	**0.35**	**0.34**	**0.28**	0.29	0.29	0.29	0.29
Skewed t	0.36	0.36	0.34	0.28	0.30	0.30	0.30	0.30
90% quantile score	1	3	5	7	9	11	13	15
Normal	0.34	0.34	0.34	0.31	0.31	0.31	0.31	0.31
Skewed Normal	0.35	0.36	0.35	**0.31**	**0.30**	**0.30**	**0.30**	**0.30**
Symmetric t	**0.34**	**0.34**	**0.33**	0.31	0.31	0.31	0.31	0.31
Skewed t	0.35	0.35	0.34	0.32	0.31	0.31	0.31	0.31
CRPS	1	3	5	7	9	11	13	15
Normal	1.12	1.12	1.09	0.99	0.96	0.96	**0.96**	**0.96**
Skewed Normal	1.14	1.13	1.10	0.99	**0.95**	**0.96**	0.96	0.96
Symmetric t	**1.12**	**1.11**	**1.08**	0.99	0.96	0.96	0.96	0.96
Skewed t	1.14	1.14	1.09	**0.98**	0.97	0.97	0.97	0.97

Notes: This table reports the average losses of the 10% and the 90% quantile scores in the top and the middle panel. The panel at the bottom shows the average continuously ranked probability score (CRPS) over the out-of-sample evaluation period. The columns indicate the week of a quarter at which the nowcast is formed. Gray areas represent the 90% model confidence set and bold letters are the lowest average losses within each column. All models are estimated on an expanding window. The evaluation period is 1990:Q1 to 2019:Q4, consisting of 120 observations.

6.4.2 A quantile matching approach

From a forecasting perspective, incorporating skewness and fatness does not enhance the predictive performance of the SPF-GARCH model. This lack of improvement could be attributed to conditional GDP growth potentially exhibiting symmetry and not being excessively fat-tailed. Alternatively, the flexibility offered by the skew t-distribution introduces increased estimation uncertainty, which might not be justified in a forecasting exercise with relatively small sample sizes. This section aims to provide further insights into the influence of conditioning information on the distribution of GDP growth, taking an in-sample perspective.

To illustrate the impact of the conditioning information on the conditional GDP growth distribution, I employ the semi-parametric quantile matching approach introduced by Adrian et al. (2019). First, I estimate quantile regressions of GDP growth on a constant and one explanatory variable, i.e., either the quarterly lag of NFCI or the SPF, for the 10%, 25%, 75%, and 90% quantiles. Given the in-sample approach, I use the current vintages of GDP growth and the NFCI, and the median nowcasts of the SPF. The sample

studied covers the period from 1973:Q1 to 2019:Q4.

Second, for each quarter, the parameters of the skew t-distribution are obtained through minimizing the squared distance between the estimated quantiles, denoted as $\hat{Q}^{\tau}_{t|x_{t-1}}$, where x_{t-1} represents either the NFCI or the median SPF nowcast, and the quantile function of the skew t-distribution $F^{-1}(\tau, \mu_t, \sigma_t, \xi_t, \nu_t)$ based on the density defined in equation (14). Therefore, the parameters of the skew t-distribution are computed according to

$$\{\hat{\mu}_t, \hat{\sigma}_t, \hat{\xi}_t, \hat{\nu}_t\} = \underset{\mu,\sigma,\xi,\nu}{\operatorname{argmin}} \sum_{\tau} \left(\hat{Q}^{\tau}_{t|x_{t-1}} - F^{-1}(\tau; \mu, \sigma, \xi, \nu)\right)^2, \tag{16}$$

where $\tau = \{0.1, 0.25, 0.75, 0.9\}$ represents the respective quantiles.

Figure A.4 illustrates the evolution of the implied mean and variance of the skew t-distribution. Panels (a) and (b) show large changes in the implied mean over time, with more pronounced swings when conditioning on the SPF compared to the NFCI. Additionally, panels (c) and (d) depict time variation in the implied variance. When conditioning on the NFCI, the skew t-distribution exhibits notable spikes in the variance during the 1970s, 1980s, and the global financial crisis, reflecting its limited ability to account for dynamics in the conditional mean compared to the SPF.

Figure A.5 compares the impact of the conditioning information on the implied skewness in panels (a) and (b) and the kurtosis in panels (c) and (d). Conditional on both the NFCI and the SPF, there is evidence in favor of time-varying skewness. However, when the SPF median nowcasts are incorporated, there is less pronounced variation. Regarding the kurtosis, both approaches indicate some level of excess kurtosis, implying that conditional GDP growth is moderately fat-tailed. In general, there appears to be no clear link between the dynamics in the kurtosis and specific economic events of turmoil.

In summary, the in-sample evidence demonstrates that the shape of the skew t-distribution crucially depends on the conditioning information. While conditioning on the NFCI results in predominant time-variation in the implied variance and skewness, incorporating the SPF median nowcasts suggests that dynamics in the conditional mean are the most notable feature of the conditional GDP growth distribution. Although there are some movements in the skewness, the findings based on the SPF median nowcasts are more in line with the evidence provided by Brownlees and Souza (2021) and Carriero et al. (2022). The resulting shape of GDP growth appears slightly negatively skewed and mildly fat-tailed, which helps to explain why departing from normality does not yield notable out-of-sample predictive improvements in the previous section.

6.5 Robustness analysis

6.5.1 Mean versus median SPF forecasts

Regarding the measure of the SPF consensus projections, both the mean and the median of the panelists' forecasts could be employed. So far, the median SPF forecast was treated as the consensus forecast, which is more robust to extreme point predictions of individual forecasters of the SPF. This section compares the mean and the median SPF projections as conditioning information for the SPF-GARCH model.

Table B.1 shows the out-of-sample results of the SPF-GARCH model incorporating either the mean or the median SPF projections. For the 10% and 90% quantiles, the differences in the prediction accuracy are small and insignificant. For the predictive density, the median SPF results in significant, albeit small, improvements. Table B.2 exhibits similar results for quantile regressions. Thus, the findings of this paper are robust to the choice of measure to construct point forecasts from the SPF panelists.

6.5.2 Forecaster disagreement

The SPF consensus forecasts derived from individual point predictions offer a useful summary of the economic outlook provided by the survey panelists. However, relying solely on a single measure of central tendency overlooks the rich information contained within the survey data.

One approach to capturing the cross-sectional dispersion in expectations among forecasters is to consider their *disagreement*. Abel et al. (2016) suggest the interquartile range (IQR) as a measure of forecasters' disagreement according to

$$SPF_{t,w}^{IQR} = SPF_{t,w}^{0.75} - SPF_{t,w}^{0.25},$$

where $SPF_{t,w}^{0.25}$ and $SPF_{t,w}^{0.75}$ denote the 25% and 75% quantiles of the panelists' point forecasts. The SPF also elicits probabilistic predictions through histogram forecasts, allowing for deriving measures of subjective *uncertainty*. However, the SPF asks respondents about fixed-event density forecasts, i.e., GDP growth for the current and the following calendar year. While disagreement is not a direct measure of uncertainty (see, for example, Abel et al., 2016; Glas, 2020), fixed-horizon disagreement is readily available from individual point predictions.

Table B.3 presents the out-of-sample results when incorporating the IQR into the SPF-GARCH and the QR-SPF model. Adding the interquartile range to the conditional volatility of the SPF-GARCH model significantly improves the nowcasts of Growth-at-Risk and results in the lowest CRPS. Thus, periods of increased dispersion among survey

panelists positively affect the conditional variance, which helps explain downside risk, although only by a small margin. At the 90% quantile, the SPF-GARCH augmented by the interquartile range in the conditional mean performs best. However, these refinements are small, implying that out-of-sample nowcast improvements are mainly due to the SPF consensus forecasts rather than disagreement among panelists.

The QR-SPF model including the IQR is mostly contained in the MCS. Especially at the 90% quantile, conditional on the one-step ahead forecasts, including disagreement in the conditioning information of quantile regressions results in the lowest forecast errors. Conditional on the SPF nowcasts, the GARCH specifications tend to perform better. Overall, the evidence suggests that survey disagreement is informative, but improvements over the models exploiting the median SPF forecasts are modest.

6.5.3 Consensus forecast error distribution

An alternative strategy to gauge uncertainty around consensus forecasts is to exploit the historical forecast errors (Reifschneider and Tulip, 2019). One straightforward approach is to estimate the conditional τ-quantile of GDP growth using the τ-quantile of past forecast errors added to the median SPF forecast. Thus, the nowcast is estimated as

$$\hat{Q}^{\text{error}}_{t|t,w}(\tau) = SPF_{t,w} + \hat{q}^{\tau}_{e^{SPF}_{t,w}}, \tag{17}$$

where $SPF_{t,w}$ is the median SPF one-step ahead forecast or nowcast, depending on week w within quarter t, and $\hat{q}^{\tau}_{e^{SPF}_{t,w}}$ denotes the empirical τ-quantile of the real-time SPF median forecast error, $e^{SPF}_{t,w} \equiv y_{t,w} - SPF_{t,w}$.

Adams et al. (2021) show that financial conditions help explain dynamics in the distribution of the median SPF forecast errors over time. They employ quantile regressions to model the τ-quantile of the SPF median forecast error distribution conditional on the NFCI according to

$$Q^{\text{NFCI}}_{e^{SPF}_{t,w}}(\tau) \equiv Q_{e^{SPF}_{t,w}}(\tau|x_{t,w}) = x_{t,w}\beta^{\tau}_{w},$$

where $x_{t,w}$ contains a constant and the real-time NFCI in week w of quarter t. The predicted quantile of GDP growth is obtained as

$$\hat{Q}^{\text{NFCI}}_{t|t,w}(\tau) = SPF_{t,w} + \hat{Q}^{\text{NFCI}}_{e^{SPF}_{t,w}}(\tau). \tag{18}$$

Table B.4 compares the out-of-sample accuracy of the SPF-GARCH model with the approaches outlined in equations (17) and (18). The SPF-GARCH model demonstrates superior forecasting performance for downside risk. However, conditional on the SPF

nowcasts, the 90% quantile forecasts based on $\hat{Q}^{\text{error}}_{t|t,w}(\tau)$ and $\hat{Q}^{\text{NFCI}}_{t|t,w}(\tau)$ show lower average quantile losses compared to SPF-GARCH, yet these differences are not statistically significant. Furthermore, the conditional density forecasts indicate similar predictive performances among all three methods. Consequently, the SPF-GARCH model remains the most suitable choice, particularly for forecasting Growth-at-Risk.

6.5.4 NFCI at higher frequency

Up to now, the NFCI has been analyzed only at a quarterly frequency, utilizing observations from the latest available vintage within week w. Since this approach discards numerous observations, another alternative is to aggregate weekly NFCI by means of a mixed-data sampling (MIDAS) approach. The results of Carriero et al. (2022) indicate that the information flow from weekly financial indicators does not substantially enhance the nowcasting accuracy compared to monthly aggregates, although it is not harmful either. Castelnuovo and Mori (2022) thus propose to use quantile regressions with three monthly lags, i.e., $x_{t,w} = (NFCI_{t,w-1}, NFCI_{t,w-5}, NFCI_{t,w-9})$. They estimate the model quarterly and find modest yet significant improvements. Therefore, this monthly aggregation scheme, denoted by $NFCI^{m}_{t,w}$, is tested in real-time. Despite the findings of Carriero et al. (2022), a weekly MIDAS approach is employed to assess the weekly information flow more efficiently. The weekly NFCI data is aggregated according to

$$NFCI^{\theta}_{t,w} = \sum_{i=1}^{m} \theta_i(\lambda_1, \lambda_2) NFCI_{t,w-i},$$

where $\theta_i(\lambda_1, \lambda_2)$ is the unrestricted beta weight

$$\theta_i(\lambda_1, \lambda_2) = \frac{(i/(m+1))^{\lambda_1 - 1} (1 - i/(m+1))^{\lambda_2 - 1}}{\sum_{k=1}^{m} (i/(m+1))^{\lambda_1 - 1} (1 - i/(m+1))^{\lambda_2 - 1}}, \tag{19}$$

with $\lambda_1 \geq 1$, $\lambda_2 \geq 1$, and m is the lag length of weekly NFCI realizations. Motivated by using three monthly lags, the weekly lag length is chosen to include twelve weeks, i.e., $m = 12$.

Table B.5 presents the out-of-sample results for the SPF-GARCH model, both without NFCI and with quarterly NFCI incorporated into the conditional mean. In addition, this table displays the outcomes for the aggregated versions, $NFCI^{m}_{t,w}$ and $NFCI^{\theta}_{t,w}$, respectively. Monthly aggregation initially shows mild nowcasting gains at the 10% quantile compared to weekly aggregation, aligning with previous studies by Castelnuovo and Mori (2022) and Carriero et al. (2022). However, following the release of the SPF nowcast, the utilization of NFCI at a higher frequency does not notably enhance the prediction accuracy of GaR. At the 90% quantile and in terms of the CRPS, the SPF-

GARCH model without NFCI tends to perform best. Nevertheless, the differences in the average losses are generally small and insignificant.

Similar insights are obtained when the NFCI is used as conditioning information in the volatility of the SPF-GARCH model (Table B.6). Regarding quantile regressions, monthly and weekly aggregation does not help increase the nowcasting accuracy of Growth-at-Risk either (Table B.7). Overall, when taking the SPF into account, the real-time analysis shows that there is little reason to employ a MIDAS approach to the NFCI.

6.5.5 SPF recession probability

The SPF also asks respondents about the probability for a decline in real GDP in the current quarter and the following four quarters, also referred to as the "Anxious Index". Especially at short horizons, the average recession probability across panelists provides accurate forecasts for economic downturns (Lahiri and Wang, 2013). Depending on the week w of the quarter t, let $SPF^{\text{rec}}_{t,w}$ denote the one-step ahead or the nowcast of the SPF mean probability for a negative growth rate in quarter t.

To assess the potential benefits of incorporating the subjective recession probability, Table B.8 presents the out-of-sample results integrating $SPF^{\text{rec}}_{t,w}$ into both the SPF-GARCH model and the QR-SPF approach. When conditioned on the one-step ahead SPF forecasts, the utilization of the anxious index yields marginal improvements in the predictive accuracy for downside risks within the quantile regression framework. However, the SPF-GARCH model is never outperformed and provides the smallest average losses across the entire distribution.

6.5.6 Rolling window estimation

To address concerns related to asymptotically vanishing estimation uncertainty due to a recursive estimation scheme (Giacomini and White, 2006), Table B.9 replicates the main results of this paper using a rolling window of 85 observations, corresponding to the number of SPF surveys available in the first quarter of 1990. Overall, the key findings remain robust. The increased estimation uncertainty affects the power of the MCS at the 10% quantile, resulting in nonsignificant results. Nevertheless, the SPF-GARCH model consistently demonstrates the lowest losses once the nowcast is available. For the 90% quantile score and the CRPS, the differences remain significant.

7 Conclusion

This study investigates the effectiveness of using SPF median projections for nowcasting quarterly GaR in the US. Nowcasts are constructed using both AR-GARCH type models

and quantile regressions. In a real-time out-of-sample forecasting evaluation, incorporating SPF consensus forecasts into the conditional mean of the GARCH model demonstrates superior predictive performance for GaR and the conditional density of GDP growth, especially following the release of the SPF nowcasts mid-quarter.

The SPF consensus projections are informative across the entire distribution of GDP growth, and the results based on the SPF-GARCH model, assuming normality, suggest notable time variation in both the lower and upper quantiles of GDP growth. Departing from symmetry, and incorporating fatness, does not enhance the forecasting accuracy for downside risk. While this does not imply that the conditional distribution of GDP growth is symmetric, these findings indicate that, given the short samples at hand, a more flexible distribution is not justified in the context of nowcasting GaR.

The SPF consensus forecasts emerge as highly informative for tail risk and the entire predictive distribution of GDP growth, particularly at short horizons. This study advocates for the inclusion of SPF consensus forecasts in the macroeconomic toolkit for real-time assessment of downside risks. While the focus is on US GDP, future research could extend these findings to the ECB SPF as well as explore Inflation-at-Risk and Unemployment-at-Risk.

References

Abel, J., R. Rich, J. Song, and J. Tracy (2016). The measurement and behavior of uncertainty: Evidence from the ECB survey of professional forecasters. *Journal of Applied Econometrics 31*(3), 533–550.

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

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

Adrian, T., F. Grinberg, N. Liang, S. Malik, and J. Yu (2022). The term structure of Growth-at-Risk. *American Economic Journal: Macroeconomics 14*(3), 283–323.

Amburgey, A. J. and M. W. McCracken (2023). On the real-time predictive content of financial condition indices for growth. *Journal of Applied Econometrics 38*(2), 137–163.

Andrade, P., E. Ghysels, and J. Idier (2014). Inflation risk measures and their informational content. *Available at SSRN 2439607*.

Andreou, E., E. Ghysels, and A. Kourtellos (2013). Should macroeconomic forecasters use daily financial data and how? *Journal of Business & Economic Statistics 31*(2), 240–251.

Ang, A., G. Bekaert, and M. Wei (2007). Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics 54*(4), 1163–1212.

Azzalini, A. and A. Capitanio (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 65*(2), 367–389.

Bańbura, M., F. Brenna, J. Paredes, and F. Ravazzolo (2021). Combining Bayesian VARs with survey density forecasts: Does it pay off?

Bańbura, M., D. Giannone, M. Modugno, and L. Reichlin (2013). Now-casting and the real-time data flow. In *Handbook of Economic Forecasting*, Volume 2, pp. 195–237. Elsevier.

Bassetti, F., R. Casarin, and M. Del Negro (2023). Inference on probabilistic surveys in macroeconomics with an application to the evolution of uncertainty in the survey of professional forecasters during the covid pandemic. In *Handbook of Economic Expectations*, pp. 443–476. Elsevier.

Brave, S. and R. A. Butters (2018). Diagnosing the financial system: Financial conditions and financial stress. *29th issue (June 2012) of the International Journal of Central Banking*.

Brownlees, C. and A. B. Souza (2021). Backtesting global growth-at-risk. *Journal of Monetary Economics 118*, 312–330.

Carriero, A., T. E. Clark, and M. Marcellino (2020). Capturing macro-economic tail risks with bayesian vector autoregressions. *Journal of Money, Credit and Banking*.

Carriero, A., T. E. Clark, and M. Marcellino (2022). Nowcasting tail risk to economic activity at a weekly frequency. *Journal of Applied Econometrics 37*(5), 843–866.

Castelnuovo, E. and L. Mori (2022). Uncertainty, skewness, and the business cycle through the midas lens.

Clark, T. E., M. W. McCracken, and E. Mertens (2020). Modeling time-varying uncertainty of multiple-horizon forecast errors. *Review of Economics and Statistics 102*(1), 17–33.

Clements, M. P. (2010). Explanations of the inconsistencies in survey respondents' forecasts. *European Economic Review 54*(4), 536–549.

Clements, M. P. (2014). Probability distributions or point predictions? Survey forecasts of US output growth and inflation. *International Journal of Forecasting 30*(1), 99–117.

Clements, M. P., R. W. Rich, and J. S. Tracy (2023). Surveys of professionals. In *Handbook of Economic Expectations*, pp. 71–106. Elsevier.

De Nicolò, G. and M. Lucchetta (2017). Forecasting tail risks. *Journal of Applied Econometrics 32*(1), 159–170.

Delle Monache, D., A. De Polis, and I. Petrella (2023). Modeling and forecasting macroeconomic downside risk. *Journal of Business & Economic Statistics*, 1–27.

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

Elliott, G., I. Komunjer, and A. Timmermann (2008). Biases in macroeconomic forecasts: Irrationality or asymmetric loss? *Journal of the European Economic Association 6*(1), 122–157.

Fagiolo, G., M. Napoletano, and A. Roventini (2008). Are output growth-rate distributions fat-tailed? Some evidence from OECD countries. *Journal of Applied Econometrics 23*(5), 639–669.

Faust, J. and J. H. Wright (2013). Forecasting inflation. In *Handbook of Economic Forecasting*, Volume 2, pp. 2–56. Elsevier.

Ferrara, L., M. Mogliani, and J.-G. Sahuc (2022). High-frequency monitoring of growth at risk. *International Journal of Forecasting 38*(2), 582–595.

Figueres, J. M. and M. Jarociński (2020). Vulnerable growth in the euro area: Measuring the financial conditions. *Economics Letters 191*, 109126.

Ganics, G., B. Rossi, and T. Sekhposyan (2020). From fixed-event to fixed-horizon density forecasts: Obtaining measures of multihorizon uncertainty from survey density forecasts. *Journal of Money, Credit and Banking*.

Ghysels, E., P. Santa-Clara, and R. Valkanov (2004). The MIDAS touch: Mixed data sampling regression models.

Ghysels, E., A. Sinko, and R. Valkanov (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews 26*(1), 53–90.

Giacomini, R. and H. White (2006). Tests of conditional predictive ability. *Econometrica 74*(6), 1545–1578.

Giannone, D., L. Reichlin, and D. Small (2008). Nowcasting: The real-time informational content of macroeconomic data. *Journal of Monetary Economics 55*(4), 665–676.

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

Glas, A. (2020). Five dimensions of the uncertainty–disagreement linkage. *International Journal of Forecasting 36*(2), 607–627.

Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association 106*(494), 746–762.

Gneiting, T. and A. E. Raftery (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association 102*(477), 359–378.

Hansen, P. R., A. Lunde, and J. M. Nason (2011). The model confidence set. *Econometrica 79*(2), 453–497.

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

Krüger, F. and H. Plett (2024). Prediction intervals for economic fixed-event forecasts. *Annals of Applied Statistics*, Forthcoming.

Lahiri, K. and J. G. Wang (2013). Evaluating probability forecasts for GDP declines using alternative methodologies. *International Journal of Forecasting 29*(1), 175–190.

Mincer, J. A. and V. Zarnowitz (1969). The evaluation of economic forecasts. In *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, pp. 3–46. NBER.

Plagborg-Møller, M., L. Reichlin, G. Ricco, and T. Hasenzagl (2020). When is growth at risk? *Brookings Papers on Economic Activity 2020*(1), 167–229.

Prasad, M. A., S. Elekdag, M. P. Jeasakul, R. Lafarguette, M. A. Alter, A. X. Feng, and C. Wang (2019). *Growth at risk: Concept and application in IMF country surveillance*. International Monetary Fund.

Reichlin, L., G. Ricco, and T. Hasenzagl (2020). Financial variables as predictors of real growth vulnerability.

Reifschneider, D. and P. Tulip (2019). Gauging the uncertainty of the economic outlook using historical forecasting errors: The Federal Reserve's approach. *International Journal of Forecasting 35*(4), 1564–1582.

Rossi, B. and T. Sekhposyan (2019). Alternative tests for correct specification of conditional predictive densities. *Journal of Econometrics 208*(2), 638–657.

Sheppard, K. (2009). MFE MATLAB function reference financial econometrics. *URL https://www.kevinsheppard.com/files/code/matlab/mfe-toolbox-documentation.pdf*.

Stark, T. et al. (2010). Realistic evaluation of real-time forecasts in the survey of professional forecasters. *Federal Reserve Bank of Philadelphia Research Rap, Special Report 1*.

Wallis, K. F. (1986). Forecasting with an econometric model: The 'ragged edge' problem. *Journal of Forecasting 5*(1), 1–13.

West, K. D. (1996). Asymptotic inference about predictive ability. *Econometrica 64*(5), 1067–1084.

Zarnowitz, V. (1985). Rational expectations and macroeconomic forecasts. *Journal of Business & Economic Statistics 3*(4), 293–311.

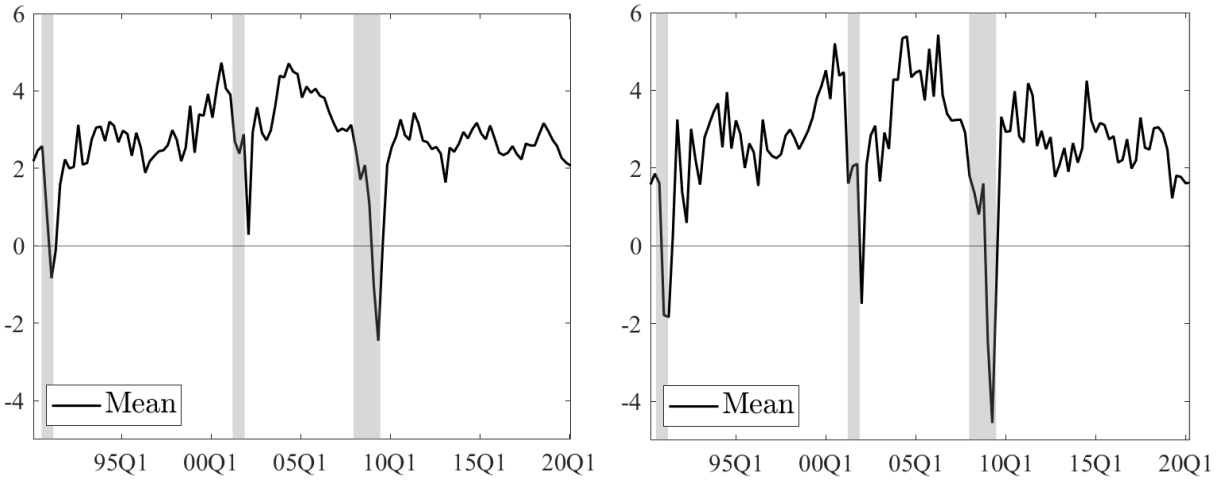
Appendices

A Figures

Figure A.1: Mean and variance of the skew t-distribution

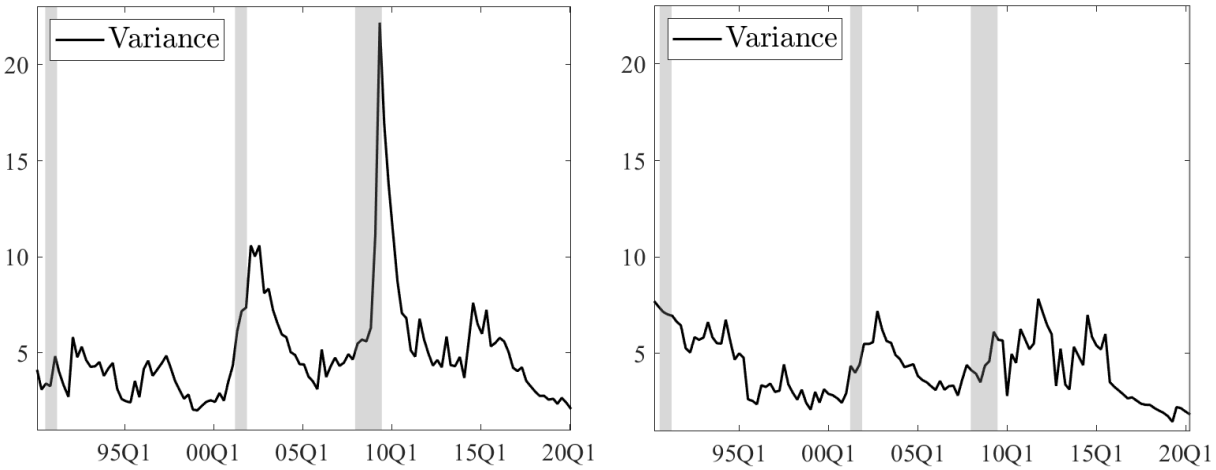
(a) Mean of skew t: $\text{SPF}_{t|t-1}$

(b) Mean of skew t: $\text{SPF}_{t|t}$



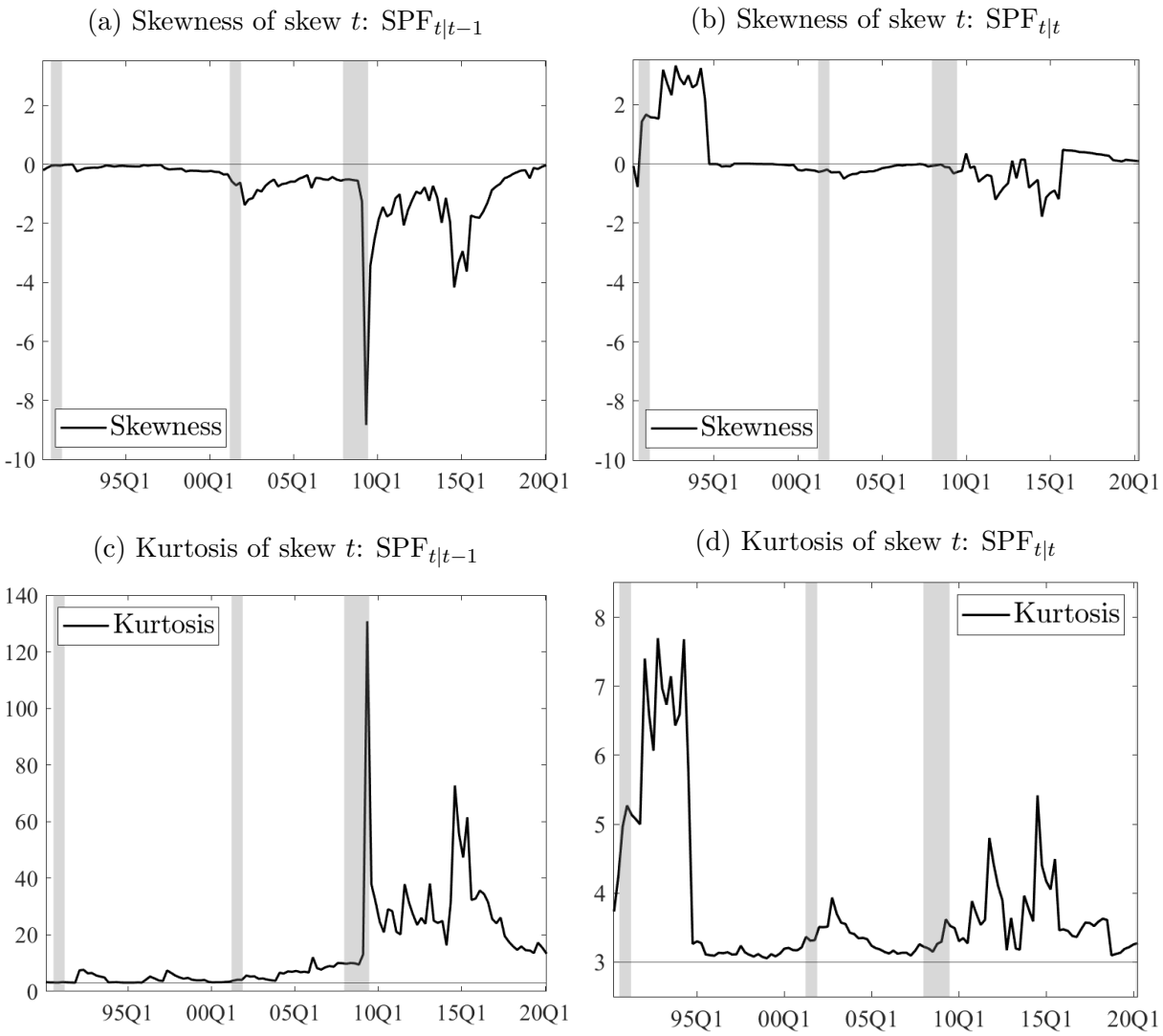
(c) Variance of skew t: $\text{SPF}_{t|t-1}$

(d) Variance of skew t: $\text{SPF}_{t|t}$



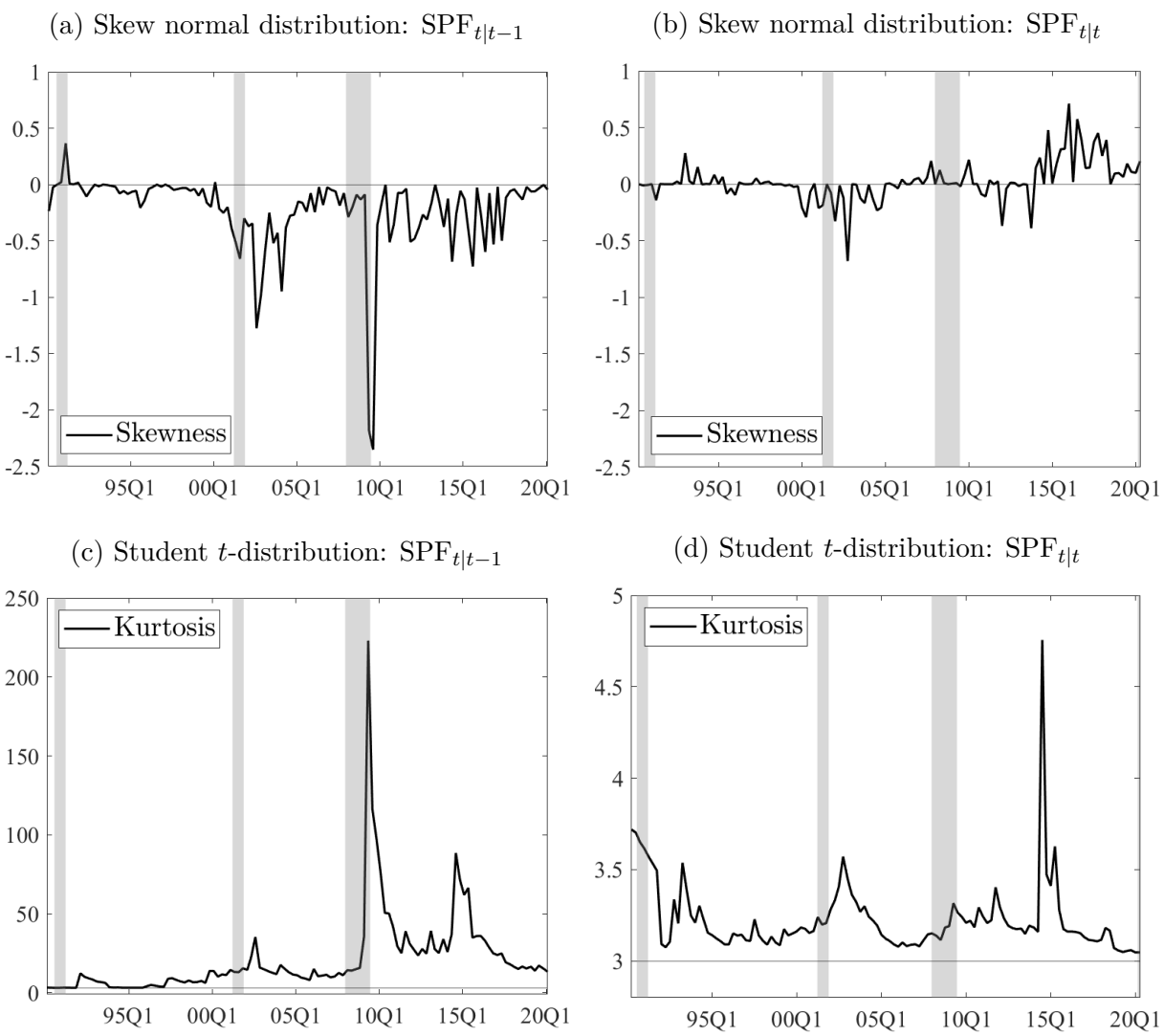
Notes: This figure illustrates real-time estimates of the mean and variance of the skew t-distribution over the out-of-sample evaluation period from 1990:Q1 to 2019:Q4 for the SPF-GARCH model. Panels (a) and (c) utilize the SPF one-step ahead median forecasts, while panels (b) and (d) employ the nowcasts. The shaded areas indicate periods of US recessions as identified by the GDP-based recession indicator.

Figure A.2: Skewness and kurtosis of the skew t-distribution



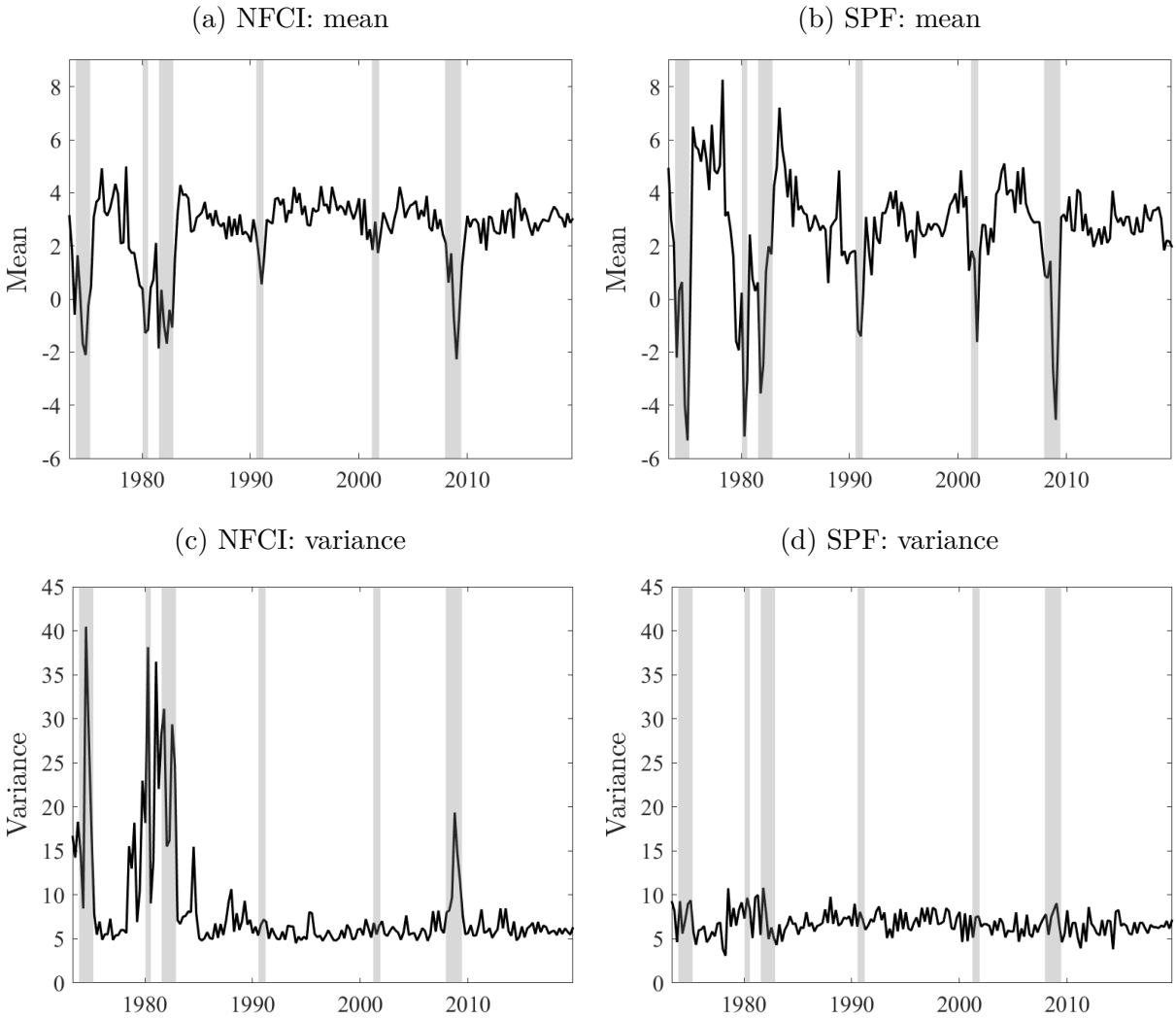
Notes: This figure illustrates real-time estimates of the skewness and kurtosis of the skew t-distribution over the out-of-sample evaluation period from 1990:Q1 to 2019:Q4 for the SPF-GARCH model. Panels (a) and (c) utilize the SPF one-step ahead median forecasts, while panels (b) and (d) employ the nowcasts. The shaded areas indicate periods of US recessions as identified by the GDP-based recession indicator.

Figure A.3: Skewness of the skew normal and kurtosis of the t-distribution



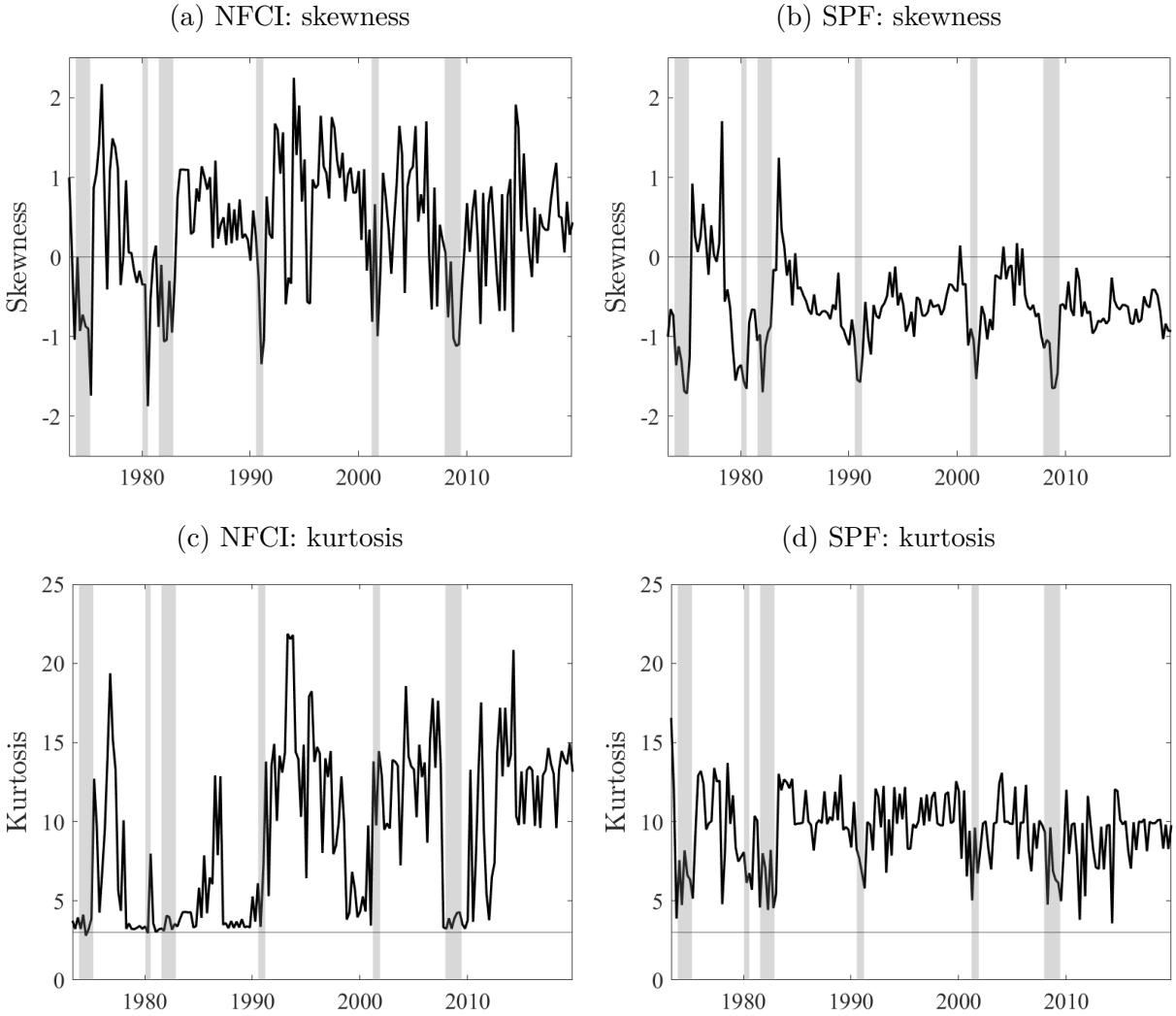
Notes: This figure illustrates real-time estimates of the skewness of the skew normal distribution and the kurtosis of the Student t-distribution over the out-of-sample evaluation period from 1990:Q1 to 2019:Q4 for the SPF-GARCH model. Panels (a) and (c) utilize the SPF one-step ahead median forecasts, while panels (b) and (d) employ the nowcasts. The shaded areas indicate periods of US recessions as identified by the GDP-based recession indicator.

Figure A.4: In-sample mean and variance



Notes: This figure illustrates the in-sample mean and variance from 1973:Q1 to 2019:Q4 based on the semi-parametric quantile matching approach outlined in Section 6.4.2. As conditioning information, panels (a) and (c) utilize the latest vintage of the NFCI, while panels (b) and (d) employ the SPF median nowcasts. The shaded areas indicate periods of US recessions as identified by the GDP-based recession indicator.

Figure A.5: In-sample skewness and kurtosis



Notes: This figure illustrates the in-sample skewness and kurtosis from 1973:Q1 to 2019:Q4 based on the semi-parametric quantile matching approach outlined in Section 6.4.2. As conditioning information, panels (a) and (c) utilize the latest vintage of the NFCI, while panels (b) and (d) employ the SPF median nowcasts. The shaded areas indicate periods of US recessions as identified by the GDP-based recession indicator.

B Tables

Table B.1: SPF-GARCH using median versus mean SPF

	week							
10% quantile score	1	3	5	7	9	11	13	15
mean SPF	0.36	0.35	0.35	0.28	**0.29**	**0.29**	**0.29**	**0.29**
median SPF	**0.36**	**0.35**	**0.35**	**0.28**	0.29	0.29	0.29	0.29
90% quantile score	1	3	5	7	9	11	13	15
mean SPF	0.34	0.34	0.34	0.32	0.31	0.31	0.31	0.31
median SPF	**0.34**	**0.34**	**0.34**	**0.31**	**0.31**	**0.31**	**0.31**	**0.31**
CRPS	1	3	5	7	9	11	13	15
mean SPF	1.13	1.12	1.10	1.01	0.98	0.98	0.98	0.98
median SPF	**1.12**	**1.12**	**1.09**	**0.99**	**0.96**	**0.96**	**0.96**	**0.96**

Notes: This table reports the average losses of the 10% and the 90% quantile scores in the top and the middle panel. The panel at the bottom shows the average continuously ranked probability score (CRPS) over the out-of-sample evaluation period. The columns indicate the week of a quarter at which the nowcast is formed. Gray areas represent the 90% model confidence set and bold letters are the lowest average losses within each column. All models are estimated on an expanding window. The evaluation period is 1990:Q1 to 2019:Q4, consisting of 120 observations.

Table B.2: QR-SPF using median versus mean SPF

	week							
10% quantile score	1	3	5	7	9	11	13	15
mean SPF	**0.43**	**0.43**	**0.42**	**0.34**	**0.31**	**0.31**	**0.31**	**0.31**
median SPF	0.44	0.44	0.43	0.35	0.31	0.31	0.31	0.31
90% quantile score	1	3	5	7	9	11	13	15
mean SPF	0.39	0.39	0.39	0.36	0.34	0.34	0.34	0.34
median SPF	**0.39**	**0.39**	**0.38**	**0.36**	**0.33**	**0.33**	**0.33**	**0.33**
CRPS	1	3	5	7	9	11	13	15
mean SPF	1.18	1.18	1.15	1.05	1.01	1.01	1.01	1.01
median SPF	**1.17**	**1.17**	**1.14**	**1.03**	**1.00**	**1.00**	**1.00**	**1.00**

Notes: This table reports the average losses of the 10% and the 90% quantile scores in the top and the middle panel. The panel at the bottom shows the average continuously ranked probability score (CRPS) over the out-of-sample evaluation period. The columns indicate the week of a quarter at which the nowcast is formed. Gray areas represent the 90% model confidence set and bold letters are the lowest average losses within each column. All models are estimated on an expanding window. The evaluation period is 1990:Q1 to 2019:Q4, consisting of 120 observations.

Table B.3: Out-of-sample evaluation using the SPF interquantile range

	week							
10% quantile score	1	3	5	7	9	11	13	15
SPF-GARCH	0.36	0.35	0.35	0.28	0.29	0.29	0.29	0.29
$+x^{\mu}_{t,w} = SPF^{IQR}_{t,w}$	0.36	0.36	0.35	0.28	0.29	0.29	0.29	0.29
$+x^{\sigma}_{t,w} = SPF^{IQR}_{t,w}$	**0.35**	**0.35**	**0.34**	**0.27**	**0.28**	**0.28**	**0.28**	**0.28**
QR-SPF	0.44	0.44	0.43	0.35	0.31	0.31	0.31	0.31
$+SPF^{IQR}_{t,w}$	0.37	0.37	0.36	0.31	0.28	0.29	0.29	0.29
90% quantile score	1	3	5	7	9	11	13	15
SPF-GARCH	0.34	0.34	0.34	0.31	0.31	0.31	0.31	0.31
$+x^{\mu}_{t,w} = SPF^{IQR}_{t,w}$	0.33	0.33	0.33	**0.31**	**0.30**	**0.30**	**0.30**	**0.30**
$+x^{\sigma}_{t,w} = SPF^{IQR}_{t,w}$	0.35	0.35	0.34	0.32	0.31	0.31	0.31	0.31
QR-SPF	0.39	0.39	0.38	0.36	0.33	0.33	0.33	0.33
$+SPF^{IQR}_{t,w}$	**0.33**	**0.32**	**0.29**	0.31	0.32	0.32	0.32	0.32
CRPS	1	3	5	7	9	11	13	15
SPF-GARCH	1.12	1.12	1.09	0.99	0.96	0.96	0.96	0.96
$+x^{\mu}_{t,w} = SPF^{IQR}_{t,w}$	1.12	1.12	1.08	0.99	0.95	0.96	0.95	0.95
$+x^{\sigma}_{t,w} = SPF^{IQR}_{t,w}$	**1.11**	**1.11**	1.09	**0.99**	**0.95**	**0.96**	**0.95**	**0.95**
QR-SPF	1.17	1.17	1.14	1.03	1.00	1.00	1.00	1.00
$+SPF^{IQR}_{t,w}$	1.13	1.13	**1.08**	1.00	0.97	0.97	0.97	0.97

Notes: This table reports the average losses of the 10% and the 90% quantile scores in the top and the middle panel. The panel at the bottom shows the average continuously ranked probability score (CRPS) over the out-of-sample evaluation period. The columns indicate the week of a quarter at which the nowcast is formed. Gray areas represent the 90% model confidence set and bold letters are the lowest average losses within each column. All models are estimated on an expanding window. The evaluation period is 1990:Q1 to 2019:Q4, consisting of 120 observations.

Table B.4: Out-of-sample evaluation using SPF forecast errors

	week							
10% quantile score	1	3	5	7	9	11	13	15
SPF-GARCH	**0.36**	**0.35**	**0.35**	**0.28**	**0.29**	**0.29**	**0.29**	**0.29**
$\hat{Q}^{\text{error}}_{t\|t,w}(0.1)$	0.43	0.43	0.43	0.38	0.38	0.38	0.38	0.38
$\hat{Q}^{\text{NFCI}}_{t\|t,w}(0.1)$	0.36	0.36	0.35	0.34	0.36	0.36	0.36	0.38
90% quantile score	1	3	5	7	9	11	13	15
SPF-GARCH	**0.34**	**0.34**	**0.34**	**0.31**	0.31	0.31	0.31	0.31
$\hat{Q}^{\text{error}}_{t\|t,w}(0.9)$	0.40	0.40	0.40	0.34	0.30	0.30	0.30	0.30
$\hat{Q}^{\text{NFCI}}_{t\|t,w}(0.9)$	0.35	0.34	0.34	0.31	**0.29**	**0.28**	**0.28**	**0.28**
CRPS	1	3	5	7	9	11	13	15
SPF-GARCH	1.12	1.12	**1.09**	**0.99**	**0.96**	**0.96**	**0.96**	**0.96**
$\hat{Q}^{\text{error}}_{t\|t,w}(\tau)$	1.13	1.13	1.13	1.04	1.03	1.03	1.03	1.03
$\hat{Q}^{\text{NFCI}}_{t\|t,w}(\tau)$	**1.09**	**1.09**	1.09	1.02	1.03	1.04	1.04	1.05

Notes: This table reports the average losses of the 10% and the 90% quantile scores in the top and the middle panel. The panel at the bottom shows the average continuously ranked probability score (CRPS) over the out-of-sample evaluation period. The columns indicate the week of a quarter at which the nowcast is formed. Gray areas represent the 90% model confidence set and bold letters are the lowest average losses within each column. All models are estimated on an expanding window. The evaluation period is 1990:Q1 to 2019:Q4, consisting of 120 observations.

Table B.5: Out-of-sample evaluation of MIDAS-NFCI in the conditional mean

	week							
10% quantile score	1	3	5	7	9	11	13	15
SPF-GARCH	0.36	0.35	0.35	**0.28**	0.29	0.29	0.29	0.29
$+x^{\mu}_{t,w} = NFCI_{t,w}$	0.37	0.35	0.34	0.28	**0.27**	**0.27**	**0.27**	**0.27**
$+x^{\mu}_{t,w} = NFCI^{m}_{t,w}$	**0.35**	**0.33**	**0.33**	0.30	0.30	0.30	0.31	0.30
$+x^{\mu}_{t,w} = NFCI^{\theta}_{t,w}$	0.38	0.36	0.35	0.29	0.27	0.27	0.27	0.27
90% quantile score	1	3	5	7	9	11	13	15
SPF-GARCH	0.34	**0.34**	0.34	**0.31**	**0.31**	**0.31**	**0.31**	**0.31**
$+x^{\mu}_{t,w} = NFCI_{t,w}$	0.37	0.36	0.34	0.32	0.31	0.31	0.31	0.31
$+x^{\mu}_{t,w} = NFCI^{m}_{t,w}$	**0.34**	0.35	**0.32**	0.33	0.32	0.32	0.33	0.32
$+x^{\mu}_{t,w} = NFCI^{\theta}_{t,w}$	0.38	0.38	0.36	0.33	0.32	0.31	0.31	0.32
CRPS	1	3	5	7	9	11	13	15
SPF-GARCH	**1.12**	**1.12**	1.09	**0.99**	**0.96**	**0.96**	**0.96**	**0.96**
$+x^{\mu}_{t,w} = NFCI_{t,w}$	1.18	1.16	1.09	1.01	0.97	0.97	0.97	0.98
$+x^{\mu}_{t,w} = NFCI^{m}_{t,w}$	1.14	1.14	**1.07**	1.04	1.01	1.00	1.02	1.01
$+x^{\mu}_{t,w} = NFCI^{\theta}_{t,w}$	1.20	1.21	1.14	1.01	0.98	0.98	0.97	0.98

Notes: This table reports the average losses of the 10% and the 90% quantile scores in the top and the middle panel. The panel at the bottom shows the average continuously ranked probability score (CRPS) over the out-of-sample evaluation period. The columns indicate the week of a quarter at which the nowcast is formed. Gray areas represent the 90% model confidence set and bold letters are the lowest average losses within each column. All models are estimated on an expanding window. The evaluation period is 1990:Q1 to 2019:Q4, consisting of 120 observations.

Table B.6: Out-of-sample evaluation of MIDAS-NFCI in the conditional volatility

	week							
10% quantile score	1	3	5	7	9	11	13	15
SPF-GARCH	0.36	0.35	0.35	0.28	0.29	0.29	0.29	0.29
$+x^{\sigma}_{t,w} = NFCI_{t,w}$	**0.35**	**0.34**	**0.33**	**0.28**	**0.28**	**0.28**	**0.28**	0.28
$+x^{\sigma}_{t,w} = NFCI^{m}_{t,w}$	0.60	1.07	0.42	0.32	0.28	0.28	0.29	0.31
$+x^{\sigma}_{t,w} = NFCI^{\theta}_{t,w}$	0.36	0.35	0.34	0.28	0.28	0.28	0.28	**0.27**
90% quantile score	1	3	5	7	9	11	13	15
SPF-GARCH	**0.34**	**0.34**	**0.34**	**0.31**	**0.31**	0.31	**0.31**	**0.31**
$+x^{\sigma}_{t,w} = NFCI_{t,w}$	0.37	0.36	0.36	0.34	0.32	0.32	0.32	0.31
$+x^{\sigma}_{t,w} = NFCI^{m}_{t,w}$	0.60	1.06	0.41	0.34	0.32	**0.31**	0.32	0.32
$+x^{\sigma}_{t,w} = NFCI^{\theta}_{t,w}$	0.36	0.37	0.36	0.34	0.32	0.32	0.32	0.32
CRPS	1	3	5	7	9	11	13	15
SPF-GARCH	**1.12**	**1.12**	**1.09**	**0.99**	**0.96**	**0.96**	**0.96**	**0.96**
$+x^{\sigma}_{t,w} = NFCI_{t,w}$	1.14	1.14	1.10	1.01	0.98	0.98	0.98	0.98
$+x^{\sigma}_{t,w} = NFCI^{m}_{t,w}$	1.54	2.42	1.20	1.04	0.99	0.98	0.99	1.00
$+x^{\sigma}_{t,w} = NFCI^{\theta}_{t,w}$	1.16	1.16	1.11	1.02	0.98	0.98	0.98	0.98

Notes: This table reports the average losses of the 10% and the 90% quantile scores in the top and the middle panel. The panel at the bottom shows the average continuously ranked probability score (CRPS) over the out-of-sample evaluation period. The columns indicate the week of a quarter at which the nowcast is formed. Grey areas represent the 90% model confidence set and bold letters are the lowest average losses within each column. All models are estimated on an expanding window. The evaluation period is 1990:Q1 to 2019:Q4, consisting of 120 observations.

Table B.7: Out-of-sample evaluation of quntile regression with MIDAS-NFCI

	week							
10% quantile score	1	3	5	7	9	11	13	15
QR-SPF	0.44	0.44	0.43	0.35	0.31	0.31	0.31	0.31
$+NFCI_{t,w}$	**0.29**	**0.30**	**0.29**	**0.28**	**0.27**	**0.27**	0.28	0.28
$+NFCI^{m}_{t,w}$	0.32	0.30	0.33	0.30	0.28	0.29	0.30	**0.27**
$+NFCI^{\theta}_{t,w}$	0.36	0.34	0.30	0.28	0.27	0.28	**0.28**	0.28
90% quantile score	1	3	5	7	9	11	13	15
QR-SPF	0.39	0.39	0.38	0.36	**0.33**	**0.33**	0.33	**0.33**
$+NFCI_{t,w}$	0.41	0.39	0.36	0.35	0.34	0.34	0.34	0.34
$+NFCI^{m}_{t,w}$	0.40	0.39	**0.35**	**0.35**	0.34	0.33	**0.31**	0.33
$+NFCI^{\theta}_{t,w}$	**0.37**	**0.38**	0.36	0.35	0.34	0.34	0.33	0.34
CRPS	1	3	5	7	9	11	13	15
QR-SPF	1.17	1.17	1.14	1.03	1.00	1.00	1.00	1.00
$+NFCI_{t,w}$	**1.17**	1.17	**1.12**	1.02	**0.98**	0.99	**0.98**	0.99
$+NFCI^{m}_{t,w}$	1.17	**1.16**	1.14	1.03	1.00	1.01	0.98	0.99
$+NFCI^{\theta}_{t,w}$	1.18	1.19	1.13	**1.01**	0.99	**0.99**	0.98	**0.99**

Notes: This table reports the average losses of the 10% and the 90% quantile scores in the top and the middle panel. The panel at the bottom shows the average continuously ranked probability score (CRPS) over the out-of-sample evaluation period. The columns indicate the week of a quarter at which the nowcast is formed. Gray areas represent the 90% model confidence set and bold letters are the lowest average losses within each column. All models are estimated on an expanding window. The evaluation period is 1990:Q1 to 2019:Q4, consisting of 120 observations.

Table B.8: Out-of-sample evaluation using the SPF recession probability

	week							
10% quantile score	1	3	5	7	9	11	13	15
SPF-GARCH	0.36	0.35	0.35	0.28	**0.29**	**0.29**	**0.29**	**0.29**
$+x^{\mu}_{t,w} = SPF^{\text{rec}}_{t,w}$	0.35	0.35	0.34	**0.28**	0.29	0.29	0.29	0.29
$+x^{\sigma}_{t,w} = SPF^{\text{rec}}_{t,w}$	0.38	0.38	0.37	0.31	0.30	0.30	0.30	0.30
QR-SPF	0.44	0.44	0.43	0.35	0.31	0.31	0.31	0.31
$+SPF^{\text{rec}}_{t,w}$	**0.33**	**0.33**	**0.33**	0.29	0.29	0.29	0.29	0.29
90% quantile score	1	3	5	7	9	11	13	15
SPF-GARCH	**0.34**	**0.34**	**0.34**	**0.31**	**0.31**	**0.31**	**0.31**	**0.31**
$+x^{\mu}_{t,w} = SPF^{\text{rec}}_{t,w}$	0.34	0.34	0.34	0.32	0.32	0.31	0.31	0.31
$+x^{\sigma}_{t,w} = SPF^{\text{rec}}_{t,w}$	0.37	0.37	0.36	0.34	0.32	0.31	0.32	0.32
QR-SPF	0.39	0.39	0.38	0.36	0.33	0.33	0.33	0.33
$+SPF^{\text{rec}}_{t,w}$	0.38	0.38	0.36	0.34	0.33	0.33	0.33	0.33
CRPS	1	3	5	7	9	11	13	15
SPF-GARCH	**1.12**	1.12	1.09	**0.99**	**0.96**	**0.96**	**0.96**	**0.96**
$+x^{\mu}_{t,w} = SPF^{\text{rec}}_{t,w}$	1.12	**1.11**	**1.09**	0.99	0.97	0.97	0.97	0.97
$+x^{\sigma}_{t,w} = SPF^{\text{rec}}_{t,w}$	1.15	1.15	1.12	1.01	0.97	0.97	0.97	0.97
QR-SPF	1.17	1.17	1.14	1.03	1.00	1.00	1.00	1.00
$+SPF^{\text{rec}}_{t,w}$	1.14	1.14	1.11	1.01	0.98	0.99	0.98	0.98

Notes: This table reports the average losses of the 10% and the 90% quantile scores in the top and the middle panel. The panel at the bottom shows the average continuously ranked probability score (CRPS) over the out-of-sample evaluation period. The columns indicate the week of a quarter at which the nowcast is formed. Gray areas represent the 90% model confidence set and bold letters are the lowest average losses within each column. All models are estimated on an expanding window. The evaluation period is 1990:Q1 to 2019:Q4, consisting of 120 observations.

Table B.9: Rolling window out-of-sample evaluation using NFCI and SPF

	week							
10% quantile score	1	3	5	7	9	11	13	15
AR-GARCH	0.42	0.41	0.39	0.39	0.35	0.35	0.36	0.36
$+x^{\mu}_{t,w} = NFCI_{t,w}$	0.44	0.42	0.36	0.37	0.42	0.37	0.30	0.33
$+x^{\sigma}_{t,w} = NFCI_{t,w}$	0.41	0.38	0.35	0.35	0.33	0.34	0.33	0.34
$+x^{\mu}_{t,w} = SPF_{t,w}$	0.38	0.38	0.36	**0.29**	**0.29**	**0.29**	**0.29**	**0.29**
QR-AR	0.45	0.46	0.43	0.43	0.42	0.42	0.42	0.42
$+NFCI_{t,w}$	**0.32**	**0.33**	**0.33**	0.32	0.31	0.33	0.33	0.33
$+SPF_{t,w}$	0.42	0.42	0.40	0.34	0.31	0.31	0.31	0.31
90% quantile score	1	3	5	7	9	11	13	15
AR-GARCH	0.39	0.39	0.33	0.33	0.33	0.34	0.34	0.34
$+x^{\mu}_{t,w} = NFCI_{t,w}$	0.42	0.41	0.39	0.41	0.48	0.43	0.39	0.41
$+x^{\sigma}_{t,w} = NFCI_{t,w}$	0.44	0.45	0.38	0.38	0.37	0.37	0.37	0.37
$+x^{\mu}_{t,w} = SPF_{t,w}$	**0.34**	**0.34**	**0.32**	**0.30**	**0.29**	**0.30**	**0.30**	**0.30**
QR-AR	0.40	0.40	0.35	0.35	0.36	0.36	0.37	0.37
$+NFCI_{t,w}$	0.42	0.42	0.36	0.35	0.35	0.36	0.37	0.38
$+SPF_{t,w}$	0.37	0.37	0.34	0.32	0.31	0.31	0.32	0.32
CRPS (equal weights)	1	3	5	7	9	11	13	15
AR-GARCH	1.23	1.21	1.13	1.13	1.12	1.12	1.13	1.13
$+x^{\mu}_{t,w} = NFCI_{t,w}$	1.29	1.27	1.16	1.18	1.30	1.21	1.12	1.16
$+x^{\sigma}_{t,w} = NFCI_{t,w}$	1.31	1.32	1.16	1.17	1.16	1.17	1.16	1.18
$+x^{\mu}_{t,w} = SPF_{t,w}$	**1.16**	**1.16**	**1.08**	**0.98**	**0.95**	**0.96**	**0.96**	**0.96**
QR-AR	1.24	1.24	1.15	1.15	1.14	1.14	1.15	1.15
$+NFCI_{t,w}$	1.22	1.25	1.17	1.17	1.16	1.17	1.17	1.17
$+SPF_{t,w}$	1.19	1.19	1.10	1.00	0.97	0.98	0.98	0.98

Notes: This table reports the average losses of the 10% and the 90% quantile scores in the top and the middle panel. The panel at the bottom shows the average continuously ranked probability score (CRPS) over the out-of-sample evaluation period. The columns indicate the week of a quarter at which the nowcast is formed. Gray areas represent the 90% model confidence set and bold letters are the lowest average losses within each column. All models are estimated on a rolling window with 85 observations. The evaluation period is 1990:Q1 to 2019:Q4, consisting of 120 observations.