

Chapter 14

Survey-Based Forecast Combination

14.1 Survey-Based Point Forecast Combination

A number of groups regularly survey economic and financial forecasters and publish “consensus” forecasts, typically the mean or median – essentially the average! – of the forecasters surveyed. (The median has some desirable robustness to outliers.) The consensus forecasts often perform very well relative to the individual forecasts.

The Survey of Professional Forecasters (SPF) is the leading U.S. consensus macroeconomic forecast. It has been produced each quarter since the late 1960s; currently it is produced by the Federal Reserve Bank of Philadelphia. A similar Survey of Professional Forecasters for Europe has been produced each quarter since the late 1990s; it is produced by the European Central Bank.

Another leading U.S. consensus forecast is the Livingston Survey, which is now also maintained by the Federal Reserve Bank of Philadelphia. It is only bi-annual but has been recorded for more than half a century. There are also many surveys done in the private sector.

14.1.1 Surveys and the Wisdom of Crowds

As emphasized in Surowiecki's *Wisdom of Crowds* (Surowiecki (2004)), wise “crowdsourcing” depends on balanced aggregation across disparate information sources. So we need: (1) independent, or at least imperfectly dependent, people, so that there's actually something to aggregate, and (2) a dispassionate aggregation mechanism, so that we avoid “groupthink.” Surveys are often good at (1), and certainly they're very good at (2). Other more exotic dispassionate aggregation mechanisms include Google's “pagerank” algorithm and open-source software coding (e.g., Linux).

14.1.2 Delphi, Focus Groups, and Related Methods

The “Delphi method” is a forecasting technique that sometimes proves useful in very difficult forecasting situations not amenable to quantification, such as new-technology forecasting. The basic idea is to survey a panel of experts anonymously, reveal the distribution of opinions to the experts so they can revise their opinions, repeat the survey, and so on. Typically the diversity of opinion is reduced as the iterations proceed. However, Delphi may be problematic insofar as it is actually *based* on groupthink. “Focus groups” maybe even worse, as certain individuals may dominate the group. At the same time, it's not clear that we should *dispense* with such techniques; they may be of some value.

14.1.3 Cross-Sectional Forecast Dispersion vs. True Uncertainty

The two are very different, and in principle unrelated, even if they are often positively correlated in practice.

In particular, the cross-sectional distribution of survey point forecasts is *not* a density forecast, combined or otherwise. Rather, it's simply the cross-sectional distribution of survey point forecasts. Density forecasts, combined

or otherwise, cannot generally be obtained from surveys of point forecasts. For that, we need a survey density, not point, forecast from each participant, from which a combined survey density forecast may be constructed.¹ (See Chapter ***).

14.2 Survey-Based Density Forecast Combination

14.3 Exercises, Problems and Complements

1. Wiki Surveys.

See [Wiki Surveys](#). It's not really a survey; rather, it's idea generation by pairwise comparison. Very interesting and evidently useful, even if naive in its methods for reconstructing preferences from pairwise rankings.

2. Issues in survey design.

(a) Time series of cross sections vs. panels.

Both the SPF and the Livingston Survey are time series of cross sections as opposed to true panels, insofar as both the number and composition of underlying forecasters has evolved over time. Other surveys like the Panel Study of Income Dynamics have true panel structure. Panel structure is preferable when possible, but it's not always possible, as with the SPF.

(b) Framing survey questions to turn individual responses into combined forecasts.

[Rothchild and Wolfers \(2013\)](#) makes the interesting observation that election surveys are more accurate when the respondents are not asked for whom they intend to vote, but rather whom they expect to win. A natural interpretation is that each response to the latter

¹There are, however, conditions under which the cross-sectional distribution of point forecasts can be interpreted as a density forecast. See ***.

survey is actually an average response over the respondent's friends, thereby making the effective sample size much larger than the nominal size N (more like $10N$, say). That is, each response in the latter survey is not a forecast, but rather a combined forecast.

3. Using surveys to assess forecastability.

As we saw earlier in Chapter ***, forecastability involves comparing estimates of “best” forecast accuracy to “naive” forecast accuracy. The question arises as to what to use for the best forecast. A strong case can be made for using a combined forecast from a survey or market. For example, to assess GDP forecastability, we might use MSE (assuming quadratic loss) from a reputable survey of professional forecasters as “best forecast” accuracy, and a historical GDP sample variance as “naive forecast” accuracy.

4. Forecastability assessment using surveys.

One could take a *survey-based* approach, based on the predictions of competitive professional forecasters. Conditional upon the assumption that the reported forecasts are optimal, those data can be used for inferences about predictability. The survey-based approach is of interest because the information sets used by actual forecasters are likely much richer than simple univariate histories. They are surely multivariate, for example, and they also contain hard-to-quantify subjective information. The survey-based approach does rely on a crucial and disputable assumption (optimality of reported forecasts), but so too does the model-based approach (adequacy of the fitted model). The key point is that the assumptions made by the two approaches are different, and that the approaches therefore naturally complement one another.

A number of relevant surveys exist, including the former Survey of Professional Forecasters by the Federal Reserve Bank of Philadelphia (see

Croushore 1993). These surveys focus on the major macroeconomic aggregates, such as real GDP growth. It would be interesting to use those forecasts to compute survey-based estimates of predictability, and to compare the survey-based and model-based estimates.

5. Adaptive crowdsourcing.

Sometimes forecast combination, particularly when done with surveys or markets, is called “crowdsourcing.” Sometimes “adaptive crowdsourcing” is appealing.

- (a) Traditional forecast combination, but with time-varying combining weights, is a form of adaptive crowdsourcing.
- (b) Another example is the judging of a science fair, with judges re-allocated as various projects are eliminated.

14.4 Notes

Useful web sites:

Macro / Finance Surveys

[U.S. Survey of Professional Forecasters \(SPF\)](#). Quarterly.

[European Survey of Professional Forecasters](#).

[Livingston Survey](#). Goes way back, spanning many business cycles. Biennial.

[Blue Chip](#)

Micro Surveys

[Panel Study of Income Dynamics](#).

[Michigan Survey of Consumer Sentiment](#). Well-known indexes of Consumer Sentiment, Current Economic Conditions, and Consumer Expectations.

Companies

[Consensus Economics](#). Private-sector forecast combination.

[Blue Chip](#)

[Other](#)

[Wiki Surveys](#)

[HuffPost Pollster](#) discusses and aggregates polling data. Combined surveys!