
# Time-Varying Macroeconomic Forecast Uncertainty: Sweden and the Covid-19 Pandemic*

Sebastian Ankargren†

September, 2020

* I am grateful for comments by Erika Färnstrand Damsgaard, Urban Hansson Brusewitz, Ylva Hedén Westerdahl, Birol Kanik, Markus Sigonius and Pär Österholm. I have benefited from code made available by Clark, McCracken and Mertens (2020), but any errors are mine. The opinions expressed in this article are the sole responsibility of the author and do not necessarily reflect the views of the National Institute of Economic Research.

† National Institute of Economic Research, Fleminggatan 7, Box 12090, 102 23 Stockholm, Sweden. E-mail: sebastian.ankargren@konj.se Phone: +46 453 59 55.

WORKING PAPER NO 153, SEPTEMBER 2020
PUBLISHED BY THE NATIONAL INSTITUTE OF ECONOMIC RESEARCH (NIER)

**NIER** prepares analyses and forecasts of the Swedish and international economy and conducts related research. **NIER** is accountable to the Ministry of Finance, and like other government agencies, has an independent status and is responsible for the assessments that it publishes.

The **Working Paper** series consists of publications of research reports and other detailed analyses. The reports may concern macroeconomic issues related to the forecasts of the institute, research in environmental economics, or problems of economic and statistical methods.

All publications can be downloaded from the NIER website.

NATIONAL INSTITUTE OF ECONOMIC RESEARCH, FLEMINGGATAN 7, SE-102 23 STOCKHOLM
+46 8 453 59 00, REGISTRATOR@KONJ.SE, WWW.KONJ.SE/EN

ISSN 1650-996X

# Abstract

I construct time-varying uncertainty around forecasts for the Swedish economy published by the National Institute of Economic Research. Using several statistics for evaluating interval and density forecasts of GDP growth and unemployment in 2014--2019, I find that time-varying uncertainty has performed better than standard measures of unconditional uncertainty derived from normal distributions with constant variance. Compared to t distributions, however, the results are not as clear, but speak in favor of time-varying uncertainty. Using Bayesian model comparisons, the results show that time-varying uncertainty has been particularly important for GDP growth, whereas correlations between revisions have played a bigger role for the unemployment rate. I study forecast uncertainty during the initial phase of the Covid-19 episode and find that the rise in forecast uncertainty was unprecedented.



**JEL classification code**: C11, C32, C53, C54

Keywords: Stochastic volatility, Density forecast, Interval forecast, Bayesian

# Sammanfattning

Jag beräknar tidsvarierade osäkerhet kring prognoser gjorda av Konjunkturinstitutet. Med hjälp av flera olika utvärderingsmått för att utvärdera intervall- och täthetsprognoser för BNP-tillväxt och arbetslöshet under 2014–2019 finner jag att tidsvarierande osäkerhet har varit mer lämpligt än vanligen använda mått baserade på normalfördelning med konstant varians. Jämfört med *t*-fördelning är dock resultaten inte lika entydiga, men talar till den tidsvarierande osäkerhetens fördel. En bayesiansk modelljämförelse visar att tidsvarierande osäkerhet har varit särskilt viktigt för BNP-tillväxten, medan korrelation mellan prognosrevideringar spelat en större roll för arbetslösheten. Jag studerar prognososäkerheten under den initiala fasen av covid-19 och kommer fram till att uppgången i prognososäkerhet saknade motstycke.

# Contents

# 1 Introduction

Forecasts are uncertain, and uncertainty about the future is the reason why forecasting is conducted. Point forecasts have historically been the focal point in policy forecasting, but the Bank of England and Sveriges Riksbank pioneered the use of interval forecasts, known as fan charts, in the late 1990s (Britton et al., 1998; Blix and Sellin, 1998). Interval forecasts put less emphasis on a single number and clearly indicate that outcomes within an entire range are plausible. Fan charts are today presented by many central banks and other professional forecasters to give an indication of the level of uncertainty that their forecasts are associated with. For a summary of institutions and central banks that present fan charts, see Ohnsorge et al. (2016).

This paper employs the method proposed by Clark et al. (2020) to estimate time-varying levels of forecast uncertainty. The method relies on stochastic volatility, which is ubiquitous in today's macroeconometric literature, to cope with changes in forecast uncertainty. In contrast to more traditional fan chart approaches, Clark et al. (2020) only apply the machinery of stochastic volatilities to the forecast errors indirectly. The method operates on the most recent nowcasting error, as well as revisions of forecasts for future outcomes. Since times associated with heightened forecast uncertainty are unexpected, these often lead to large forecast revisions between successive forecasting rounds. These large revisions are picked up as a signal of increased forecast uncertainty, which directly spills over to predictive distributions and intervals. Conversely, when the economy offers few surprises and revisions are small, the method allows for forecast uncertainty to decrease. The decrease leads to sharper predictive distributions and intervals that are more informative.

I use the time-varying model to characterize forecast uncertainty surrounding quarterly forecasts made by the National Institute of Economic Research, one of Sweden's most prominent forecasters. The variables that I study are GDP growth and the unemployment rate. My results show that for the evaluation period 2014–2019, time-varying uncertainty improves density and interval forecasts compared to normal distributions with constant variance. I

also make comparisons with a thick-tailed alternative based on a $t$ distribution. Time-varying uncertainty for the most part dominates, but not always. In a Bayesian model comparison framework, the results show that time variation has been more important than contemporaneous correlation between revisions for characterizing GDP growth forecast uncertainty.[1] I obtain the opposite result for the unemployment rate. Finally, I show that forecast uncertainty for both variables increased substantially because of the Covid-19 pandemic.

Uncertainty can be a powerful tool when forecasters communicate projections.[2] Clements (2018) compares the conditional predictive densities provided by the Survey of Professional Forecasters to unconditional density forecasts. In similar work, Knüppel and Schultefrankenfeld (2019) evaluate inflation uncertainty forecasts from several central banks. Both Clements (2018) and Knüppel and Schultefrankenfeld (2019) find that forecasters tend to be underconfident at short horizons and overconfident at longer horizons. This result supports previous findings in the literature that uncertainty forecasts often overstate the level of uncertainty, see Clements (2004); Wallis (2004). Knüppel (2014) accounts for correlation between forecast errors at different horizons and provide evidence that doing so can have a large impact on estimated forecast uncertainty. Similarly, Knüppel (2018) highlights the sensitivity to small samples that occurs when the forecast horizon is increased, and no or few historical forecast errors are available. Exploiting correlation between forecast errors can mitigate estimation errors considerably.

Clark et al. (2020) develop their method based on nowcasting errors and forecast revisions. The more standard way is basing descriptions of forecast uncertainty on historical forecast errors in conjunction with a model that prescribes a variance that is constant over time, see for example Sveriges Riksbank (2007); Reifschneider and Tulip (2017). Intervals with constant variance give a description of the average, or unconditional, uncertainty of a

[1] Revisions that originate in the same period are allowed to be correlated, but serial correlation between revisions from different points in time is not permitted. The latter would imply inefficient forecasts, but throughout I assume forecasts to be efficient.

[2] In addition to uncertainty, many forecasters also describe risks and whether they are balanced. However, Knüppel and Schultefrankenfeld (2012) show that the statistical value of the quantitative risk assessments published by Bank of England and Sveriges Riksbank is small.

forecaster's projections. However, a forecaster wishing to communicate a quantitative perception of the *current* level of forecast uncertainty may be dissatisfied using *average* levels of uncertainty as a proxy.[3] The reason is not only that the intervals—or, more generally, the dispersions in the full predictive distributions—are perceived to be too narrow in times of turbulence, but also that they are too wide in more predictable and certain times. Thus, uncertainty intervals and predictive distributions that accurately can give a characterization of the level of forecast uncertainty in real-time can be highly useful tools that complement point forecasts and qualitative descriptions of uncertainties. I give some evidence in that direction in this work.

# 2 Modeling Time-Varying Uncertainty of Forecast Errors

The estimation approach I use to describe time-varying uncertainties of forecast errors is based on the techniques developed by Clark et al. (2020). The time index $t$ refers to time in quarters, and I operate at the end of the quarter so that the outcome for $t-1$ is known at time $t$. Let $\hat{y}_{t|t-h}$ denote a forecast for time $t$ produced at time $t-h$, where $h = 0, \ldots, H$. Due to publication delays, $y_t$ is not observed at time $t$ and $\hat{y}_{t|t}$ denotes its forecast, that is the nowcast. Let also $e_{t|t-h} = y_t - \hat{y}_{t|t-h}$ denote the forecast error for the $h$ horizon forecast. Finally, $\mu_{t|t-h} = \hat{y}_{t|t-h} - \hat{y}_{t|t-h-1}$ represents the forecast revision for time $t$ between forecasts made at the successive points in time $t-h-1$ and $t-h$.

The main idea behind the method is that the forecast error $e_{t|t-h}$ can be decomposed

[3] One way to communicate the current level of uncertainty is to relate it to the historical level and whether the current situation is perceived to be more, less or equally uncertain as the past. Reifschneider and Tulip (2017) discuss how the members of the Federal Open Market Committee use such an approach.

into a nowcasting error and the sum of forecast revisions as

$$e_{t|t-h} = e_{t|t} + \sum_{i=1}^{h} \mu_{t|t-h+i}, \quad h = 1, \ldots, H. \tag{1}$$

I maintain an assumption that the forecasts are optimal in the sense of Mincer and Zarnowitz (1969). Consequently, the sequence $\{\mu_{t|t-h}\}_{h=0}^{H}$ is a martingale difference sequence. The implication is that the terms in (1) are uncorrelated.[4]

## 2.1 Econometric Model

I here provide a brief overview of the method I use and relegate some of the details to Appendix A. The objective of the econometric analysis is to obtain a useful description of the posterior distribution $p(e_{T|T}, \ldots, e_{T+H|T}|\eta)$.[5] The methodology achieves this goal by postulating a model for the terms present in (1) that is itself subsequently exploited in order to target the posterior distribution of interest. Let first

$$\eta_t = \begin{pmatrix} e_{t-1|t-1} \\ \mu_{t|t} \\ \vdots \\ \mu_{t+H-1|t} \end{pmatrix}.$$

The data vector $\eta_t$ is assumed to have mean zero so that forecasts are unbiased. While the sequence of random variables $\{\eta_t\}$ is uncorrelated over time, it is not independent. The methodology prescribes a relatively standard multivariate stochastic volatility model, see, for instance, the seminal work by Primiceri (2005); Cogley and Sargent (2005), that is specified

[4]Optimal is here taken to mean that $E(\hat{y}_{t|t-h+s}|\Omega_{t-h+q}) = \hat{y}_{t|t-h+q}$, $s > q \geq 0$ and $\Omega_t$ is the information as of time $t$. The assumption directly implies that the expectation of a future revision $\mu_{t|t-h+s}$ is zero given a previous revision $\mu_{t|t-h+q}$, since $E(\mu_{t|t-h+s}|\Omega_{t-h+q}) = E(\hat{y}_{t|t-h+s}|\Omega_{t-h+q}) - \hat{y}_{t|t-h+q} = 0$.

[5]The predictive distribution for the variable of interest is obtained by centering the distribution around the forecasts $\hat{y}_{T|T}, \ldots, \hat{y}_{T+H|T}$.

as

$$\eta_t = A\Lambda_t^{0.5}\epsilon_t \tag{2}$$

$$\epsilon_t \sim \text{iid}\, \text{N}(0, I_{H+1})$$

$$A = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ a_{2,1} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{H+1,1} & a_{H+1,2} & \cdots & 1 \end{pmatrix}$$

$$\Lambda_t = \text{diag}(\exp\{\lambda_{1,t}\}, \ldots, \exp\{\lambda_{H+1,t}\}).$$

Clark et al. (2020) use separate random walk processes for the log-volatilities $\lambda_{i,t}$ in their main specification, connected through a possibly non-diagonal covariance matrix for the log-volatility innovations. Because of a much shorter data set, I confine myself in this analysis to their alternative specification based on a single common factor.[6] Let the evolution of the common volatility factor be given by

$$\lambda_{0,t} = \lambda_{0,t-1} + \nu_t, \quad \nu_t \sim \text{N}(0,1), \quad \lambda_{0,0} = 0. \tag{3}$$

The $H+1$ equation-specific log-volatilities are assumed to relate linearly to the common factor through

$$\lambda_{i,t} = \lambda_{i,0} + \beta_i \lambda_{0,t}.$$

The role of $\lambda_{i,0}$ is to accommodate varying levels of baseline volatility for the elements of $\eta_t$, whereas $\beta_i$ allows for various degrees of pass-through of the common factor on the volatilities of the elements of $\eta$. The lower triangular matrix $A$, on the other hand, introduces correlation

[6] However, instead of fixing one of the loadings and letting the variance of the factor innovation be free, I let all loadings be freely estimated and set the variance of the factor innovation to one.

among the elements of $\eta_t$. I investigate the relative importance of time variation ($\beta_i \neq 0$) and correlation ($a_{i,j} \neq 0$) in Section 3.3.

# 3 Statistics for Evaluating Descriptions of Uncertainty

I compare the interval and density forecasts produced using the time-varying method with those made by alternative methods to evaluate the potential appropriateness of describing forecast uncertainty in a time-varying manner. A number of statistics can be used to this end. In this paper, I use empirical coverage rates and interval scores to assess the relative merits of the methods' interval forecasts. Furthermore, the quality of the respective density forecasts are compared by means of the log predictive and continuous ranked probability scores. Finally, the model with time-varying uncertainty is compared to restricted alternatives using marginal likelihoods.

## 3.1 Interval Forecasts

### 3.1.1 Empirical Coverage Rates

A well-calibrated interval forecast should have an empirical coverage rate that is close to its nominal $100(1-\alpha)\%$ rate. I therefore compare the methods' empirical coverage rates and their closeness to their target levels by computing a sequence of 'hits' and 'misses', that is,

$$I_t^{\alpha}(h) = \begin{cases} 1 & \text{if} \quad y_t \in [l_{t|t-h}^{\alpha},\, u_{t|t-h}^{\alpha}], \\ 0 & \text{otherwise,} \end{cases}$$

where $l_{t|t-h}^{\alpha}$ and $u_{t|t-h}^{\alpha}$ are the lower and upper ends of the interval forecast for horizon $h$ with nominal coverage rate $100(1-\alpha)\%$. The empirical coverage rate is the average of the hit sequence, $\frac{1}{T_h}\sum_{t=1}^{T_h} I_t^{\alpha}(h)$, where $T_h$ is the number of forecasts for horizon $h$ in the evaluation period.

### 3.1.2 Interval Scores

In general, a good forecast balances calibration and sharpness appropriately (Gneiting et al., 2007).[7] For example, an interval forecaster that with probability $1-\alpha$ selects the forecast $(\hat{y}_{t|t-h} - c_1, \hat{y}_{t|t-h} + c_1)$, with $c_1$ being a very large number, and with probability $\alpha$ instead utilizes the interval forecast $(\hat{y}_{t|t-h} - c_2, \hat{y}_{t|t-h} + c_2)$, $c_2$ being infinitesimally small, will have an empirical coverage rate that on average is close to the nominal target. However, the forecast is inarguably useless because of the forecaster's sole focus on calibration without any attention to sharpness.

Gneiting and Raftery (2007) propose an alternative scoring rule for evaluating interval forecasts that takes into account both calibration and sharpness. The proposed interval score is

$$S^{\text{int}}_{\alpha,h} = \frac{1}{T_h}\sum_{t=1}^{T_h} S^{\text{int}}_{\alpha}(l^{\alpha}_{t|t-h}, u^{\alpha}_{t|t-h}; y_t)$$

$$S^{\text{int}}_{\alpha}(u, l; y) = (u-l) + \frac{2}{\alpha}(l-y)\mathbb{1}\{y < l\} + \frac{2}{\alpha}(y-u)\mathbb{1}\{y > u\}.$$

The score gives a more accurate picture of the quality of an interval forecast than the empirical coverage rate alone does in that it also acknowledges the informativeness of an interval. If two intervals both contain the outcome, the score rewards the shorter. Similarly, if both intervals do not contain the outcome, an additional penalty is incurred that is proportional to the size of the deviation of the outcome from the endpoint.

## 3.2 Density Forecasts

To evaluate density forecasts, I employ the commonly used scoring rules log score (LS) and continuous ranked probability score (CRPS). Both scores are negatively oriented and defined

[7] Calibration is the statistical consistency between forecasts and realized values. Sharpness refers to the concentration in the predictive distributions. The goal of interval or density forecasts is thus to maximize sharpness subject to the constraint that the forecasts are well calibrated.

as

$$\begin{aligned}\text{LS}_{t,h} &= -\log[f_{t|t-h}(e_{t|t-h})] \\ \text{CRPS}_{t,h} &= \int_{-\infty}^{\infty} \left[F_{t|t-h}(z) - \mathbb{1}(e_{t|t-h} < z)\right] \mathrm{d}z,\end{aligned}$$

where $f_{t|t-h}(e_{t|t-h}) = f(e_{t|t-h}|\eta_{t-h}, ..., \eta_1)$ is the predictive density for time $t$ formed at $t-h$ evaluated at the realized forecast error $e_{t|t-h}$. The corresponding cumulative distribution function is $F_{t|t-h}(e_{t|t-h}) = \int_{-\infty}^{e_{t|t-h}} f_{t|t-h}(z)\,\mathrm{d}z$. The predictive distributions are not available in closed form, but estimates can be computed based on the draws of parameters and volatilities using a mixture-of-parameters formulation that Krüger et al. (2020) show produces accurate estimates.

## 3.3 Bayesian Model Comparison via the Marginal Likelihood

To compare models in a Bayesian framework, the most common statistic is the marginal likelihood

$$p(\eta) = \int p(\eta|\Theta)p(\Theta)\,\mathrm{d}\Theta,$$

where $\Theta$ contains the unknown parameters and volatilities to be integrated out. Assuming models with equal prior probabilities, the posterior probability of the model is proportional to the marginal likelihood. I estimate the marginal likelihood using bridge sampling as proposed by Meng and Wong (1996), see also Gronau et al. (2017).[8]

[8]Bridge sampling relies on a proposal distribution for which I use a multivariate $t$ distribution with 10 degrees of freedom. Its mean and scale matrix are set equal to the sample mean and covariance of the first half of the post-burn-in MCMC draws, and the remaining draws are used in the evaluation steps of the bridge sampling iterations following the suggestion by Overstall and Forster (2010). Changing the marginal likelihood estimator to the modified harmonic mean proposed by Gelfand and Dey (1994) with a truncated normal distribution for tuning function as suggested by Geweke (1999) yields results that are very similar.

# 4 Data

The data consist of quarterly forecasts for Sweden made by the National Institute of Economic Research (NIER). The NIER publishes reports on the Swedish economy in March, June, October, and December.[9] The reports contain forecasts of a wide range of variables; in this work, I focus on GDP growth and the unemployment rate as these generally receive attention from media and the general public. The data range from 2002Q3 to 2020Q2 as the NIER began publishing quarterly forecasts in the third quarter of 2002. The forecast horizon varies between reports, but all reports include forecasts for at least five quarters ahead. I use data through 2019Q4 for evaluations in Section 5 to exclude the Covid-19 pandemic, and study this episode separately in Section 6. I use the first available vintages for computing nowcasting errors for GDP growth. The unemployment rate is not revised, so the choice of vintage is immaterial for the results.

# 5 Results

In this section, I use the statistics described in Section 3 to evaluate the interval and density forecasts produced by the model for time-varying uncertainty. The evaluation period is 2014Q1–2019Q4 and is pseudo-real-time; I recursively add a new quarter to the dataset, re-estimate the model, and compute uncertainty forecasts. I let the maximum forecast horizon be $H = 5$, which is the largest choice permitted by the data, so that I obtain predictive distributions for the nowcast $y_{T|T}$ through the forecast five quarters ahead, $y_{T+5|T}$. One possible explanation for the potential success of time-varying uncertainties is that, unconditionally, the predictive distributions are thick-tailed. I compare the model with time-varying uncertainty to predictive distributions and interval forecasts derived from normal and Student's $t$ distributions, respectively, to give an indication of whether this conjecture has any explanatory power. Both approaches take forecast errors at different horizons to be inde-

[9]The October publication was first released in 2017. I use forecasts from reports published in August prior to its commencement.

pendent. I estimate their unknown parameters—variance, and scale and degrees of freedom, respectively—recursively using maximum likelihood for the out-of-sample evaluation.

## 5.1 GDP Growth

Figure 1 gives a first glance of the level of time variation of GDP forecast uncertainty. Each panel presents the absolute value of the elements of $\eta_t$ together with the full-sample estimates of volatilities $\lambda_{i,t}^{0.5}$ as well as real-time estimates. In the absence of correlation across elements of $\eta_t$, that is if $A = I_{H+1}$, $\lambda_{i,t}^{0.5}$ is equal to the standard deviation at time $t$ of the $i$th component of $\eta_t$. The figure shows an increase in the volatilities around 2008–2009 with large nowcast errors through 2015. The peak of the volatility is estimated to be in 2011 in the midst of a period of large nowcast errors and substantial revisions of forecasts at longer horizons. Revisions made in the first half of 2020 clearly stand out and display magnitudes not previously seen.

### 5.1.1 Interval Forecasts

To assess the calibration of the intervals, I display empirical coverage rates in Figure 2. The figure clearly shows that the intervals based on the normal distribution are underconfident, a result that is in line with previous literature (Clements, 2004; Knüppel and Schultefrankenfeld, 2019). Intervals based on Student's $t$ do much better, with empirical coverage rates that are never further from the nominal coverage rate than the normal distribution's. Intervals derived from time-varying predictive densities are even closer to the nominal levels, and are never worse than intervals based on normality or Student's $t$. Overall, all three types of intervals are too conservative. Intervals with nominal coverage rate 90% have with only one exception included every outcome.

While coverage rates assesses calibration in isolation, it is important to consider the crucial balance between calibration and sharpness. To account for both of these concepts, Figure 3 shows the intervals evaluated using the interval score (Gneiting and Raftery, 2007).

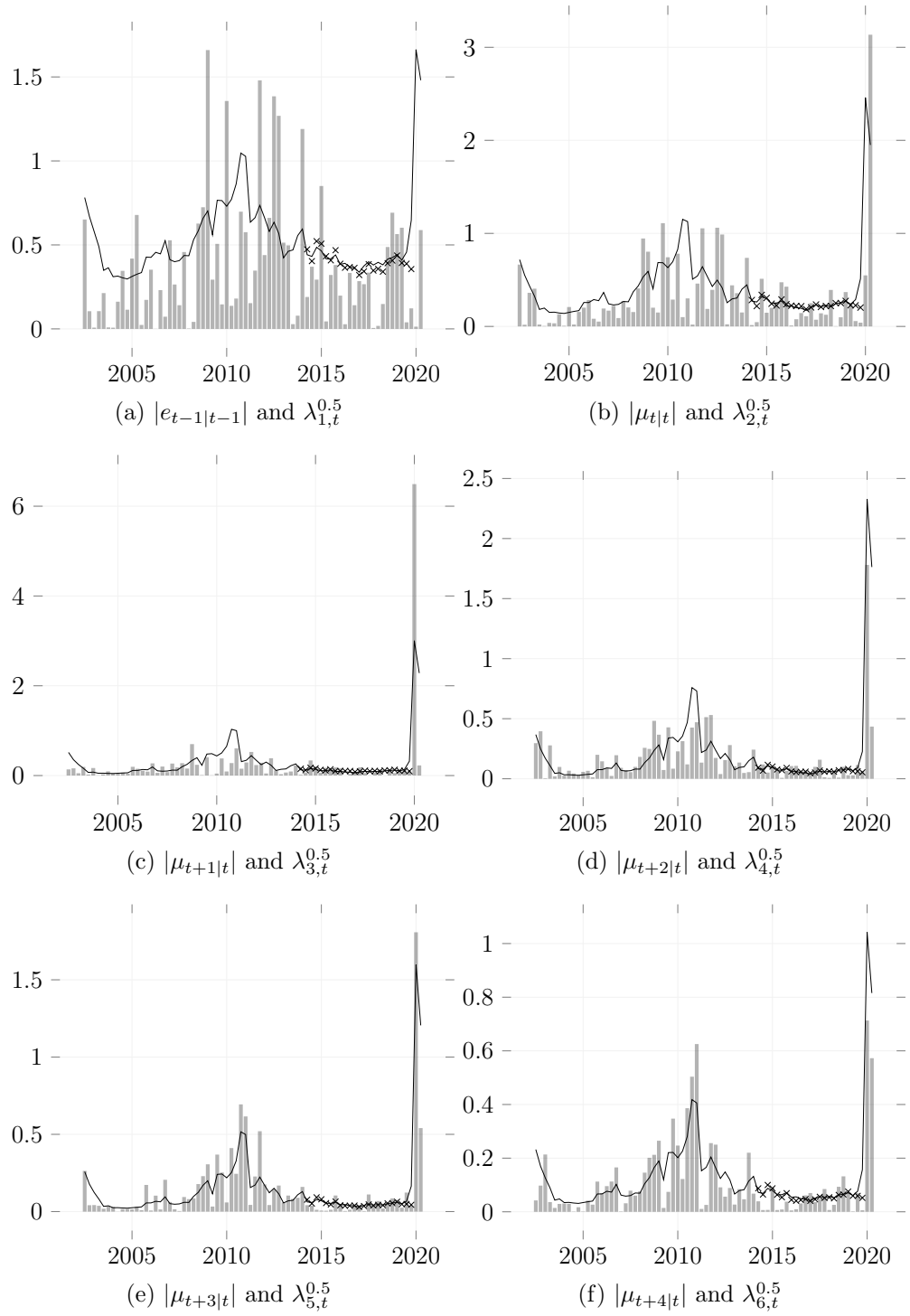
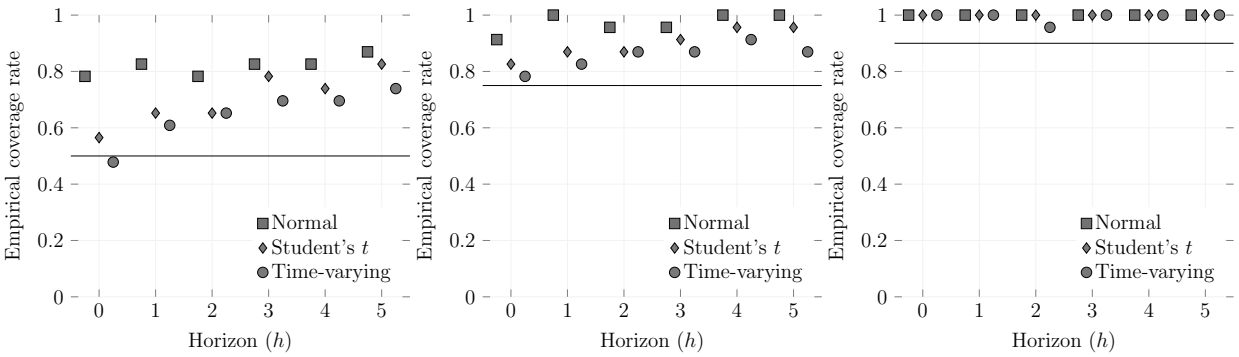
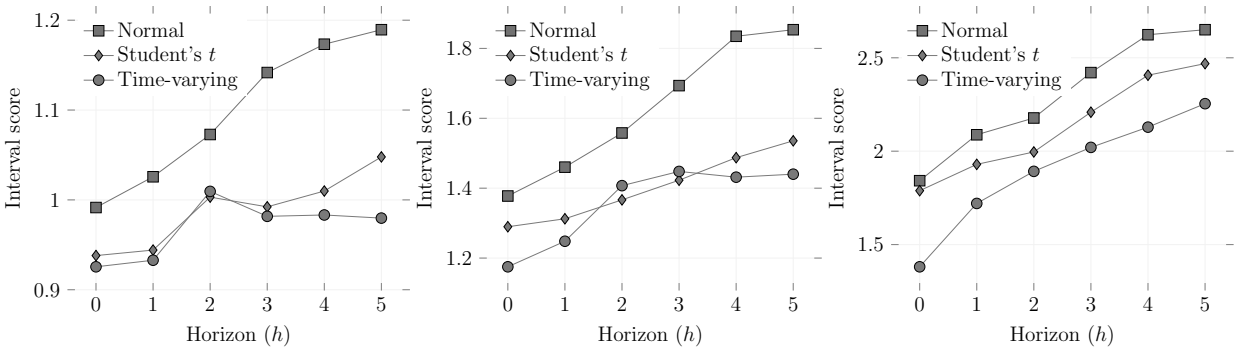



Figure 1: Absolute values of components of $\eta_t$ (bars), and volatilities (lines) for GDP growth. Crosses denote real-time estimates of $\lambda^{0.5}_{i,t}$, whereas lines display full-sample (smoothed) estimates.



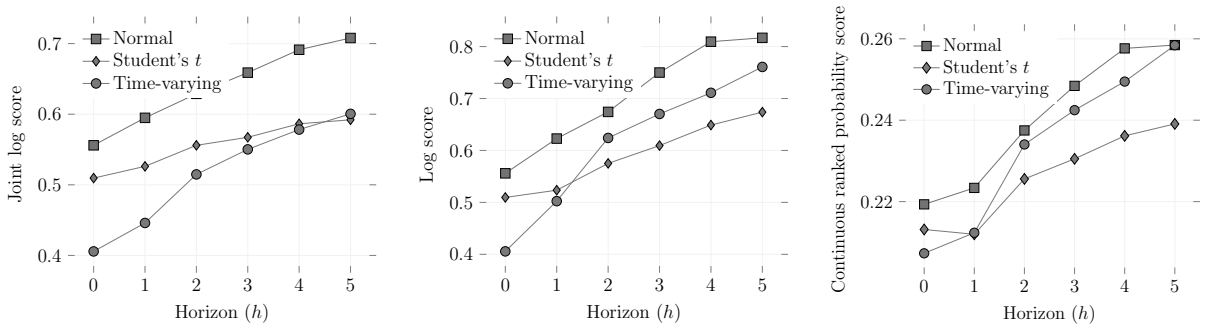

(a) 50% intervals (b) 75% intervals (c) 90% intervals

Figure 2: Empirical coverage rates, GDP growth




(a) 50% intervals (b) 75% intervals (c) 90% intervals

Figure 3: Interval scores, GDP growth




(a) Log score, jointly from horizon 0 to $h$. The joint score is divided by $h+1$ for comparative purposes.

(b) Log score, univariate

(c) Continuous ranked probability score

Figure 4: Density scores, GDP growth

The interval scores in Figure 3 generally favor time-varying uncertainty. It is slightly superior to the Student's $t$ intervals for 50% intervals at the initial two horizons, and with an increased differential at the longest horizons. Time-varying uncertainty is somewhat inferior to Student's $t$ for horizons two and three for 75% intervals, but better at all horizons in the 90% case. The difference to the normality-based intervals is relatively much larger, especially for 50% intervals. The results indicate that time-varying uncertainty and Student's $t$ balance calibration and sharpness better.

### 5.1.2 Density Forecasts

The intervals in the preceding evaluation are all derived from full predictive distributions. To analyze the performance of the estimated predictive distributions, I present the two popular scoring rules log score and continuous ranked probability score in Figure 4.

The figure shows that the log score computed jointly over forecast horizons ranks time-varying forecast uncertainty first. However, in contrast to the interval score, both the univariate log score and the continuous ranked probability score favor Student's $t$ for horizons beyond the first two. For the initial nowcast horizon, time-varying uncertainty is preferred by both scoring rules. These results indicate that accounting for dependencies across horizons improves the joint predictive distribution. Compared to normality-based predictive

distributions, time-varying uncertainty is superior also for the univariate horizon-specific scores.

## 5.2 Unemployment Rate

Figure 5 displays the nowcast errors and forecast revisions for the unemployment rate together with volatility estimates. Compared to Figure 1, the volatility estimates for the unemployment rate to a lesser degree track the size of the revisions at longer horizons. The discrepancy suggests that $\lambda_{i,t}^{0.5}$ may not approximate the standard deviation of the forecast revisions at longer horizons, indicating that the matrix $A$ is likely not diagonal. The volatilities show two historical peaks in 2006 and 2008 during periods of large nowcast errors and major forecast revisions. Volatilities thereafter decreased until late 2019 when large revisions inflated them anew, and in the beginning of 2020 the Covid-19 pandemic led to further increases.

### 5.2.1 Interval Forecasts

Figure 6 reveals that intervals leveraging time-varying uncertainty have, for the most part, been preferable according to the empirical coverage rates. Their empirical coverage rates are closer to the nominal rates with only two exceptions. Both exceptions pertain to overconfidence, with coverage rates that are somewhat too low relative the nominal rates. Student's $t$ produces intervals that have better coverage properties than normality-based intervals, but the differences are smaller for unemployment than for GDP growth (Figure 2).

Figure 7 presents the interval score, that balances calibration and sharpness, for intervals for the unemployment rate. All three intervals obtain scores that are indistinguishable at the nowcasting and one-quarter horizons, but diverge for subsequent horizons. Time-varying uncertainty leads to the best score for all horizons and coverage probabilities. Student's $t$ fares comparatively better than the normal distribution, but the difference is small for the 90% intervals.

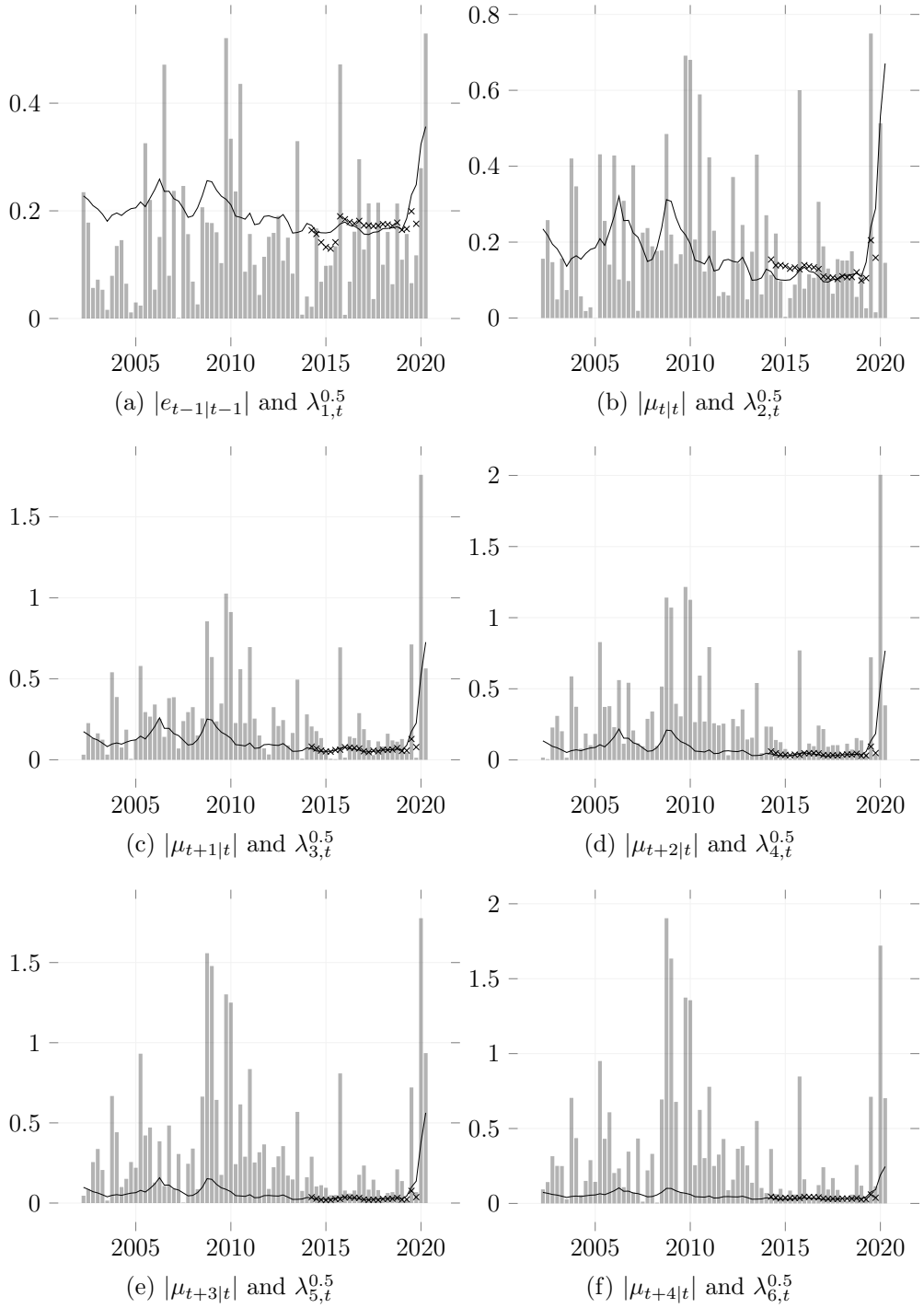
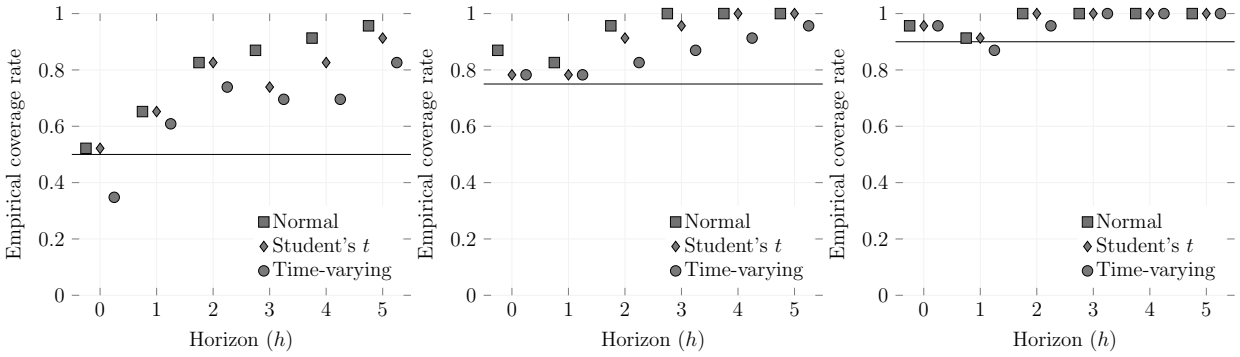
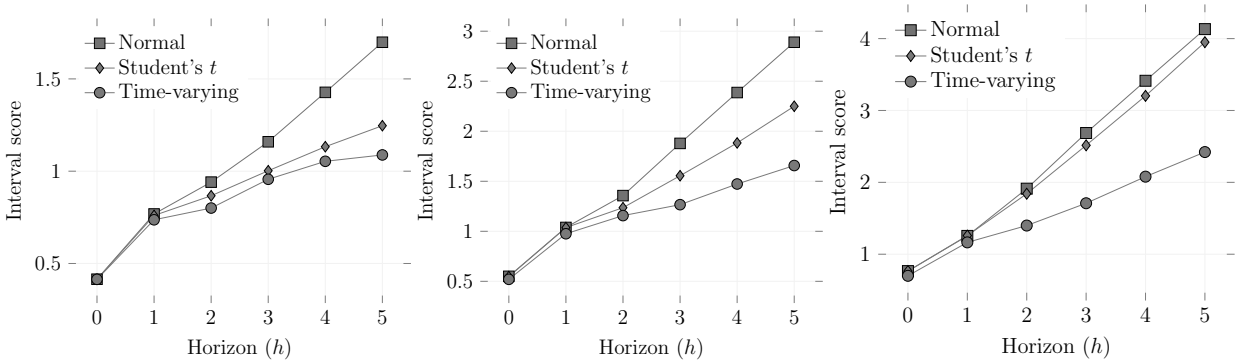



Figure 5: Absolute values of components of $\eta_t$ (bars), and volatilities (lines) for Unemployment rate. Crosses denote real-time estimates of $\lambda^{0.5}_{i,t}$, whereas lines display full-sample (smoothed) estimates.



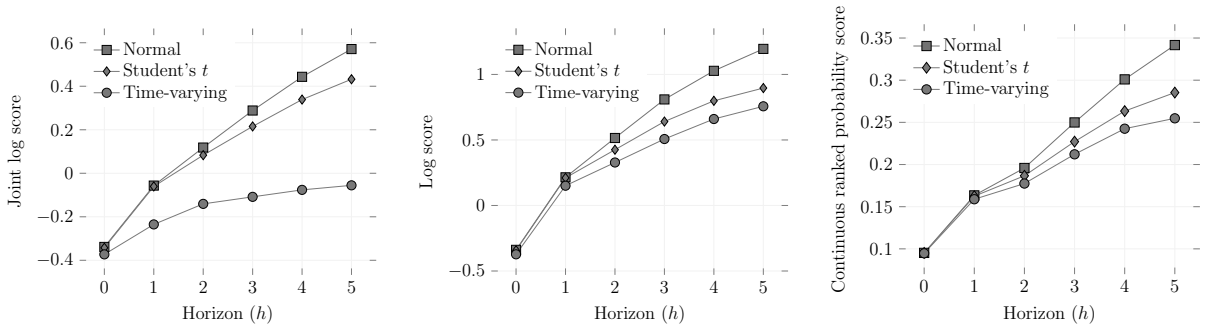

(a) 50% intervals (b) 75% intervals (c) 90% intervals

Figure 6: Empirical coverage rates, Unemployment rate




(a) 50% intervals (b) 75% intervals (c) 90% intervals

Figure 7: Interval scores, Unemployment rate




(a) Log score, jointly from horizon 0 to $h$. The joint score is divided by $h+1$ for comparative purposes.

(b) Log score, univariate

(c) Continuous ranked probability score

Figure 8: Density scores, Unemployment rate

### 5.2.2 Density Forecasts

The log and continuous ranked probability scores for the unemployment rate forecasts are displayed in Figure 8. In contrast to the results for GDP growth (Figure 4), time-varying uncertainty uniformly leads to better scores. The difference between the three methods is minor for the first two horizons in terms of the univariate scores, but for longer horizons the scores start to diverge. The joint log score shows a greater difference between time-varying uncertainty, and normal and Student's $t$ than what the univariate scores show. This increased differential suggests that correlations between forecast errors enhances the joint predictive distributions.

## 5.3 Marginal Likelihood

The results for GDP growth and unemployment in the preceding sections show that time-varying forecast uncertainty leads to statistical improvements for a variety of statistics. In this section, I compare the baseline model for time-varying uncertainty described in Section 2 to restricted alternatives.

The model from Section 2 includes two key features that the constructions based on normal and Student's $t$ distributions lack: correlation between horizons, and time-varying

Table 1: Log marginal likelihoods

| | Full model | No correlation | No time variation |
|---|---|---|---|
| *GDP growth* | | | |
| | 105.0 | 82.0 | -9.2 |
| | [105.0, 105.1] | [82.0, 82.0] | [-9.2, -9.2] |
| *Unemployment rate* | | | |
| | 355.6 | -62.7 | 305.4 |
| | [355.6, 355.7] | [-62.7, -62.7] | [305.4, 305.4] |

*Note:* The 'no correlation' alternative is the model equipped with the restriction that $A = I_{H+1}$, whereas the 'no time variation' alternative has $\beta_i = 0$. The reported range, in brackets, is the minimum and maximum log marginal likelihoods obtained from 10 independent runs. The models are estimated using 100,000 draws in the MCMC algorithm, where the first 50,000 are discarded. Marginal likelihoods are estimated using bridge sampling with 100,000 importance sampling draws from a multivariate $t$ distribution with 10 degrees of freedom and mean and scale equal to the sample mean and covariance based on the first 25,000 retained MCMC draws. The estimation sample is 2002Q3–2019Q4.

variances. Table 1 presents marginal likelihoods for the baseline model, and two restricted variants that shut off correlations and time variation, respectively. The former model restricts $A = I_{H+1}$ so the nowcast error and revisions are conditionally independent, and the latter restricts $\beta_i = 0$ to remove time-varying variances. Priors are kept unchanged elsewhere.

The marginal likelihoods in Table 1 show that the model without restrictions receives the largest marginal likelihood for both GDP and unemployment. The table shows that the relative importance of correlation and time variation is different for the two variables. Restricting the correlation, that is $A$, is less consequential for GDP growth than restricting the time-varying variances to be constant as indicated by the higher marginal likelihood for the 'No correlation' alternative. The result is the opposite for the unemployment rate; restricting correlations reduces the marginal likelihood more than restricting time variation. These results state more formally what Figure 4 and 8 also suggest: For the unemployment rate, the slope of the joint log score curve is notably flatter than the slope of the univariate

log score curve, whereas for GDP growth the slopes are almost the same. This indicates that the joint log score for GDP growth is close to a joint log score computed assuming independence across horizons. The joint log score for the unemployment rate is much further from the average of the univariate scores, implying that a sizable part of the joint log score improvement can be attributed to correlations between forecast errors. Thus, the results of Bayesian model comparisons and density forecast evaluations both point towards a larger importance of correlation between horizons for the unemployment rate. More generally, the results make it clear that forecast errors for different variables may have different properties, and thus potentially also require different treatments.

# 6 Forecast Uncertainty During Covid-19

The rapid spread of the Covid-19 pandemic caused an economic situation that was widely described as extraordinarily uncertain. The NIER published regular forecasts on December 19, 2019, April 1, 2020, and June 17, 2020. Because of the quickly changing circumstances, an updated forecast was released April 29, 2020. The peak of the downturn was expected to hit during the second quarter of 2020, and the forecasts for quarterly GDP growth were revised from 0.2 (December) to -6.3 (April 1) to -11.2 (April 29) to -9.5 (June 17). For the unemployment rate, the forecast changed from 7.2 (December) to 8.9 (April 1) to 11.5 (April 29) to 8.8 (June 17).

The forecasting environment was undoubtedly perceived to be incredibly uncertain. In the Swedish Economy Report published on April 1 (National Institute of Economic Research, 2020), the NIER wrote:

> The global economy is expected to contract in 2020, and Swedish GDP falls by more than 3 per cent this year in the NIER's base scenario. However, there is extreme uncertainty about future developments.

As another sign of the heightened uncertainty, Sveriges Riksbank decided to publish two

scenarios instead of a single forecast in its Monetary Policy Report published on April 28 (Sveriges Riksbank, 2020), writing:

> The great uncertainty over the course of the pandemic means that the Riksbank, in this report, has chosen to discuss developments on the basis of two different scenarios, rather than a single specific forecast.

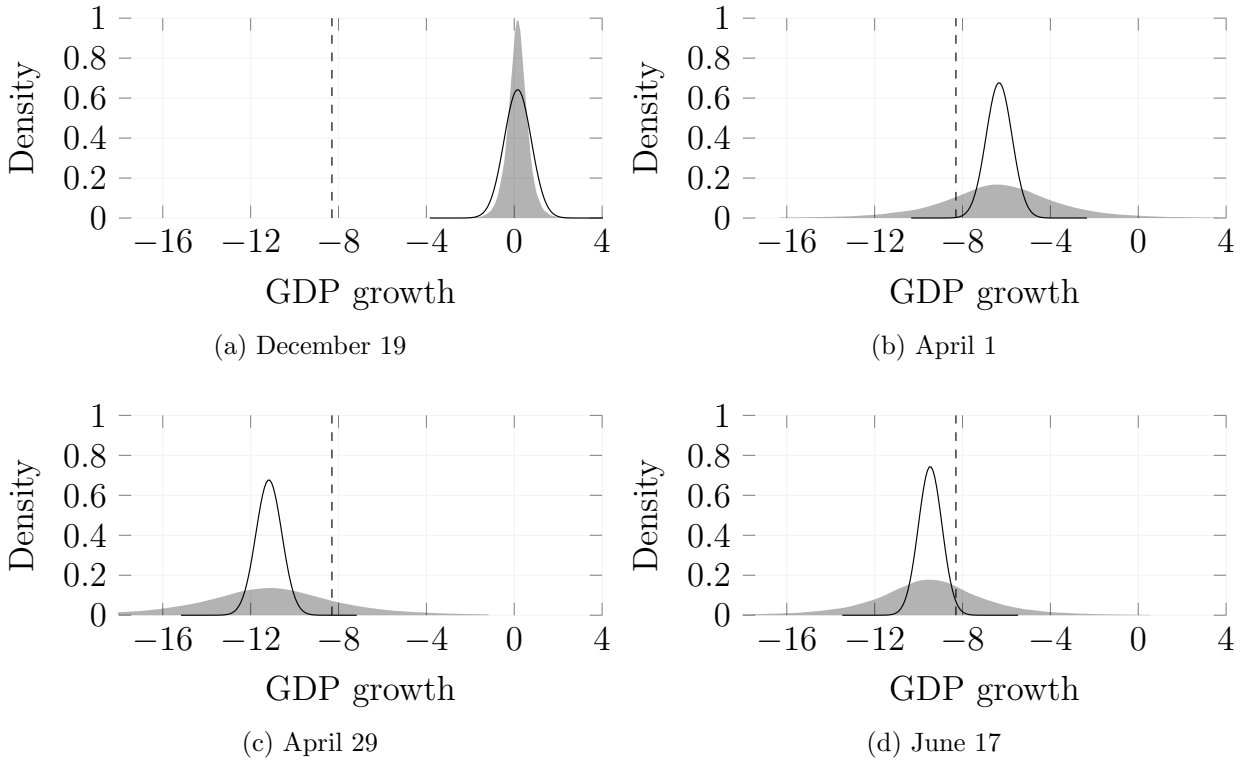
Similarly, the European Central Bank (ECB) decided not to present uncertainty intervals around their projections in their June projection (European Central Bank, 2020). They justified the decision saying that:

> This reflects the fact that the standard computation of the ranges (based on historical projection errors) would not, in the present circumstances, provide a reliable indication of the unprecedented uncertainty surrounding the current projections.

Because the ECB's uncertainty intervals are based on unconditional moments of the forecast error distribution, they would severely underestimate the prevailing uncertainty.

Figure 9 and 10 show predictive densities incorporating time-varying uncertainty for GDP growth and unemployment, overlaid with predictive densities obtained from fitting zero-mean normal distributions to the historical forecast errors. The figures show the tremendous impact the pandemic-induced revisions had on forecast uncertainty. In December, the predictive densities for both GDP growth and unemployment were more concentrated than the corresponding densities based on normal distributions, indicating that forecast uncertainty was lower than usual. In April, the rise in uncertainty was remarkable, particularly for the GDP growth forecast.[10] The unconditional normal distribution predictive density was, by construction, unaffected by the large revisions, but nevertheless gave the visual impression of a forecast that was much more certain than it arguably was. Forecasts for GDP growth that were made during the pandemic and that employed time-varying uncertainty were better

[10]The rise in forecast uncertainty was counteracted by a decrease in forecast horizon between December and April. Despite the decrease in forecast horizon, uncertainty increased.




(a) December 19

(b) April 1

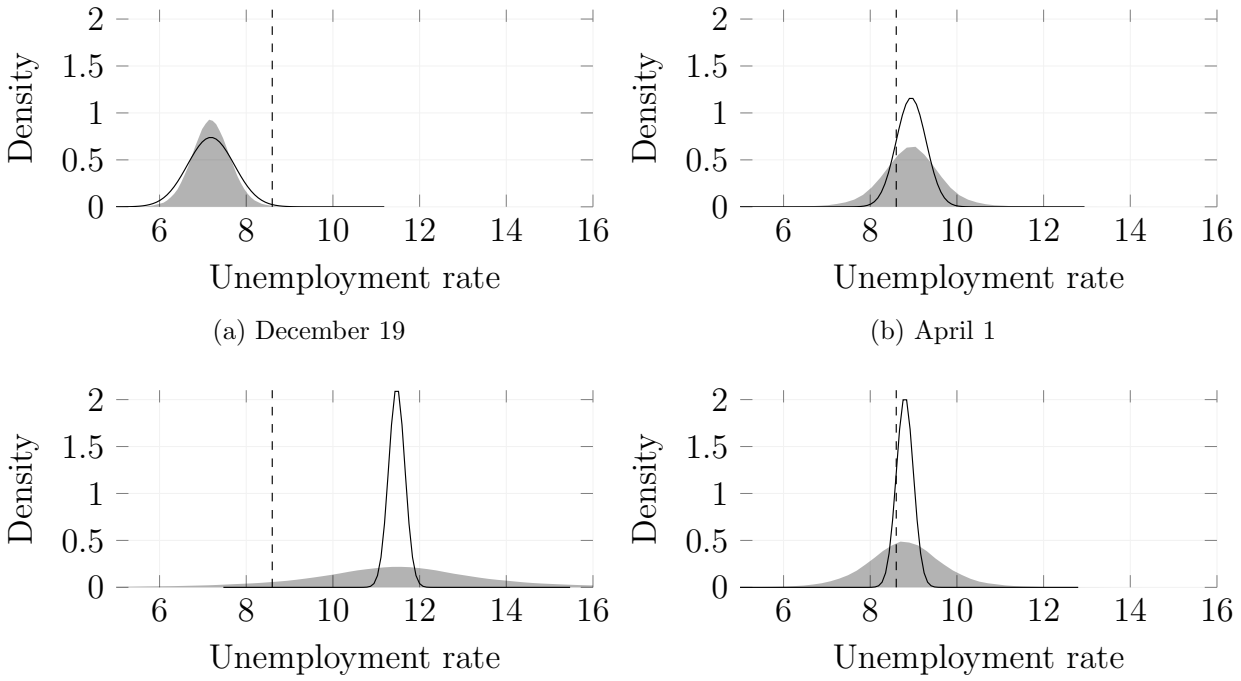
(c) April 29

(d) June 17

Figure 9: Predictive distributions for GDP growth in 2020Q2. Solid lines show unconditional predictive distributions based on normality, vertical dashed line shows the first published outcome. The December forecast is a $h = 2$ step forecast, April forecasts are $h = 1$ forecasts, and the June forecast a $h = 0$ forecast (nowcast).

according to the logarithmic score, as evidenced by higher predictive distributions at the value of the outcome. The forecast errors were unusually large, which was anticipated using time-varying uncertainty. For the unemployment rate, the normality-based unconditional uncertainty forecasts scored better for the April 1 and June 17 forecasts. The April 29 forecast, however, was more than 7 standard deviations away from the outcome. The description of forecast uncertainty using unconditional levels of uncertainty was ex-post clearly inappropriate for the April 29 forecast.

To see the impact of the revisions between successive forecasts, Figure 11 and 12 plot the 50%, 75% and 90% intervals for GDP growth and unemployment. The left panels show intervals based on time-varying uncertainty, and the right panels show unconditional




(a) December 19

(b) April 1

(c) April 29

(d) June 17

Figure 10: Predictive distributions for the unemployment rate in 2020Q2. Solid lines show unconditional predictive distributions based on normality, vertical dashed line shows the first published outcome. The December forecast is a $h = 2$ step forecast, April 1 a $h = 1$ forecast, and April 29 and June forecasts are $h = 0$ forecasts (nowcasts).

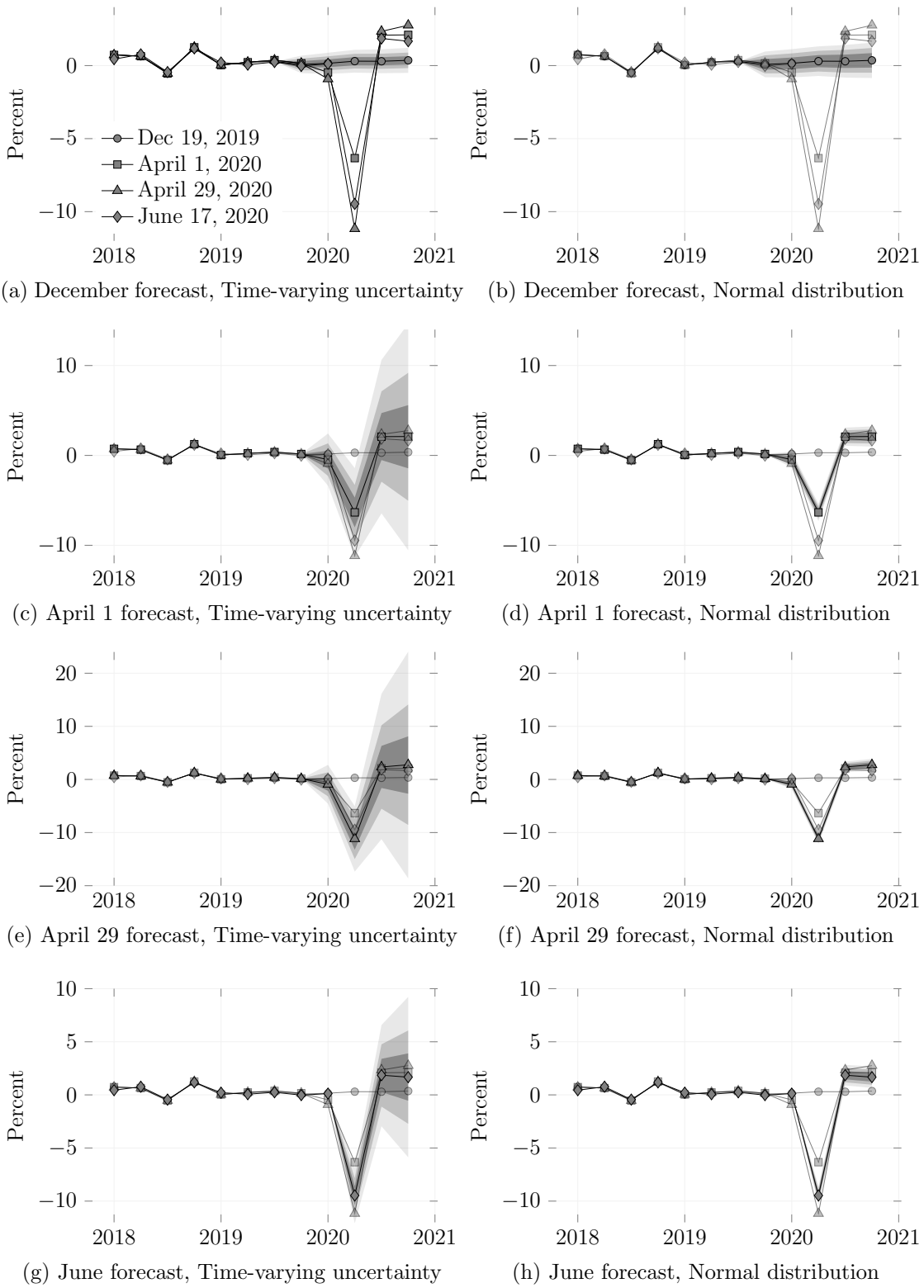
normality-based uncertainty intervals.

Intervals based on time-varying uncertainty convey the message that the forecasts made in April were unprecedentedly uncertain, saying that the margin of error was expected to be much higher than otherwise. For the April 1 forecast, the 90% interval for GDP growth in 2020Q2 ranged from approximately -1 to -11. While the intervals for 2020Q2 are wide, the fans expand even further for the later half of 2020. Since gauging the present was extremely challenging in April, it is not surprising that the uncertainty around the forecasts for the remainder of 2020 is so large so as to essentially say that anything is possible. The corresponding normality-based intervals for the April 1 GDP forecast are barely visible because of the large scales. The revisions for the April 29 forecast for both GDP and unemployment increased the level of uncertainty further, extending the range of outcomes that were plausible given the size of the revisions that had occurred.

A major challenge during the initial phase of the pandemic was that severe economic effects were widely expected, but available traditional data sources did not show any signs of an impending crisis. Figure 9–12 illustrate the benefit of basing the estimation of uncertainty on revisions—because revisions are forward-looking, the measure of time-varying uncertainty can react in real time as soon as changes are made to the forecasts. Mapping qualitative descriptions of uncertainty to quantitative measures is certainly challenging, but the estimated time-varying forecast uncertainty appears to accompany the sentiments voiced by the NIER and the Riksbank well insofar as painting a clear picture of unparalleled levels of uncertainty.

# 7 Conclusion

I have studied whether time-varying uncertainty is useful for describing the uncertainty surrounding the National Institute of Economic Research's forecasts for GDP growth and unemployment in Sweden. My results indicate that the answer is yes, time-varying uncer-



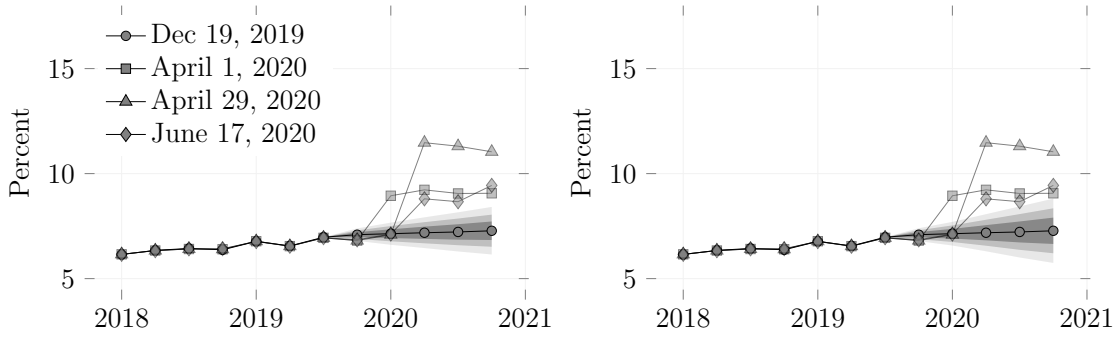

(a) December forecast, Time-varying uncertainty (b) December forecast, Normal distribution

(c) April 1 forecast, Time-varying uncertainty (d) April 1 forecast, Normal distribution

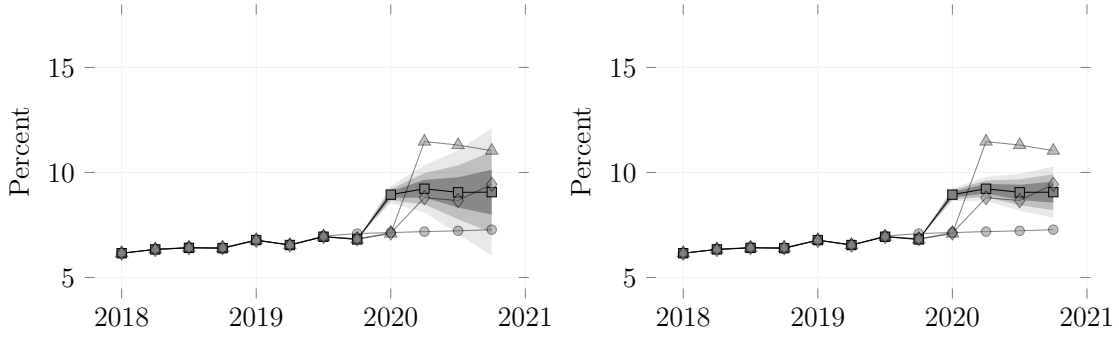
(e) April 29 forecast, Time-varying uncertainty (f) April 29 forecast, Normal distribution

(g) June forecast, Time-varying uncertainty (h) June forecast, Normal distribution

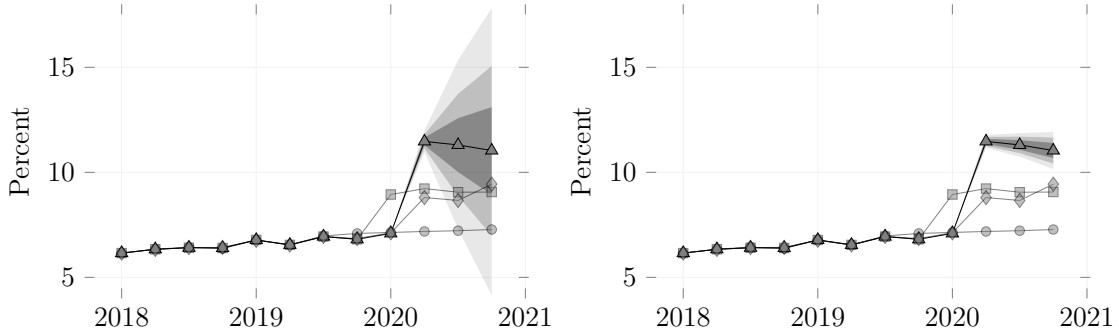
Figure 11: Fan charts for GDP growth. The shaded areas display 50%, 75% and 90% pointwise intervals.



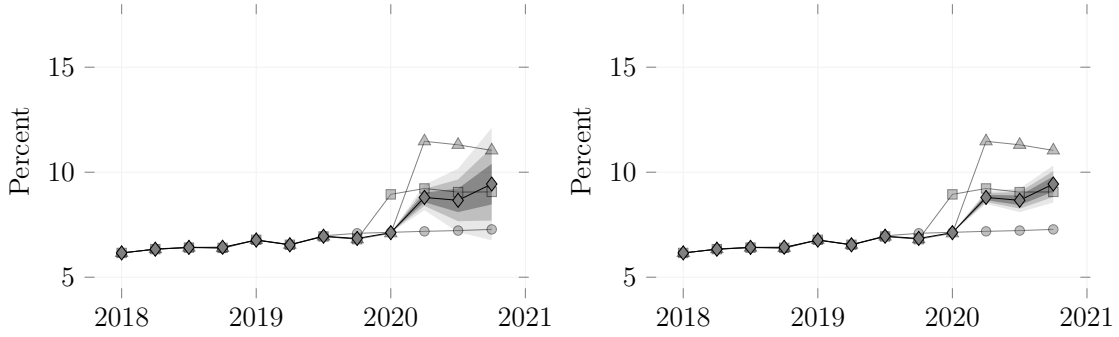

(a) December forecast, Time-varying uncertainty (b) December forecast, Normal distribution




(c) April 1 forecast, Time-varying uncertainty (d) April 1 forecast, Normal distribution




(e) April 29 forecast, Time-varying uncertainty (f) April 29 forecast, Normal distribution




(g) June forecast, Time-varying uncertainty (h) June forecast, Normal distribution

Figure 12: Fan charts for the unemployment rate. The shaded areas display 50%, 75% and 90% pointwise intervals.

tainty improves the statistical properties of uncertainty intervals and predictive distributions. I find enhancements with respect to multiple statistics and for both variables, GDP growth and unemployment, that I study. My sample starts in 2002 and is thus relatively short. The results should be particularly encouraging for those forecasters and policy institutions, of which there are many, that lack an extensive record of quarterly forecasting errors to rely on for estimation.

I provide additional evidence of the relevance of time-varying uncertainty in an application to the Covid-19 pandemic. The application highlights a policy value of the approach that is more communicative than statistical. A challenge for forecasters in times of distress is to convey to the public that their forecasts are more uncertain than what would be expected on average. By showing unconditional, normality-based uncertainty intervals based on past performance, average forecast uncertainty becomes the focus although average uncertainty, as during the Covid-19 pandemic, may be a poor indicator for the prevailing level of uncertainty. In contrast, my application shows that time-varying uncertainties react forcefully and in real time because of the large revisions caused by the pandemic.

A time-varying characterization of forecast uncertainty therefore appears to behave precisely as desired: During the last five years, forecast uncertainty was relatively speaking lower, and in the evaluation during this period the approach based on time-varying uncertainty resulted in better statistical properties. However, when the pandemic hit and caused an extraordinary level of uncertainty, the method immediately picked this up to yield unprecedented levels of uncertainty. As the method brings both statistical and communicative improvements, it should therefore be an appealing tool for many forecasters.

# References


Blix, M. and Sellin, P. (1998). Uncertainty Bands for Inflation Forecasts. Working Paper 65, Sveriges Riksbank.

Britton, E., Fisher, P., and Whitley, J. (1998). The Inflation Report Projections: Understanding the fan chart. *Bank of England Quarterly Bulletin*, 38(1):30–37.

Clark, T. E., McCracken, M. W., and Mertens, E. (2020). Modeling Time-Varying Uncertainty of Multiple-Horizon Forecast Errors. *The Review of Economics and Statistics*, 102(1):17–33, doi:10.1162/rest_a_00809.

Clements, M. P. (2004). Evaluating the Bank of England Density Forecasts of Inflation. *Economic Journal*, 114(498):844–866.

Clements, M. P. (2018). Are Macroeconomic Density Forecasts Informative? *International Journal of Forecasting*, 34(2):181–198, doi:10.1016/j.ijforecast.2017.10.004.

Cogley, T. and Sargent, T. J. (2005). Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US. *Review of Economic Dynamics*, 8(2):262–302, doi:10.1016/j.red.2004.10.009.

Del Negro, M. and Primiceri, G. E. (2015). Time Varying Structural Vector Autoregressions and Monetary Policy: A Corrigendum. *The Review of Economic Studies*, 82(4):1342–1345, doi:10.1093/restud/rdv024.

Durbin, J. and Koopman, S. J. (2002). A Simple and Efficient Simulation Smoother for State Space Time Series Analysis. *Biometrika*, 89(3):603–616, doi:10.1093/biomet/89.3.603.

European Central Bank (2020). Eurosystem Staff Macroeconomic Projections for the Euro Area. https://www.ecb.europa.eu/pub/pdf/other/ecb.projections202006_eurosystemstaff~7628a8cf43.en.pdf.

Gelfand, A. E. and Dey, D. K. (1994). Bayesian Model Choice: Asymptotics and Exact Calculations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 56(3):501–514, doi:10.1111/j.2517-6161.1994.tb01996.x.

Geweke, J. (1999). Using Simulation Methods for Bayesian Econometric Models: Inference, Development, and Communication. *Econometric Reviews*, 18(1):1–73, doi:10.1080/07474939908800428.

Gneiting, T., Balabdaoui, F., and Raftery, A. E. (2007). Probabilistic Forecasts, Calibration and Sharpness. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(2):243–268, doi:10.1111/j.1467-9868.2007.00587.x.

Gneiting, T. and Raftery, A. E. (2007). Strictly Proper Scoring Rules, Prediction, and Estimation. *Journal of the American Statistical Association*, 102(477):359–378, doi:10.1198/016214506000001437.

Gronau, Q. F., Sarafoglou, A., Matzke, D., Ly, A., Boehm, U., Marsman, M., Leslie, D. S., Forster, J. J., Wagenmakers, E.-J., and Steingroever, H. (2017). A Tutorial on Bridge Sampling. *Journal of Mathematical Psychology*, 81:80–97, doi:10.1016/j.jmp.2017.09.005.

Jordan, A., Krüger, F., and Lerch, S. (2019). Evaluating Probabilistic Forecasts with scoringRules. *Journal of Statistical Software*, 90(12):1–37, doi:10.18637/jss.v090.112.

Kim, S., Shephard, N., and Chib, S. (1998). Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. *The Review of Economic Studies*, 65(3):361–393, doi:10.1111/1467-937X.00050.

Knüppel, M. (2014). Efficient Estimation of Forecast Uncertainty Based on Recent Forecast Errors. *International Journal of Forecasting*, 30(2):257–267, doi:10.1016/j.ijforecast.2013.08.004.

Knüppel, M. (2018). Forecast-Error-Based Estimation of Forecast Uncertainty When the Horizon is Increased. *International Journal of Forecasting*, 34(1):105–116, doi:10.1016/j.ijforecast.2017.08.006.

Knüppel, M. and Schultefrankenfeld, G. (2012). How Informative Are Central Bank Assessments of Macroeconomic Risks? *International Journal of Central Banking*, 8(3):87–139.

Knüppel, M. and Schultefrankenfeld, G. (2019). Assessing the Uncertainty in Central Banks' Inflation Outlooks. *International Journal of Forecasting*, 35(4):1748–1769, doi:10.1016/j.ijforecast.2019.03.014.

Krüger, F., Lerch, S., Thorarinsdottir, T. L., and Gneiting, T. (2020). Predictive Inference Based on Markov Chain Monte Carlo Output. *Preprint available at arXiv*, https://arxiv.org/abs/1608.06802.

Meng, X.-L. and Wong, W. H. (1996). Simulating Ratios of Normalizing Constants via a Simple Identity: A Theoretical Exploration. *Statistica Sinica*, 6(4):831–860.

Mincer, J. and Zarnowitz, V. (1969). The Evaluation of Economic Forecasts. In Mincer, J., editor, *Economic Forecasts and Expectations: Analysis of Forecasting Behavior and Performance*, pages 3–46. National Bureau of Economic Research.

National Institute of Economic Research (2020). Swedish Economy Report, April. https://www.konj.se/english/publications/swedish-economy-report/swedish-economy/2020-04-01-particularly-deep-economic-downturn-in-the-wake-of-covid-19.html.

Ohnsorge, F. L., Stocker, M., and Some, M. Y. (2016). Quantifying Uncertainties in Global Growth Forecasts. Policy Research Working Paper Series 7770, The World Bank.

Omori, Y., Chib, S., Shephard, N., and Nakajima, J. (2007). Stochastic Volatility with Leverage: Fast and Efficient Likelihood Inference. *Journal of Econometrics*, 140(2):425–449, doi:10.1016/j.jeconom.2006.07.008.

Overstall, A. M. and Forster, J. J. (2010). Default Bayesian Model Determination Methods for Generalised Linear Mixed Models. *Computational Statistics & Data Analysis*, 54(12):3269–3288, doi:10.1016/j.csda.2010.03.008.

Primiceri, G. E. (2005). Time Varying Structural Vector Autoregressions and Monetary Policy. *The Review of Economic Studies*, 72(3):821–852, doi:10.1111/j.1467-937X.2005.00353.x.

Reifschneider, D. L. and Tulip, P. (2017). Gauging the Uncertainty of the Economic Outlook Using Historical Forecasting Errors: The Federal Reserve's Approach. Finance and Economics Discussion Series 2017-020, Board of Governors of the Federal Reserve System, doi:10.17016/FEDS.2017.020.

Sveriges Riksbank (2007). Monetary Policy Report 2007:1, February. http://archive.riksbank.se/Upload/Dokument_riksbank/Kat_publicerat/Rapporter/2007/ppr_07_01_nyeng.pdf.

Sveriges Riksbank (2020). Monetary Policy Report 2020:2, April. https://www.riksbank.se/en-gb/monetary-policy/monetary-policy-report/2020/monetary-policy-report-april-20202/.

Wallis, K. F. (2004). An Assessment of Bank of England and National Institute Inflation Forecast Uncertainties. *National Institute Economic Review*, 189(1):64–71.

# A Econometric Details

## A.1 Prior Distributions

The unknown parameters of the model are $\theta = \{\lambda_{1:H+1,0}, \beta, a\}$, where

$$\lambda_{1:H+1,0} = (\lambda_{1,0}, \ldots, \lambda_{H+1,0})'$$
$$\beta = (\beta_1, \ldots, \beta_{H+1})',$$

and $a$ is the lower triangular part of $A$, that is,

$$a = (a_{2,1}, a_{3,1}, \ldots, a_{H+1,1}, a_{3,2}, \ldots, a_{H+1,2}, \ldots, a_{H+1,H})'.$$

Because of a limited amount of data, I use weakly informative priors. The prior distributions for the intercepts $\lambda_{i,0}$ are independent normal distributions $\lambda_{i,0} \sim$ iid $\mathrm{N}(0, 10)$. For the loadings $\beta_i$, I use $\beta_i \sim$ iid $\mathrm{N}(0, 0.5)$. For the lower triangular elements of $a$, I let $a_j \sim$ iid $\mathrm{N}(0, 1)$.

## A.2 Estimation Using Markov Chain Monte Carlo

Estimation proceeds using a standard Gibbs sampling algorithm that has as its stationary distribution

$$p(a, \beta, \lambda_{1:H+1,0}, \lambda_{0,1:T} | \eta),$$

where $\lambda_{0,1:T} = (\lambda_{0,1}, \ldots, \lambda_{0,T})'$. At iteration $r$, the Markov Chain Monte Carlo (MCMC) algorithm cycles through the following steps.

**Sample** $a$ Given $\lambda_{1:H+1,0}$, $\beta$ and $\lambda_{0,1:T}$, the equation-specific volatilities $\lambda_{i,1:T}$ are known. A sample of the full vector $a$ can be obtained by recursively sampling parameters equation

by equation, see, for instance, Cogley and Sargent (2005). To provide some intuition, the first equation in (2) is

$$e_{t-1|t-1} = \exp(0.5\lambda_{1,t})\epsilon_{1,t},$$

and so $\epsilon_{1,t}$ is known conditional on $\exp(0.5\lambda_{1,t})$. The second equation is

$$\mu_{t|t} = a_{2,1}\exp(0.5\lambda_{1,t})\epsilon_{1,t} + \exp(0.5\lambda_{2,t})\epsilon_{2,t}.$$

Multiplying both sides of the equation by $\exp(-0.5\lambda_{2,t})$ yields a standard homoscedastic Bayesian linear regression model. The error $\epsilon_{2,t}$ is also known given $a_{2,1}$. Sampling continues analogously for the remaining elements of $a$ equation by equation.

**Sample** $(\lambda_{1:H+1,0}, \lambda_{0,1:T})$ Using the Omori et al. (2007) refinement of the Kim et al. (1998) mixture approximation, the log of the squared forecast errors can be written as

$$\begin{aligned}\eta_{i,t}^2 &= \lambda_{i,0} + \beta_i\lambda_{0,t} + \epsilon_{i,t}^2 \\ &\approx \lambda_{i,0} + \beta_i\lambda_{0,t} + m_{r_{i,t}} + s_{r_{i,t}}\nu_{i,t}, \quad \nu_{i,t} \sim \mathrm{N}(0,1),\end{aligned}$$

where $r_{i,t} \in \{1, 2, \ldots, 10\}$ are mixture indicators and $m_{r_{i,t}}$ and $s_{r_{i,t}}$ are the means and standard deviations of the mixtures. The indicators enable the use of a conditionally linear state-space model with normally distributed errors although $\log(\epsilon_{i,t}^2)$ is non-normally distributed.

Let $\tilde{y}_{i,t} = \log(\eta_{i,t}^2) - m_{r_{i,t}}$, and $\tilde{y}_t = (\tilde{y}_{1,t}, \ldots, \tilde{y}_{H+1,t})'$. The mixture approximation yields the state-space model

$$\begin{aligned}\tilde{y}_t &= Z\lambda_t + G_t\epsilon_{y,t} \\ \lambda_t &= \lambda_{t-1} + L\epsilon_{\lambda,t},\end{aligned}$$

where $\epsilon_{y,t} \sim \text{iid}\,\text{N}(0, I_{H+1})$ is independent of $\epsilon_{\lambda,t} \sim \text{iid}\,\text{N}(0,1)$, and

$$Z = (\beta, I_{H+1}), \quad G_t = \text{diag}(s_{r_{1,t}}, s_{r_{2,t}}, \ldots, s_{r_{H+1,t}})'$$

$$L = \begin{pmatrix} 1 \\ 0_{H+1\times 1} \end{pmatrix},$$

$$\lambda_t = \begin{pmatrix} \lambda_{0,t} \\ \lambda_{1,t} \\ \vdots \\ \lambda_{H+1,t} \end{pmatrix}, \quad \lambda_{0:H+1,0} \sim \text{N}(a_0, P_0), \quad a_0 = 0_{H+2\times 1},$$

$$P_0 = \begin{pmatrix} 0 & 0_{1\times H+1} \\ 0_{H+1\times 1} & 10 I_{H+1} \end{pmatrix}.$$

A simulation smoother, see Durbin and Koopman (2002), conditional on the mixture indicators $r_{i,t}$ is used to produce a draw from the desired posterior distribution. The mixture indicators are sampled in an auxiliary step that immediately precedes the simulation smoother. For details, see Kim et al. (1998); Del Negro and Primiceri (2015).

**Sample** $\beta$ From the mixture formulation of the state-space model,

$$\frac{\log(\eta_{i,t}^2) - m_{r_{i,t}} - \lambda_{i,0}}{s_{r_{i,t}}} = \beta_i \frac{\lambda_{0,t}}{s_{r_i}} + \nu_{i,t}.$$

Given the mixture indicators $r_{i,t}$, the volatility intercepts $\lambda_{1:H+1,0}$ and the common volatility $\lambda_{0,1:T}$, Bayesian linear regression can again be used to sample from the posterior of $\beta_i$.

## A.3 Posterior Simulation of Forecast Errors

Given a draw from the MCMC output, the log-volatility factor is simulated forwards and future nowcast errors and forecast revisions simulated given the future level of volatility.

Finally, (1) is used in order to move from $p(\eta_{T+1:T+H+1}|\eta_{1:T})$ to $p(e_{T|T},\ldots,e_{T+H|T}|\eta_{1:T})$. The steps of the algorithm are:

1. Given $\lambda_{0,T}$, simulate $\{\lambda_{0,T+i}\}_{i=1}^{H}$ using (3) and form $\Lambda_t$ for $t = T+1,\ldots,T+H+1$
2. Given $A$ and $\Lambda_t$ simulate $\eta_t$ by drawing $\epsilon_t \sim \mathrm{N}(0, I_{H+1})$ and using (2) for $t = T+1,\ldots,T+H+1$
3. Given $\{\eta_{T+i}\}_{i=1}^{H+1}$, compute $\{e_{T+i-1|T}\}_{i=1}^{H+1}$ using (1)

## A.4 Evaluating Density Scores

The log score can be efficiently estimated by (Krüger et al., 2020)

$$\begin{aligned} \mathrm{LS}_h &= \frac{1}{T_h}\sum_{t=1}^{T_h}\widehat{\mathrm{LS}}_{t,h} \\ \widehat{\mathrm{LS}}_{t,h} &= -\frac{1}{R}\sum_{r=1}^{R}\log\left[f(e_{t|t-h}|\theta^{(r)},\lambda^{(r)}_{1:H+1,0},\lambda^{(r)}_{0,1:t-h},\eta_{1:t-h})\right], \end{aligned} \tag{4}$$

where $f(e_{t|t-h}|\theta^{(r)},\lambda^{(r)}_{1:H+1,0},\lambda^{(r)}_{0,1:t-h},\eta_{1:t-h})$ is the density of a normal distribution with mean 0 and standard deviation $\sigma^{(r)}_{t|t-h}$. An explicit expression for $\sigma^2_{t|t-h}$ is provided in Appendix A.5.

Analogously, the continuous ranked probability score can be computed as (Jordan et al., 2019)

$$\begin{aligned} \mathrm{CRPS}_h &= \frac{1}{T_h}\sum_{t=1}^{T_h}\widehat{\mathrm{CRPS}}_{t,h} \\ \widehat{\mathrm{CRPS}}_{t,h} &= \frac{1}{R}\sum_{r=1}^{R}\sigma^{(r)}_{t|t-h}\,\mathrm{CRPS}\left(\Phi,\frac{e_{t|t-h}}{\sigma^{(r)}_{t|t-h}}\right) \\ \mathrm{CRPS}\left(\Phi,z\right) &= z\left(2\Phi(z)-1\right)+2\phi(z)-\frac{1}{\sqrt{\pi}}, \end{aligned}$$

where $\Phi$ and $\phi$ are the standard normal cumulative and probability distribution functions,

respectively.

Both scores are univariate and marginal with respect to the horizon, and neglect the joint behavior of predictive distributions across horizons. Multiple horizons can be evaluated jointly by changing the univariate predictive distribution in (4) to the joint predictive density. Appendix A.5 provides an expression for the conditionally normal joint predictive distribution of $(e_{t|t}, \ldots, e_{t+H|t})'$.

## A.5 Conditional Variance of Future Forecast Errors

Using the law of total variance,

$$\begin{aligned}\Sigma_{t+h|t} &\equiv \mathrm{Var}(\eta_{t+h}|\theta, \lambda_{1:t}, \eta_{1:t}) \\ &= \mathrm{E}\left[\mathrm{Var}(\eta_{t+h}|\theta, \lambda_{1:t}, \lambda_{t+1:t+h}, \eta_{1:t})|\theta, \lambda_{1:t}, \eta_{1:t}\right] \\ &+ \mathrm{Var}\left[\underbrace{\mathrm{E}(\eta_{t+h}|\theta, \lambda_{1:t}, \lambda_{t+1:t+h}, \eta_{1:t})}_{=0}|\theta, \lambda_{1:t}, \eta_{1:t}\right].\end{aligned}$$

The first term is

$$\begin{aligned}&\mathrm{E}\left[\mathrm{Var}(\eta_{t+h}|\theta, \lambda_{1:t}, \lambda_{t+1:t+h}, \eta_{1:t})|\theta, \lambda_{1:t}, \eta_{1:t}\right] \\ &\qquad = A\,\mathrm{E}\left(\Lambda_{t+h}|\theta, \lambda_{1:t}, \eta_{1:t}\right)A'. \quad (5)\end{aligned}$$

The $i$th diagonal element of $\Lambda_{t+h}$ is $\exp(\lambda_{i,t+h}) = \exp(\lambda_{i,0} + \beta_i\lambda_{0,t+h})$, and so

$$\begin{aligned}\lambda_{i,t+h} &= \lambda_{i,0} + \beta_i\lambda_{0,t+h} \\ &= \lambda_{i,0} + \beta_i\left(\lambda_{0,t} + \sum_{s=1}^{h}\nu_s\right).\end{aligned}$$

The conditional distribution of the expression is $\mathrm{N}(\lambda_{i,0} + \beta\lambda_{0,t}, h\beta_i^2)$. Hence, the conditional distribution of $\exp(\lambda_{i,t+h})$ is log-normal and its conditional expectation is

$$\mathrm{E}\left(\exp\{\lambda_{i,t+h}\}|\theta, \lambda_{1:t}, \eta_{1:t}\right) = \exp\left\{\lambda_{i,0} + \beta\lambda_{0,t} + \frac{h\beta_i^2}{2}\right\}. \tag{6}$$

The conditional variance can thus be computed using (5) and (6), whereby

$$\eta_{t+h}|\theta, \lambda_{1:t}, \eta_{1:t} \sim \mathrm{N}(0, \Sigma_{t+h|t}).$$

Let $\Sigma_{t+h|t}^{(i,j)}$ denote element $(i,j)$ of $\Sigma_{t+h|t}$. The forecast error $e_{t+h|t}$ is conditionally normal with mean zero and variance

$$\sigma_{t+h|t}^2 \equiv \mathrm{Var}(e_{t+h|t}|\theta, \lambda_{1:t}, \eta_{1:t}) = \sum_{j=0}^{h} \Sigma_{t+h+1-j|t}^{(j+1,j+1)}.$$

Since $\eta_t$ and $\eta_s$, $t \neq s$, are conditionally independent and normally distributed, their joint distribution is normal. Because forecast errors $e_{t+\ell|t}$ and $e_{t+h|t}$ are just linear combinations of $\eta$, they are also jointly normal. The mean vector is zero, and element $(\ell, h)$ of the conditional covariance of $(e_{t|t}, \ldots, e_{t+H|t})'$ is

$$\begin{aligned}
\mathrm{Cov}(e_{t+\ell|t}, e_{t+h|t}) &= \mathrm{Cov}(\sum_{j=0}^{\ell} \eta_{t+\ell+1-j,j+1}, \sum_{i=0}^{h} \eta_{t+h+1-i,i+1}) \\
&= \sum_{j=0}^{\ell} \mathrm{Cov}(\eta_{t+\ell+1-j,j+1}, \eta_{t+\ell+1-i,j+h-\ell+1}) \\
&= \sum_{j=0}^{\ell} \Sigma_{t+\ell+1-j|t}^{(j+1,j+h-\ell+1)}.
\end{aligned}$$