International Journal of Forecasting 30 (2014) 99–117

Contents lists available at ScienceDirect

International Journal of Forecasting

journal homepage: www.elsevier.com/locate/ijforecast

Probability distributions or point predictions? Survey forecasts of US output growth and inflation

Michael P. Clements

ICMA Centre, Henley Business School, University of Reading and Institute for New Economic Thinking at the Oxford Martin School, University of Oxford, United Kingdom

ARTICLE INFO

Keywords:
Subjective density forecasts
First moments
Bayesian learning

ABSTRACT

We consider whether survey respondents' probability distributions, reported as histograms, provide reliable and coherent point predictions, when viewed through the lens of a Bayesian learning model. We argue that a role remains for eliciting directly-reported point predictions in surveys of professional forecasters.

1. Introduction

Survey respondents' subjective probability distributions of inflation and output growth are not always consistent with their corresponding point predictions, and when they differ, it tends to be in the direction of the point forecasts presenting a more favourable outlook. For the US Survey of Professional Forecasters (SPF), Clements (2009, 2010) and Engelberg, Manski, and Williams (2009) examine the relationship between the subjective probability distributions of the individual respondents, which are reported as histograms, and the respondents' point forecasts, for both real output growth and inflation. They conclude that the two match in the majority of cases, but when they are inconsistent, the point forecasts of output growth and inflation tend to suggest a rosier outlook: the output growth and inflation point forecasts are higher and lower, respectively, than measures of the central tendency derived from the subjective probability distributions, reported as histogram forecasts.[1]

Engelberg et al. (2009, p. 40) conclude that:

> '…point predictions may have a systematic, favorable bias. …agencies who commission forecasts should not ask for point predictions. Instead, they should elicit probabilistic forecasts.…'

However, Clements (2010) suggests that the point predictions are more accurate than measures of the central tendency derived from the probability distributions, when judged by the conventional squared-error loss, casting doubt on the recommendation that surveys need only elicit information on respondents' probability distributions when 'most likely' outcomes are of interest.

In this paper, we make use of the Bayesian learning model (BLM) for comparing the point predictions and the central tendencies of the probability forecasts. The Bayesian model is of course a widely-used model of expectations formation, and indicates how an individual ought to update his/her forecast as new information becomes available in certain circumstances. We use a simple learning model due to Kandel and Zilberfarb (1999), which allows for individuals interpreting the public information differently. Bayesian models have recently been applied by a number of authors, such as Lahiri and Sheng (2008, 2010), Manzan (2011) and Patton and Timmermann (2010), for example. The novelty of our approach is that we fit a learning model to each survey respondent's point and probability forecasts simultaneously, in order to enable

E-mail address: m.p.clements@reading.ac.uk.

[1] García and Manzanares (2007) find that the growth and inflation forecasts of the European Central Bank's Survey of Professional Forecasters follow a similar pattern, and Boero, Smith, and Wallis (2008) find that the same is generally true of the Bank of England Survey of External Forecasters, especially for the output growth forecasts.

0169-2070/$ – see front matter © 2013 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ijforecast.2013.07.010

us to interpret any differences between the two types of forecasts in terms of the parameters of the learning model—whether the differences arise because new information is weighted differently, or interpreted differently, when the two types of forecasts are updated. We develop formal tests of whether the respondents (taken together) update both types of forecasts in the same way, and consider whether the weight given to, and the interpretation of, the new information changes as expected as the forecast horizon shortens. We would expect that more weight would be given to new information when the forecast horizon is short, relative to when it is long. At long horizons, the respondents' beliefs about the expected long-run mean growth rate should hold sway, with current developments becoming more influential as the horizon shortens. The estimates of the BLM key parameters allow a simple assessment of whether the forecasts conform to these fairly minimal requirements. Manzan (2011) has recently fitted a BLM to point predictions, and found that these properties hold, albeit with heterogeneity in the model estimates across individuals. When we model an individual's histogram means and point predictions jointly, we find that the evolution of the means of the histograms is not explained well by the BLM, contrary to the findings for the point predictions.

Clements (2010, 2013) consider a number of possible explanations for the apparent inconsistencies between the point predictions and probability forecasts. Foremost among these are the possibility that the forecasters' loss functions may be asymmetric for the errors they make in their point predictions (with the implicit assumption that the histograms – and therefore the moments we derive from them – are a true reflection of the underlying subjective distributions). However, under this scenario, we ought to find the point predictions less accurate than the histogram means when evaluated in terms of squared-error loss (the histogram mean approximates the conditional expectation, which minimizes the squared-error loss, and the point prediction is deliberately skewed away from the conditional expectation, to the extent that over- and under-predictions are penalized differently). The evidence points in the opposite direction. Secondly, Clements (2010) considers whether the updating of the histogram forecast is delayed relative to that of the point forecast, by testing whether the difference between the histogram mean and the point prediction is systematically related to information which was known at the time when the forecasts were issued. However, the finding of a significant correlation is not readily interpretable in the absence of a model of expectations formation. That approach essentially tests an orthogonality condition. The BLM is a fully-specified model with readily interpretable parameters, and the focus is on whether these parameters are the same across the two types of forecast for a given individual. Thus, differences in the ways in which an individual updates their two types of forecasts as new information becomes available can be interpreted in terms of the parameters of the BLM.

When we estimate the BLM, we regard the mean as the central tendency of the histogram to be matched to the point predictions. Because the estimates of the mean will depend on the distributional assumptions we make, we undertake a sensitivity exercise; that is, we present results for an alternative distributional assumption. Our preferred approach is to fit generalized beta distributions to calculate moments from the histograms (following Engelberg et al., 2009). The alternative assumption which we make is to fit Gaussian distributions. For the normal distribution, the mean and mode coincide, so we could alternatively view this as interpreting the individuals' point predictions as their modal forecasts when we compare their point predictions and histograms. By and large, our qualitative findings are not sensitive to the distributional assumptions. This suggests that allowing the underlying subjective distributions to be asymmetric (by fitting a generalized beta distribution rather than a normal distribution) is not of any consequence, for our purpose, and also that the results are robust to viewing the point predictions as the modal forecasts.

Our general conclusion is that the use of best-practice methods to calculate continuous distributions from histograms gives estimates of means that do not constitute coherent sets of forecasts when viewed through the lens of a learning model. This suggests that there is a case for eliciting point predictions if we require measures of point predictions from the surveys.

The remainder of the paper is organized as follows. Section 2 describes the SPF forecast data, the calculation of means and the fitting of continuous distributions to the histograms. Section 3 updates the evidence of inconsistencies in the two types of forecast using the bounds approach of Engelberg et al. (2009). Section 4 outlines the application of the Bayesian learning model to the joint measurement of the central tendency of the individuals' probability distributions and their point predictions, and sets out the tests of the hypotheses of interest. Section 5 reports the empirical findings. Finally, Section 6 offers some concluding remarks.

2. Data description and distributional assumptions

We choose the SPF as our source of survey expectations because it contains information on respondents' probability distributions for inflation and output growth as well as their point forecasts for these key macro-aggregates, and spans a long historical period. It is a quarterly survey of professional forecasters of the US economy. The SPF began as the NBER-ASA survey in 1968:4, and is still running today: see for example Croushore (1993). The last data we use are for the fourth quarter of 2010, giving 169 quarterly surveys of expectations data spanning the last 40 years.[2]

The histograms are of the percentage change in the survey year relative to the previous year. We calculate matching year-on-year point forecasts as follows. The surveys provide point forecasts of the level of the variable in the current (survey quarter) and each of the next four quarters. We use the forecasts of the current and subsequent quarters, along with the actual values from the vintage of data

2 The data were downloaded on 15th September, 2011, and thus include the corrections released on August 12, 2011. Further details are available at http://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/.