Draft Preregistration - Simple Macroeconomic Forecast Distributions

Friederike Becker, Fabian Krüger and Melanie Schienle

August 16, 2024

NOTE: THIS IS A PRELIMINARY AND INCOMPLETE DRAFT THAT IS SUBJECT TO CHANGE BEFORE ITS SCHEDULED PUBLICATION IN SEPTEMBER 2024.

Background

Macroeconomic forecasts feature prominently in economic policy debates. The public forecasting hemisphere is generally dominated by point forecasts, and explicit quantification of the uncertainty surrounding these is usually unavailable. While the academic literature on macroeconomic forecasting has long studied measures of forecast uncertainty, these measures are typically hard to access for the broader public (e.g., not available in real time). This situation is unsatisfactory since information about uncertainty is essential to assess whether or not a given point forecast serves as a credible argument in policy debates. In Becker et al. (2024), we set out to provide simple forecast distributions for inflation and real GDP growth for seven advanced economies in real-time, based on a pre-existing panel of point forecasts from the International Monetary Fund. The present pre-registration protocol serves to specify a transparent set of rules and criteria for the publication and subsequent evaluation of our forecasts.

Forecast format and release schedule

We publish probabilistic fixed-event forecasts in the format of two central intervals at the 50% and 80% confidence level, requiring quantiles at the $10\%, 25\%, 75\%$ and 90% levels. The forecast targets are the annual rates of CPI inflation and GDP growth, respectively. Our forecasts will be published twice a year, once in Spring and once in Fall, each containing forecasts for the current and next year. For each combination of country, target series and target year, we thus obtain up to four forecasts in total, denoted by their horizons, that is, "Fall, current", "Spring, current", "Fall, next", and "Spring, next".

Due to our reliance on the International Monetary Fund World Economic Outlook (IMF WEO), we will closely follow its release schedule. The WEO is published biannually, once in April and once in September/October (International Monetary Fund, 2024). We aim to release our forecasts within 14 days following the respective WEO publication. From a practical perspective, the exact timing is clearly relevant for the "Fall, current" forecast, but much less so for the other horizons.
As a general point, we note the dependence of our study on the continued and timely release of IMF forecasts but deem this as uncritical, given its general reputation and its consistent publication for more than 30 years.

Construction of forecast intervals

We construct forecast intervals based on current point forecasts and past point forecast errors from the IMF WEO. Concretely, for each combination of country, target and forecast horizon, we take the 6th and 9th value from the order statistics of the last 11 absolute-transformed forecast errors and add and subtract these to the current point forecast. We thereby obtain the upper and lower endpoints of the 50% and 80% intervals, respectively. We impose the assumption that the uncertainty should not decrease with the forecast horizon within a given release and average forecast intervals between the forecast horizons when this is not naturally the case. For more details on the methodology, see Becker et al. (2024).
We require truth values for both the calculation of forecast errors and the final evaluation. The IMF WEO includes four truth values for each target year, which are published bi-annually following the initial realization. We use the respective truth value that is published in the first fall release, whenever available. In instances where it's not, we supplement this with that published in the first spring release. Most notably, this is necessary in spring for the calculation of the forecast errors for the previous target year.

Scope

We issue forecasts for four forecast horizons ("Fall, current", "Spring, current", "Fall, next", and "Spring, next"), as detailed in the previous section.
We issue forecasts for the Group of Seven countries (henceforth G7, including Canada, France, Germany, Italy, Japan, the United Kingdom and the United States), which we also studied in our retrospective evaluation (Becker et al., 2024).

Forecast evaluation

Two metrics will be used to evaluate forecasts. As our main metric, we will report central interval coverage, that is, the empirical rate at which the forecast intervals covered the true value. Empirical

interval coverage will be aggregated across countries and forecast horizons, and will be reported separately by confidence level and target series. As argued by Raftery (2016), forecast users can view this as a central criterion for trust in forecast intervals.

As a further measure of forecast performance, we evaluate with a weighted average of the respective interval scores (Gneiting and Raftery, 2007) at the 50% and 80% level. For this, we use weights as suggested by Bracher et al. (2021), that is, with $\tau_1 = 0.5$ and $\tau_2 = 0.8$, $w_1 = \frac{1-\tau_1}{2} = 0.25$ and $w_2 = \frac{1-\tau_2}{2} = 0.1$. The interval score will further be aggregated via the arithmetic mean across countries, and reported separately by target series and forecast horizon.

Benchmark forecasts

As in our retrospective evaluation in Becker et al. (2024), we plan to fit benchmark models to evaluate and compare our main method's performance to. Note that benchmark forecasts are needed to meaningfully interpret the level of interval score values. While the individual models' interval coverage can be interpreted on their own (should be as close as possible to nominal level), some natural variation can be expected, making benchmark forecasts useful additions to contextualize observed coverage values.

We plan for these benchmark models to be the same as in the aforementioned retrospective evaluation, notably an autoregressive model (AR) of order one and a Bayesian vector autoregressive model (BVAR) (Primiceri, 2005; Del Negro and Primiceri, 2015). To obtain forecast intervals from these models, we apply the same process as for the WEO forecasts, as described in section 'Construction of forecast intervals'. For the BVAR model, we additionally directly extract intervals from the model's forecast distribution. We thus obtain three sets of benchmark forecast intervals in total.

While AR is simple and transparent, BVAR is much more flexible, allowing for time-varying parameters, which however comes at the cost of increased complexity. Additionally, the directly extracted forecast intervals from the BVAR model are fully based on parametric modeling assumptions, and thus serve as a further check of the methodology.

At the time of evaluation (see section 'Evaluation period'), we plan to fit the benchmark models on quarterly data of the respective target series, obtained from the OECD. Note that this plan is subject to data availability; in particular, we reserve the right to resort to other data sources should the OECD data be unavailable, or should there become available another more suitable data source. Moreover, if data for one of the target series is unavailable at that time, the respective country will be excluded from the final evaluation, for the duration for which the data is unavailable.

Furthermore, we reserve the right to include other benchmark models at the time of evaluation.

Evaluation period

We will evaluate forecasts made between fall 2024 and including fall 2029, for target years 2024 up to 2029. For each combination of target series and country, we thus obtain six forecasts for the "Fall, current" horizon, five forecasts each for the "Spring, current" and "Fall, next" horizons, and four forecasts for the "Spring, next" horizon.

Technical Points

Forecasts will be published in the subdirectory 'forecasts' in the following GitHub repository:

https://github.com/KITmetricslab/MacroPI/

The csv-files will be clearly identified with the respective forecasts origin, for instance, the file containing forecasts made in Spring 2025 will be named 'forecasts_Spring2025.csv'. In each file, a forecast will be comprised of one row (long format), identified by the following columns:

- 'country': identifies the country using the ISO Alpha-3 code
- 'target': identifies the target series. 'gdp_growth' for real GDP Growth, 'inflation' for inflation
- 'forecast_year': the year the forecast is made. Uniquely identifies the forecast origin in conjunction with column 'forecast_season'
- 'forecast_season': the season the forecast is made, 'F' for Fall and 'S' for Spring. Uniquely identifies the forecast origin in conjunction with column 'forecast_year'
- 'target_year': the year the forecast is issued for.
- 'quantile': the quantile level, one of $0.1, 0.25, 0.75$ or 0.9. The 50% forecast interval is comprised of the 0.25 and 0.75 quantiles, the 80% forecast interval of the 0.1 and 0.9 quantiles.
- 'prediction': prediction value for the forecast instance identified by all other preceding columns.

Further, we supply historical truth data and point forecasts, both from the corresponding IMF World Economic Outlook publication. The name of the csv-file identifies the forecast origin. These csv-files contain the following columns:

- 'country': identifies the country using the ISO Alpha-3 code
- 'target': identifies the target series. 'gdp_growth' for real GDP Growth, inflation for 'inflation'
- 'target year': the year the forecast is issued for.
- 'prediction'/'truevalue': prediction (point forecasts) or true value (historic values) for the instance identified by all other preceding columns.

Miscellaneous

As part of our aim to keep forecasts easily accessible and transparent, we visualize them on the following website, based on R Shiny:

`https://probability-forecasting.shinyapps.io/macropi/`

We reserve the right to change any part of the visualization during the time of this study, including the URL. The current URL will be clearly linked to in the aforementioned GitHub repository (Becker et al., 2023).

References

Becker, F., F. Krüger, and M. Schienle (2024): "Simple Macroeconomic Forecast Distributions for the G7 Economies," Preprint, arXiv:2408.08304.

Becker, F., F. Krüger, and M. Schienle (2023): "Macroeconomic prediction intervals - GitHub repository," Available at `https://github.com/KITmetricslab/MacroPI/`.

Bracher, J., E. L. Ray, T. Gneiting, and N. G. Reich (2021): "Evaluating epidemic forecasts in an interval format," *PLOS Computational Biology*, 17, 1–15.

Del Negro, M. and G. E. Primiceri (2015): "Time varying structural vector autoregressions and monetary policy: A corrigendum," *The Review of Economic Studies*, 82, 1342–1345.

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

International Monetary Fund (2024): "World Economic Outlook Database," Data retrieved from International Monetary Fund. Downloaded via code available at `https://github.com/fredbec/imfpp`, using direct download link `https://www.imf.org/-/media/Files/Publications/WEO/WEO-Database/WEOhistorical.ashx`.

Primiceri, G. E. (2005): "Time varying structural vector autoregressions and monetary policy," *The Review of Economic Studies*, 72, 821–852.

Raftery, A. E. (2016): "Use and communication of probabilistic forecasts," *Statistical Analysis and Data Mining*, 9, 397–410.