Are macroeconomic density forecasts informative?

Michael P. Clements
ICMA Centre, Henley Business School, University of Reading, Reading RG6 6BA, United Kingdom

A R T I C L E   I N F O

Keywords:
Probability distribution forecasts
Aggregation
Kullback–Leibler information criterion

A B S T R A C T

We consider whether survey density forecasts (such as the inflation and output growth histograms of the USA Survey of Professional Forecasters) are superior to unconditional density forecasts. The unconditional forecasts assume that the average level of uncertainty that has been experienced in the past will continue to prevail in the future, whereas the SPF projections ought to be adapted to the current conditions and the outlook at each forecast origin. The SPF forecasts might be expected to outperform the unconditional densities at the shortest horizons, but it transpires that such is not the case for the aggregate forecasts of either variable, or for the majority of the individual respondents for forecasting inflation.

© 2017 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.

1. Introduction

There has been considerable interest in survey forecasts over recent years, driven in part by the opportunities that they offer to test theories of expectations-generation (see e.g. Andrade and Le Bihan (2013), Coibion and Gorodnichenko (2012, 2015) and Pesaran and Weale (2006), amongst many others) and to improve forecasting accuracy, either as direct forecasts themselves (see e.g. Ang, Bekaert, and Wei (2007), Clements (2015)) or as an adjunct to other forecasting models (see e.g. Wright (2013)). In addition to collecting point predictions, some surveys also elicit respondents’ subjective probability distributions, in the form of histograms, thus offering the promise of ‘direct’ measures of the forecast uncertainty (see e.g. Clements (2014), Giordani and Söderlind (2003)) and as an alternative to less theoretically satisfactory measures such as forecaster disagreement, as given by some measure of the cross-sectional dispersion of the point predictions (Zarnowitz & Lambros, 1987).

Our interest is in whether survey respondents are able to form probability assessments about the future values of key macro-variables (such as output growth and inflation) that are more accurate than ‘unconditional’ benchmark densities. Little is known about the way in which the information content of such subjective probability assessments varies with the forecast horizon. One might surmise that survey forecasters would outperform the benchmarks at short horizons, but that any advantage would dissipate as the horizon increases; however, to the best of my knowledge there is little evidence as to whether this is the case. Of interest are the performances of the aggregate distributions (i.e., averaging across individual respondents) and the individual forecasters’ assessments, i.e., whether combination (or aggregation) plays an important role.

Here, we consider the US Survey of Professional Forecasters (SPF). We regard the SPF densities as adding value if they are more accurate than the benchmarks, at least at short horizons. We begin by considering truly unconditional density forecasts as the benchmarks, assuming normality and estimating the mean and variance from the historical forecasts. However, these densities are rejected resoundingly both when we test whether they are specified correctly and when we compare them against the SPF densities, simply because the unconditional mean is a poor estimate of the conditional mean. Hence, the rejection of the truly unconditional densities is not surprising, and they do not constitute a sufficiently challenging benchmark for the SPF densities. We then refine the benchmark forecasts to provide a stiffer challenge: the forecast densities are centred on the median point predictions (of the SPF respondents), and hence draw on forecast origin information,
but the scale or dispersion is calculated from the historical variance of the forecast errors, as before. A comparison of the SPF densities to these benchmarks serves to investigate whether the SPF densities contain any useful information about the uncertainty or probable dispersion of future outcomes. That is, we shift the focus to second moments, acknowledging that survey forecasters are able to forecast first moments (as found by Ang et al. (2007)), for example.1

One might expect the SPF forecasts to outperform the unconditional densities at the shortest horizons, assuming the variances of the densities change over time in a way which is at least partly predictable. However, the relative improvements would be expected to diminish as the forecast horizon lengthens, as the role of current conditions in predicting future developments lessens. Our results suggest the opposite: the aggregate and individual histograms are rejected in favour of the benchmark densities at the shorter horizons, reflecting under-confidence by the survey respondents at within-year horizons, as was documented by Clements (2014). That is, the survey respondents tend to over-estimate the degree of uncertainty that they face when forecasting at the shorter horizons. We show that this is true both at the aggregate level and for individuals. Moreover, the misspecification is found to be systematic, at least at the level of the aggregate histograms, meaning that 'future' densities can be improved using a simple correction calculated from an in-sample or training set.

We should emphasize that this finding of under-confidence runs counter to the prevailing view in the literature on behavioural economics and finance (see for example the surveys by Hirshleifer (2001) and Rabin (1998)). Over-confidence is a long-established view. For example, in discussing over-confidence, Malmendier and Taylor (2015) refer to the statement by Smith (1776, Book 1, Chapter X) that ‘The over-weening conceit which the greater part of men have in their own abilities, is an ancient evil remarked by the philosophers and moralists of all ages’. Hence, the findings that we present are a challenge in some ways, but they are presented in the hope that they may foster further work in this area.

Our contribution is not alone; there has been earlier work. Giordani and Söderlind (2003) find that SPF respondents’ confidence intervals for annual inflation one year ahead have actual coverage rates that are markedly lower than the nominal, indicating over-confidence; however, these authors do not consider shorter horizons. Giordani and Söderlind (2006) investigate the US SPF real annual GDP (and GNP) forecasts over the period 1982–2003, and consider forecasts of the current-year annual growth rate made in each of the four quarters of the year (i.e., forecasts from one quarter to one year ahead, approximately). However, they report coverage rates for all four horizons taken together. Kenny, Kostka, and Masera (2012) consider the ECB’s SPF and find over-confidence in the respondents’ Euro area GDP growth and inflation forecasts, but this is for one- and two-year-ahead forecast horizons. Clements (2014) compares the ex ante uncertainty estimates of the SPF respondents (that is, uncertainty estimates calculated from their histograms) to ex post estimates, while Clements and Galvão, 2017 compare the survey estimates to model-based estimates. Both find under-confidence on the part of the shorter-horizon survey estimates. Clements (2014) essentially compares the actual forecast errors with those expected based on the ex ante assessments. The comparisons reported by Clements and Galvão (2017) show that econometric models provide more accurate estimates of the uncertainty than the survey forecasts. Their study is real-time, in the sense that the models’ datasets match the data available to the survey respondents at each forecast origin, and so do not benefit from a ‘look-ahead’ bias. Nevertheless, the failure of the survey respondents not to have used the modern econometric modelling techniques of Clements and Galvão (2017) is not surprising. Being out-performed by the benchmarks that we use in these paper, on the other hand, might call into question the value of the density forecasts.

We contribute to the literature on survey expectations in a number of ways: first, we assess the value of the forecast horizon ‘conditioning’ information by comparing the SPF histograms directly to unconditional, empirical distributions; second, we use different ways of assessing and comparing the SPF and benchmark forecasts; third, we consider whether survey forecasters are more skillful at assessing the probabilities attached to particular regions of density (corresponding to events of particular interest); and, fourth, we consider whether simple mechanical corrections of the survey forecasts improve their accuracies.

Our empirical investigation both considers whether the SPF and benchmark densities are specified correctly and provides a comparison of the two, without requiring either set to provide a close approximation of the truth. We are careful to check that our findings are not dependent on changes over time in the way in which the survey has been implemented, or on any mismatch between point predictions and histogram means (e.g., Engelberg, Manski, & Williams (2009)), and we also consider alternative loss functions.

Figs. 1 and 2 present a selective look-ahead to our results. For annual output growth and inflation respectively, they present time series of the aggregate densities and the outcomes (advance estimates) for (i) the year-ahead forecasts, made in response to the surveys held in the first quarters of the years 1982 to 2013 (top panels), and (ii) the one-quarter-ahead forecasts, made in response to the surveys in the fourth quarters of 1981 to 2013 (bottom panels). A simple visual inspection of the densities and the associated outcomes suggests that the short horizon forecasts are too dispersed: realizations outside the 80% interval (defined by the 90th and 10th percentiles) should occur a fifth of the time, but there are no such instances for either output growth or inflation. Moreover, realizations outside the interquartile range should occur half the time. For output growth, it appears that the actuals are well within

1 Knüppel and Schultefrankenfeld (2012) are interested primarily in assessing the informativeness of predictions of third moments (i.e., skewness) made by Central Banks. Our focus is on whether the second moment assessments are reasonably accurate as a precursor to the consideration of higher moments.
دریافت فوری
متن کامل مقاله

امکان دانلود نسخه تمام متن مقالات انگلیسی
امکان دانلود نسخه ترجمه شده مقالات
پذیرش سفارش ترجمه تخصصی
امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
امکان دانلود رایگان ۲ صفحه اول هر مقاله
امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
دانلود فوری مقاله پس از پرداخت آنلاین
پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات