

Superforecasting the Fed's Target Range

A look back over the last 6 months

Chris Karvetski, PhD
Senior Data and Decision Scientist, Good Judgment Inc
karvetski@goodjudgment.com
14 June 2023

Executive Summary

The Federal Reserve's target range for the federal funds rate is the single most important driver in financial markets. Anticipating inflection points in the Fed's policy has immense value. This paper compares the forecasts of two groups—those produced by Good Judgment Inc's Superforecasters and those from the futures markets using the CME FedWatch Tool¹—on the outcomes of four Federal Reserve meetings from 1 February 2023 through 14 June 2023. We find that the Superforecasters were 66% more accurate than the CME FedWatch Tool forecasts when using Brier scoring (Figure 1), and the Superforecasters anticipated the June inflection point before the CME FedWatch Tool.

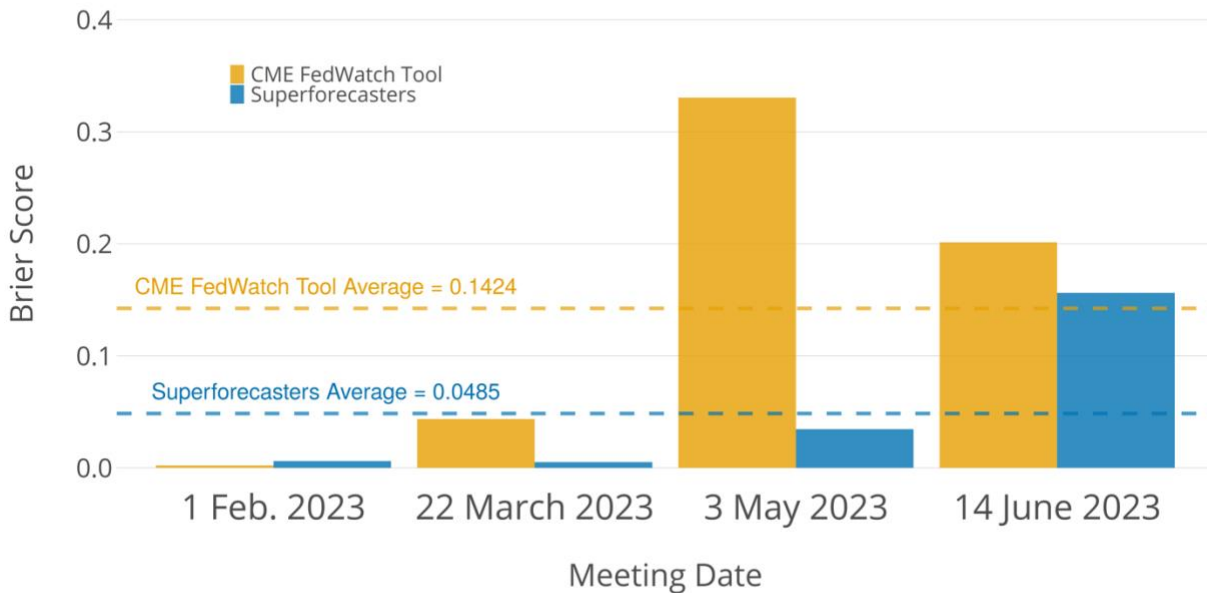
Background

Between March 2020 and January 2022, in response to the Covid-19 pandemic, the Federal Reserve kept its target range at 0.00-0.25. Since then, to cool the economy and fight persistent inflation, the Fed would raise the target range at the next seven meetings, eventually landing at 4.25-4.50 on 14 December 2022. The start of 2023 brought expectations of early hikes during the 31 January-1 February meeting, but also speculation about when the Fed might hit pause. Fed watchers' predictions would be tested over the next few months with a mixed bag of signals—strong economic data and persistently high inflation readings, on the one hand, and the collapse of banks such as Silicon Valley Bank and Signature Bank, on the other. Throughout this period, Good Judgment's Superforecasters have been beating the futures markets, signaling the Fed would continue to hike until a June pause, while markets and experts alike flipflopped on their calls. The Fed's decision to pause at its 14 June meeting gives the forecasting world a chance to pause as well and evaluate accuracy over the last six months.

¹ <https://www.cmegroup.com/markets/interest-rates/cme-fedwatch-tool.html>

Keeping Score: Futures and Supers

Brier scores of the intermeeting probability forecasts that the FOMC would cut, hold, or raise the target range.



Lower Brier scores imply greater forecasting accuracy.
Data sourced from Good Judgment Inc and the CME FedWatch Tool.

Figure 1. Accuracy results of forecasting the last four Federal Reserve meetings.

Data

Our dataset consists of daily forecasts from both the Superforecasters and the CME FedWatch Tool for the last four meetings held between 1 February 2023 and 14 June 2023. For each day of our dataset, we had two forecasts for the immediate upcoming meeting: one from the Superforecasters and one from the CME FedWatch Tool. Once a meeting occurred, the daily forecasts would then imply predictions for the outcome of the next meeting. The Superforecasters’ daily forecasts consisted of a trio of probabilities that summed to 100% on the outcomes of “Lower,” “Higher,” and “Same,” which reflected whether the Federal Reserve would cut, raise, or keep the target range the same as the last meeting.

The CME FedWatch Tool provided probabilities on partitions of the target range, from “0.00-0.25” through “12.00-12.25” in quarter-point increments. The CME FedWatch Tool probabilities also summed to 100%, and with knowledge of the current target range—which was set at the previous meeting—could be translated to equivalent outcomes of “Lower,” “Higher,” and “Same”. For example, in predicting the outcome of the 14 June 2023 meeting, the current target range was “5.00-5.25” as set at the 3 May 2023 meeting. For all forecasts after 3 May, the CME forecast probability for “5.00-5.25” for the 14 June 2023 meeting reflected the forecast on the “Same” outcome. Likewise, the sum of forecast probabilities on outcome ranges less than “5.00-5.25” (e.g., “0.00-0.25” through “4.75-5.00,” inclusive) reflected the forecast on “Lower,” and the sum of forecast probabilities on outcome ranges greater than “5.00-5.25” reflected the forecast on “Higher.”

To track accuracy, the corresponding daily forecasts on the three outcomes were scored using Brier scores, which are oriented such that lower scores imply greater accuracy. Brier scores are a common way of tracking and comparing forecasting accuracy. To calculate each daily Brier score, the two groups' daily trio of forecasts for the three outcomes were compared with the actual outcome (more technically, the daily Brier score is a summation of the squared differences between the outcome and respective forecasts). For example, if a forecast was {0%, 20%, 80% } for “Lower,” “Higher,” and “Same,” and the Federal Reserve left rates the same during the meeting being forecasted, the daily Brier score would be:

$$(0-0)^2 + (0.2-0)^2 + (0.8-1)^2 = 0.00 + 0.04 + 0.04 = 0.08.$$

To get an accuracy score for each meeting, the daily Brier scores were averaged over all forecasting days that preceded the meeting (~46 days of forecasting on average between meetings). Finally, to benchmark performance over the entire six months, we averaged these values over all four meetings.

Brier scores reward both correctness and stability in forecasting, two qualities that financial and investment decision-makers should appreciate. With regards to correctness, if a competing forecast was {0%, 10%, 90% } and was made on the same day and for the same meeting as in the example above, the Brier score for this second forecast would be 0.02, which is better than 0.08 since there is more probability on the eventual correct outcome (i.e., 90% > 80%). In terms of stability, imagine one forecasting system gave a forecast of {0%, 20%, 80% } for two consecutive days for the same meeting in the example above, yielding a two-day average Brier score of 0.08. Furthermore, imagine a second forecasting system gave a forecast of {0%, 10%, 90% } on the first day and {0%, 30%, 70% } on the second day. The second forecasting system has a two-day average forecast of {0%, 20%, 80% }—which is the same as the first system's average forecast—but the first day's Brier score is 0.02 and the second day's Brier score is 0.18, yielding a two-day average of 0.10. This value is greater, meaning *worse*, than the score of the first forecasting system (0.10 > 0.08). In other words, all else equal, over-reacting and over-correcting to information and events will hurt forecasting accuracy when using Brier scoring.

Results

The forecasting accuracy results for the four meetings of 2023 are shown in Table 1. The first four columns describe the details of each meeting, with the focal meeting date, the target range from the prior meeting, the new resulting target range, and the directional outcome. The fifth and sixth columns show the average daily probability from the CME FedWatch Tool and the Superforecasters, respectively, on the (eventual) correct outcome (whether “Lower,” “Higher,” or “Same”), and the standard deviation of these probabilities. The final two columns show the average daily Brier scores for each meeting for both the CME and the Superforecasters. Finally, the averaged Brier scores over all four meetings are shown in the bottom right: 0.1424 for the CME FedWatch Tool and 0.0485 for the Superforecasters.

The Superforecasters were 66% more accurate than the CME over the last four meetings, with the most noticeable improvements during the last three meetings (from 3/22/23 to 6/14/23) where there was the most uncertainty about the Federal Reserve's decision-making. In these

latter cases, the Superforecasters generally assigned higher probabilities to the correct outcome and had lower noise in their forecasts (as measured by standard deviation) than the CME forecasts.

Table 1. Forecasting accuracy results.

Meeting Date	Prior Range	Resulting Range	Target Range Change	CME Avg. Correct (SD)	Supers' Avg. Correct (SD)	CME Brier	Supers' Brier
2/1/23	4.25-4.50	4.50-4.75	Higher	100% (1%)	96% (3%)	0.0001	0.0041
3/22/23	4.50-4.75	4.75-5.00	Higher	94% (13%)	97% (3%)	0.0416	0.0032
5/3/23	4.75-5.00	5.00-5.25	Higher	66% (22%)	90% (8%)	0.3285	0.0325
6/14/23	5.00-5.25	5.00-5.25	Same	73% (17%)	73% (7%)	0.1994	0.1542
						0.1424	0.0485

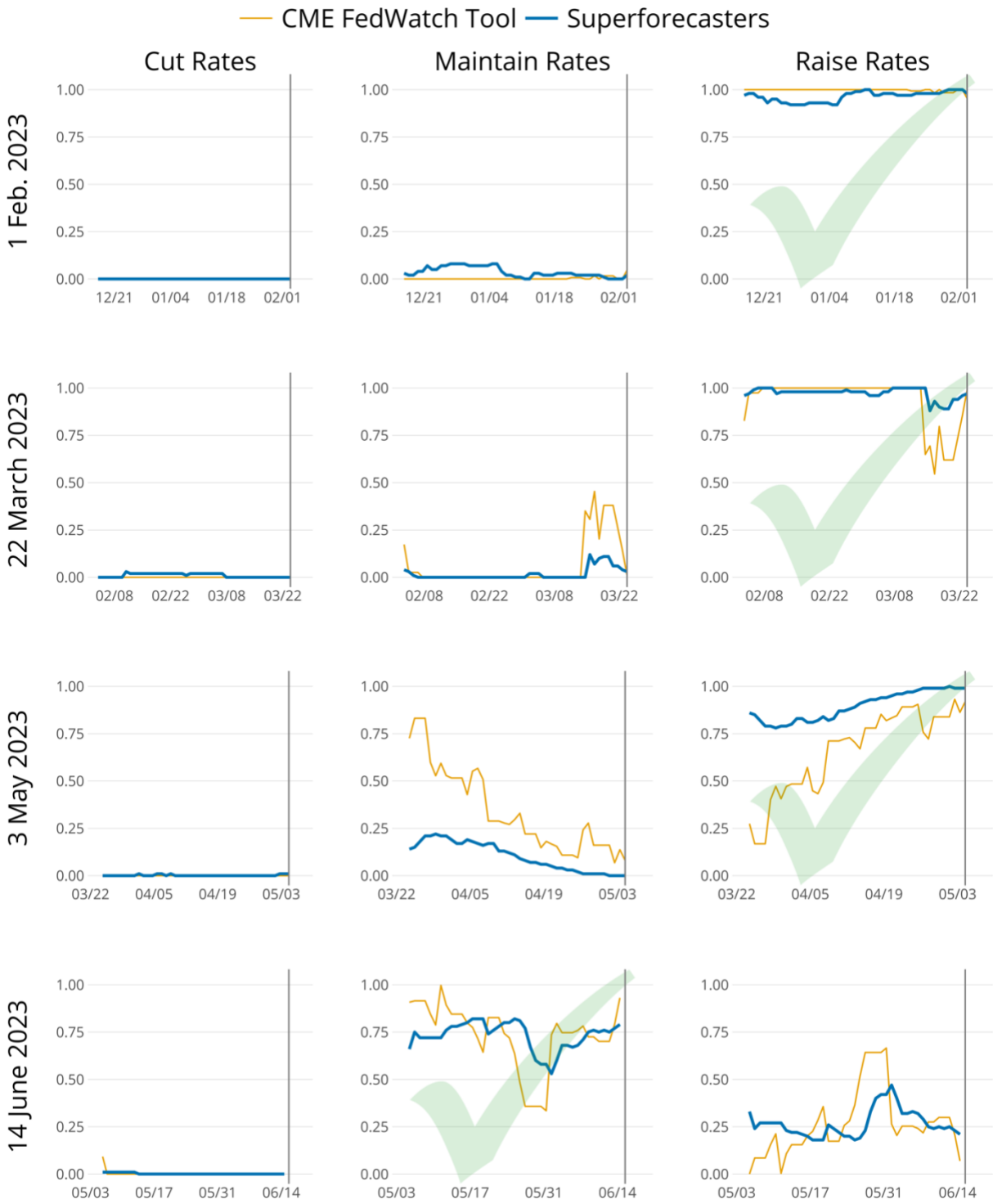
We can better understand the forecasting accuracy results by zooming in on particular meeting forecasts. Figure 2 shows the forecast probabilities from the CME FedWatch Tool and the Superforecasters for each of the three considered scenarios over the last four meetings. The first row indicates that most forecasters—both CME and Superforecasters specifically, but also a majority of prognosticators in the financial space—expected another hike to start 2023.

For the meeting to be held on 22 March 2023—the second row of the figure—both groups started out with high forecast probabilities for a raise, but the CME forecasts dropped significantly on 10 March, when Silicon Valley Bank collapsed, and would bottom out at 55% on 15 March. The Superforecasters, on the other hand, took time to process the news, and only lowered their forecast for another hike to 88% before converging back to a near-certainty of another raise.

In the third row of Figure 2, the CME forecasts did not initially expect another raise for the May meeting. The CME forecasts were less than 50% coming out of the March meeting, and did not eclipse 50% until well into April. Conversely, the Superforecasters never dropped below 78% in forecasting another raise of the target range for the May meeting, and their forecasts on the correct outcome were always higher than the CME forecasts. The May meeting—where the CME forecasts jumped the gun in predicting a pause while the Superforecasters correctly predicted another hike—would reflect the largest difference between the two groups in terms of Brier score.

Finally, the last row of Figure 2 shows the CME forecasts started high on a pause for the 14 June 2023 meeting, yet switched to a raise near the end of May, before flipping again to pause. The Superforecasters, on the other hand, favored the pause (i.e., greater than 50% on pause) for the entirety of the forecasting period.

Superforecasting® the Federal Reserve's Decision-Making



Data sourced from Good Judgment Inc and the CME FedWatch Tool.

Figure 2. Forecast probabilities for the last four Federal Reserve meetings.

Conclusion

When comparing the forecasts of two groups—Good Judgment’s Superforecasters and the futures markets using the CME FedWatch Tool—for the last four Federal Reserve meetings, we find that the Superforecasters assigned higher probabilities to the correct outcome. They were 66% more accurate than the futures markets (as measured by Brier scores) and had lower noise in their forecasts (as measured by standard deviation). An overview of the financial media during this period presents a similar picture:

- Ahead of the Fed’s March meeting, when Silicon Valley Bank collapsed, leading market observers like Goldman Sachs² said the Fed would pause, and Nomura³ said they would start to cut rates at that meeting.
- Experts like Pimco’s former chief economist Paul McCulley⁴ prematurely predicted that the Fed would go on hold in May. As the date of the meeting approached, the futures markets—as well as most market participants—came to share the Superforecasters’ view that another hike was in the cards, but lagged behind the Superforecasters by weeks.
- In the weeks heading into the June meeting, the futures were oscillating between a pause, a hike, and—during the first few days of May—possibly even a cut. Stronger economic data led experts such as Mohamed El-Erian⁵ to forecast that the Fed would continue to raise rates for at least another meeting and perhaps longer.

Good Judgment’s Superforecasters have been providing a clear signal on the Fed’s policy well before the futures and many market participants. This paper finds that the Superforecasting method is an effective tool for decision-makers. At the same time, the application of Brier scoring presents a promising avenue for comparing accuracy of various forecasters of the Fed’s policy and other consequential outcomes.

² <https://www.forbes.com/sites/siladityaray/2023/03/13/svb-collapse-fallout-goldman-analysts-forecast-no-fed-rate-hike-in-march/>

³ <https://www.bloomberg.com/news/articles/2023-03-13/nomura-predicts-rate-cut-and-qt-halt-at-upcoming-fed-meeting#xj4y7vzkg>

⁴ <https://www.cnbc.com/2023/04/11/fed-rate-hikes-will-end-next-month-former-pimco-economist-mcculley.html>

⁵ <https://www.youtube.com/watch?v=PmMhN1LNZG4>