ICMA Centre

Henley Business School
UNIVERSITY OF READING

Discussion Paper

Do US Macroeconomics Forecasters Exaggerate Their Differences?

September 2014

Michael P Clements
ICMA Centre, Henley Business School, University of Reading

Discussion Paper Number:
ICM-2014-10

The aim of this discussion paper series is to disseminate new research of academic distinction. Papers are preliminary drafts, circulated to stimulate discussion and critical comment. Henley Business School is triple accredited and home to over 100 academic faculty, who undertake research in a wide range of fields from ethics and finance to international business and marketing.

admin@icmacentre.ac.uk

www.icmacentre.ac.uk

Do US Macroeconomic Forecasters Exaggerate Their Differences?

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

September 12, 2014

Abstract

Application of the Bernhardt, Campello and Kutsoati (2006) test of herding to the calendar-year annual output growth and inflation forecasts suggests forecasters tend to exaggerate their differences, except at the shortest horizon when they tend to herd. We consider whether these types of behaviour can help to explain the puzzle that professional forecasters sometimes make point predictions and histogram forecasts which are mutually inconsistent.

JEL Classification: E37.

Keywords: macro-forecasting, imitative behaviour, histogram forecasts, point predictions.

*Michael Clements is Professor of Econometrics at the ICMA Centre, Henley Business School, University of Reading, and a member of the Institute for New Economic Thinking at the Oxford Martin School, University of Oxford. Email: m.p.clements@reading.ac.uk. The computations were undertaken using code written in Gauss 14, Aptech Systems.

1 Introduction

It is generally understood that economic forecasters may have incentives to act strategically in the sense of seeking to enhance their reputations (see, e.g., Lamont (2002), Ehrbeck and Waldmann (1996), Laster, Bennett and Geoum (1999) and Ottaviani and Sorensen (2006), *inter alia*). For example, Lamont (2002, p. 265) notes that besides minimizing squared forecast errors, agents may set their forecasts to 'optimize profits or wages, credibility, shock value, marketability, political power...'. Much of the empirical literature on testing the influence of factors other than accuracy on the determination of agents' forecasts rests on the notion of herding - whether forecasters take into account the views of others when they produce their forecasts. This may be manifest in forecasters skewing their optimal forecast towards the consensus view, or artificially exaggerating the difference between their forecast and the consensus, where optimal is to be understood in the narrow sense of maximizing the expected accuracy of the forecast (for example, minimizing the expected squared forecast error). Indeed, Lamont (2002) supposes that forecasters actual loss functions may contain terms in the difference between the forecast and the consensus, as well as conventional accuracy measures such as the absolute forecast error.

The focus of this paper is whether the respondents to the US Survey of Professional Forecasters (US-SPF) take into account the consensus view when they issue their forecasts of US output growth and inflation. Compared to studies such as Lamont (2002), where the individual forecasters are identified and their track performances are public knowledge, the SPF respondents are anonymous (although each respondent has a unique identifier, so that individuals can be tracked over time). One might suspect this would weaken the extent to which the forecasters behave strategically, but alternatively if the respondents report the same forecasts to the SPF as they make public through other spheres, these issues remain pertinent. We regard it as a matter that can only be determined by an empirical study. We wish to discover whether herding behaviour depends on the forecast horizon. Forecasters may behave differently when providing their expectations of relatively distant events compared to short-horizon forecasts. In our study the forecast horizons span one-quarter to one-year ahead forecasts. As well as exploring forecast behaviour across different horizons, we also explore behaviour by type of forecaster, and assess whether forecasters working in firms characterized as financial service providers systematically differ from those in non-financial service firms.

The behaviour of the forecasters that take part in the US-SPF is of interest in itself as part of the endeavour to better understand the actual expectations formation process of economic agents, given the pre-eminence of the US-SPF. [1] But in addition of particular interest is whether strategic behaviour is responsible in part for the inconsistencies between the respondents' reported probability distributions and point predictions, as documented by Engelberg, Manski and Williams (2009) and Clements (2009, 2010).

Testing for herding is not straightforward, as individuals' forecasts will tend to cluster together to the extent that they share the same information irrespective of whether they consider the consensus view when they form their forecasts. The influential approach of Lamont (2002) does not seek to establish whether forecasters herd or scatter (anti-herd) - i.e., whether they downplay or exaggerate their differences, but tests whether the pattern varies over the forecaster's lifetime. He finds significant evidence that herding changes as agents age: older forecasters become more radical. Lamont (2002) regresses the absolute difference between each respondent's forecast and the consensus on the age of the forecaster, and the average deviation (as well as individual fixed-effects). Despite the warning that only age-related changes in the pattern of scattering/herding are detectable (when the age variable is found to be statistically significant), some have sought (incorrectly) to make inferences about herding when the age variable is insignificant: see Ashiya and Doi (2001).

We use the non-parametric test of herding of Bernhardt *et al.* (2006) which does allow for a discrimination between scattering and herding, and has several key advantages, as discussed in section 2. Section 3 describes the forecast data, and section 4 the evidence of whether these forecasters herd or exaggerate their differences. Section 5 describes the inconsistencies between the point predictions and histogram forecasts, and how we test whether these can be explained by herding. Section 6 presents the results, and finally section 7 concludes.

2 Testing procedure

Suppose a survey respondent's h-step ahead forecast distribution is given by:

$$y_\tau \mid \Omega_{\tau-h} \sim D\left(E_{\tau-h}\left[y_\tau\right], V_{\tau-h}\left[y_\tau\right]\right),$$

where $E_{\tau-h}(y_\tau) \equiv E(y_\tau \mid \Omega_{\tau-h})$, and $V_{\tau-h}(y_\tau) \equiv Var(y_\tau \mid \Omega_{\tau-h})$. If the forecaster chooses to announce their conditional expectation $(E_{\tau-h}(y_\tau))$ as their point prediction, $F_{\tau,h}$, then $F_{\tau,h} = E_{\tau-h}(y_\tau)$, and the point forecasts are conditionally (and therefore also unconditionally)

[1] The US SPF has been an important source of data for research on expectations formation. As of April 21 2014, the Academic Bibliography maintained by the Philadelphia Fed listed 79 research papers based on the SPF forecast data: see http://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/academic-bibliography.cfm.

unbiased, since $E\left(\left[y_\tau - F_{\tau,h}\right] \mid \Omega_{\tau-h}\right) = E\left(\left[y_\tau - E_{\tau-h}\left(y_\tau\right)\right] \mid \Omega_{\tau-h}\right) = 0$.

Suppose now that the survey respondent herds in the sense that his point prediction is biased towards the consensus forecast. Denote by $C_{\tau,h}$ the consensus forecast of y_τ *available* at period $\tau - h$.[2] Hence $C_{\tau,h} \in \Omega_{\tau-h}$ is in the agent's information set by assumption. Herding occurs when the forecaster biases their forecast $F_{\tau,h}$ away from $E_{\tau-h}\left(y_\tau\right)$ in the direction of $C_{\tau,h}$, whereas 'anti-herding' is when the forecast is moved further away from the consensus forecast relative to the original position.

Bernhardt *et al.* (2006) develop a non-parametric test for herding (and anti-herding) based on the sum of two conditional probabilities:

1) the probability that the reported forecast exceeds the outcome conditional on the forecast exceeding the consensus:

$$\Pr\left(y_\tau < F_{\tau,h} \mid C_{\tau,h} < F_{\tau,h}\right) \tag{CP$_1$}$$

and 2) the probability that the reported forecast is less than the outcome conditional on the forecast being less than the consensus:

$$\Pr\left(y_\tau > F_{\tau,h} \mid C_{\tau,h} > F_{\tau,h}\right). \tag{CP$_2$}$$

Consider the first of these. If the forecaster pays no attention to the consensus view, then the probability should be one half assuming that the forecasts are unbiased. If the forecaster herds, then the reported forecast will have been moved downward towards the consensus, and the probability that the forecast exceeds the actual will be less than one half. Anti-herding would result in the probability exceeding a half. The same logic can be applied to the second conditional probability, which will be less than (exceed) one half if there is herding (anti-herding). Bernhardt *et al.* (2006, p.661-3) detail the advantages of basing a test statistic on the average of the two conditional probabilities. Two of the key advantages for macroeconomic forecasting are that the average of the two probabilities will not (incorrectly) suggest herding simply because of i) the occurrence of common or aggregate shocks that affect all forecasters or ii) the presence of systematic bias (unrelated to herding).

Consider the effect of common (unobserved) shocks on the first probability, CP$_1$. In the absence of herding, $\Pr\left(y_\tau < F_{\tau,h} \mid C_{\tau,h} < F_{\tau,h}\right) = \Pr\left(y_\tau < F_{\tau,h}\right)$. But if there is a positive shock, $\Pr\left(y_\tau < F_{\tau,h}\right) < \frac{1}{2}$. However the second probability CP$_2$ is then equal to $\Pr\left(y_\tau > F_{\tau,h}\right)$ under no herding, and $\Pr\left(y_\tau > F_{\tau,h}\right) = 1 - \Pr\left(y_\tau < F_{\tau,h}\right)$. Hence the average of the two probabilities will not be affected by unforecasted shocks.

[2] In what follows the assumption of what is known to the forecaster is shown to play an important role.

Consider forecast bias not related to herding ((Bernhardt *et al.* (2006)) motivate this in the context of analysts' forecasts of company earnings, if earlier forecasts tend to be more optimistic, say). Suppose for example that the forecaster targets the $\alpha_{\tau,h}$ percentile of y_τ, that is, $F_{\tau,h}$ is set such that $\Pr(y_\tau < F_{\tau,h}) = \alpha_{\tau,h}$. However, $\Pr(y_\tau > F_{\tau,h}) = 1 - \Pr(y_\tau < F_{\tau,h}) = 1 - \alpha_{\tau,h}$ so that the average of the two probabilities is one half irrespective of the value of $\alpha_{\tau,h}$.

2.1 Test statistic

Let $\gamma_\tau^+ = 1$ if $F_\tau > C_\tau$, and $\gamma_\tau^- = 1$ if $F_\tau < C_\tau$ (note we have dropped the forecast horizon subscript for notational simplicity). Then define the joint events as $\delta_\tau^+ = 1$ if $F_\tau > C_\tau$ and $F_\tau > y_\tau$, and $\delta_\tau^- = 1$ if $F_\tau < C_\tau$ and $F_\tau < y_\tau$. The Bernhardt *et al.* (2006) test statistic S is calculated as:

$$S = \frac{1}{2}\left[\frac{\sum_\tau \delta_\tau^+}{\sum_\tau \gamma_\tau^+} + \frac{\sum_\tau \delta_\tau^-}{\sum_\tau \gamma_\tau^-}\right] \tag{1}$$

which is asymptotically normally distributed $N\left(0.5, \frac{1}{16}\left[\left(\sum_\tau \gamma_\tau^+\right)^{-1} + \left(\sum_\tau \gamma_\tau^-\right)^{-1}\right]\right)$ under the null of no herding: that the probability of an over-prediction (under-prediction) is independent of whether $F_\tau > C_\tau$ ($F_\tau < C_\tau$). Bernhardt *et al.* (2006) show that the mean is $\frac{1}{2}$ under the null of no herding irrespective of whether or not forecasts are biased (i.e., of whether or not $\Pr(F_\tau > y_\tau) = \frac{1}{2}$), but that the variance will be over-estimated when forecasts are biased, and also when where the forecast errors are correlated, so that the test statistic will be conservative. That is, the test will be under-sized so that a rejection of the null is strong evidence against the no-herding null.

3 Survey Data

The US Survey of Professional Forecasters (SPF) is our source of expectations. The SPF began as the NBER-ASA survey in 1968:4 and runs to the present day.[3] It is a quarterly survey of macroeconomic forecasters of the US economy, eliciting information on point predictions for a number of macro variables as well as the respondents' histograms for inflation and output growth. We use the SPF in conjunction with Real Time Data Set for Macroeconomists (RTDSM) run by the Federal Reserve Bank of Philadelphia (see Croushore and Stark (2001)) to re-create the data the respondents would have had access to when filing their survey returns,

[3] Since June 1990 it has been run by the Philadelphia Fed, renamed as the Survey of Professional Forecasters (SPF): see Zarnowitz (1969) and Croushore (1993).

as described below.[4]

Tests of herding are applied to the point predictions of the calendar year annual growth rate of real GDP (i.e., the percentage change between the level of output in year t relative to year $t-1$) and of the annual inflation rate. For the Q1 surveys, we sum the forecasts of the current quarter and the forecasts of the next three quarters, and divide by the data for the previous year's four quarters, taken from the RTDSM.[5,6] For Q2 surveys the approach is the same except the value for the preceding quarter (Q1) is now data, and similarly for surveys made in the third and fourth quarters of the year. So we have forecasts of annual inflation made in Q1 through to Q4 of the year. For the first-quarter surveys the forecast horizon is just under a year, whereas for fourth-quarter surveys it is just under one quarter.

The SPF also provides histograms of annual inflation and output growth in the current year relative to the previous year. The histograms match the point predictions in terms of forecast target (the annual change) and the forecast horizon, and we discuss the use of these forecasts in section 5.

For both variables we have calendar year point predictions from the 181 quarterly surveys from 1968:Q4 to 2013:Q4. Inflation histogram forecasts are available over the same period, but the real output growth histograms are only from 1981:Q3 onwards.[7]

There are a number of advantages to using these forecasts for our purposes. Firstly, the forecasts are of the (calendar) year-on-year real GDP growth rate and inflation rate. These are clear, unequivocal measures of activity and prices and forecasts of these measures are very much in the public eye.

Secondly, they have a fixed-event dimension. This means there are multiple forecasts of the same target event made at different points in time (equivalently, forecasts of the same target

[4]Later vintages of data contain revisions and definitional changes (see e.g., Landefeld, Seskin and Fraumeni (2008) for a discussion of the revisions to US national accounts data).

[5]As of 1981:3, forecasts of the level of the output price for the current year were provided. Summing the quarterly forecasts allows us to use data back to 1968:4.

[6]The point forecasts of the growth rate are calculated using the actual data for the previous year from the RTDSM available in the quarter of the survey. The one exception is that the RTDSM for 1996Q1 is missing the value for 1995Q4. In constructing the year-on-year point forecast growth rates for the respondents to the 1996Q1 survey we use the previous-quarter forecasts (of 1995Q4).

[7]Prior to 1981:3 the point predictions for output referred to nominal output, but a series for real output has been imputed (by the Philadelphia Fed) from the forecasts of nominal output and the deflator. The results reported in the paper make use of the imputed real growth forecasts. But probability distributions for real output growth are not given for the period before 1981:Q3. Hence when the computations require both point predictions and histograms the available sample is reduced. In addition, the reliability of the survey data for some survey quarters is questionable. The online documentation provided by the Philadelphia Feb: 'Documentation for the Philadelphia Fed's Survey of Professional Forecasters', http://www.phil.frb.org/econ/spf/. The problematic survey quarters are 1968.4, 1969.4, 1970.4, 1971.4, 1972.3, 1972.4, 1973.4, 1975.4, 1976.4, 1977.4, 1978.4, 1979.2, 1979.3 and 1979.4 for inflation, and 1985.1 and 1986.1 for both variables.

made at different horizons). Hence we can use as the consensus for forecasting period τ at time $\tau - h$ an average of the individual forecasts of τ made in the previous quarter, $\tau - h - 1$. It may be reasonable to assume that individual forecasters are aware of the current forecasts made by their fellow forecasters, and produce their forecasts cognizant of the prevailing view, given the media coverage accorded to the forecasts and pronouncements on the current state and likely evolution of key macro indicators.[8] We report the results of calculating the consensus based on last quarter's forecasts and the current quarter's forecasts. The former strategy is more conservative, and we discuss potential problems with the second strategy when we consider the results.

Formally, we let $F_{i\tau,h+1} = E\left(y_\tau \mid \mathcal{I}_{i,\tau-(h+1)}\right)$ denote each individual i's forecast (where $i = 1, \ldots, N_\tau$) of y_τ made at time $\tau - (h+1)$. We take the consensus to be the cross-sectional median as is commonly done, and denote this by $C_{\tau,h,-1}$, where the '-1' subscript indicates it is the previous quarter's consensus of y_τ (relative to survey quarter $\tau - h$). The current consensus is denoted $C_{\tau,h,0}$, and we will continue to use $C_{\tau,h}$ when we do not need to discriminate between the two. The key question of interest remains whether survey respondents' forecasts made at $\tau - h$ (of y_τ) are influenced by the consensus (either $C_{\tau,h,0}$ or $C_{\tau,h,-1}$).

The SPF allows us to identify forecasters by the industry they work in, beginning with the 1990:Q2 survey onwards. Individuals are classified as either working in a firm characterized as a financial service provider, or a non-financial service provider, with a third category for those for whom insufficient information is available to make the designation. A number of papers have considered whether various aspects of forecasters' use of information differs by forecaster 'type'. For example, Carroll (2003) argues that consumers acquire new information more slowly than professional forecasters, whilst Coibion and Gorodnichenko (2012) find that information acquisition and processing does not differ systematically across consumers, firms, central bankers and professional forecasters. Our comparison between different 'types' of professional forecasters might be thought less likely to throw up significant differences in behaviour than between consumers and professional forecasters, for example, but nevertheless it is an empirical question whether financial and non-financial forecasters' herding behaviour differs.

In addition to the point forecasts of the calendar year annual growth rates, the probability distribution forecasts of these same quantities (expressed in the form of probabilities of inflation (say) lying in certain pre-specified intervals) allows us to calculate higher forecast moments of these two variables. There is no reason in principle to confine an investigation of herding

[8] For example, a respondent to a fourth quarter survey filing their forecast of annual inflation in mid November is likely to have more up-to-date information on the expected annual inflation rate than the consensus view given by the survey responses made in mid August.

to first moments: if survey respondents are uncertain about the uncertainty surrounding their central projections, they might well take a lead from the consensus level of uncertainty (as given in the previous quarter's aggregate histogram, for example, or as the average of the individual respondents' forecast standard deviations). However the approach to testing for herding requires actual values against which the forecasts can be compared. Estimates of the actual values of conditional variances or standard deviations could be obtained from squared forecast errors or realized variance measures using higher-frequency data,[9] but we do not consider these issues here, and instead confine our attention to first moment estimates for which actual values are readily available.

4 Empirical findings: Evidence of herding when SPF respondents report their point predictions

The results are shown in table 1. Consider the results when we use the current consensus, $C_{\tau,h,0}$. The results suggest anti-herding: forecasters deliberately move their reported forecasts further from the consensus, that is, they exaggerate the difference between their forecast and the consensus. This is true of both inflation and output growth, and holds at all forecast horizons: 1-year ahead, down to 1-quarter ahead.[10] The lower bound on the statistic (1) in all cases exceeds one half, indicating that the finding is statistically significant (in a two-sided test at the 5% level). Across all horizons, the sum of the two probabilities is 0.68 for inflation, and 0.64 for output growth.

When the consensus is calculated from the previous quarter's forecasts, $C_{\tau,h,-1}$, the sums of the conditional probabilities are generally closer to a half, suggesting the SPF respondents pay less attention to the gap between their current forecasts and the lagged consensus when setting their current forecasts. Nevertheless, for inflation we still reject the null in favour of anti-herding for the second and third quarter forecasts, and interestingly, find evidence of herding at the shortest horizons (corresponding to the fourth quarter surveys). For output growth there is also evidence of anti-herding in responding to the second and third quarter surveys, but again, clear evidence of herding at the shortest horizon.

The results for the previous quarter's consensus provide the more conservative test of (anti-)herding. They may under-estimate each forecaster's knowledge of the views of others: failure to

[9] For example, Andersen and Bollerslev (1998) describe the practice of using squared returns to proxy actual return volatility, and there is a large literature on the use of realized volatility measures (see, e.g., Andersen, Bollerslev, Diebold and Labys (2001, 2003)).

[10] The two component probabilities of the S-statistic defined in (1) exceed a half in all cases.

reject the non-herding null might simply reflect the fact that the lagged consensus is out-dated information. But by the same token, these results are unlikely to falsely indicate a dependence on the views of others by overstating the forecaster's information set.

The results for the lagged consensus suggest that herding behaviour depends on the forecast horizon. For forecasting both variables, there is a tendency to move towards the consensus for the shortest horizon (one-quarter ahead), but at longer horizons forecasters tend to exaggerate their differences. There is a literature on the dispersion of forecasters' expectations, and how this varies with the forecast horizon (see Lahiri and Sheng (2008) and Patton and Timmermann (2010), *inter alia*). For example, Patton and Timmermann (2010) show that dispersion is greater at longer horizons. Our findings on herding are broadly consistent with these patterns, but we leave a more formal analysis of herding and disagreement for future research.

Tables 2 and 3 report results separately for forecasters who can be characterized as working for financial services firms and for non-financial services firms, for the sub-sample of surveys from 1990:Q2 for which identification of forecaster type is possible. The results for both inflation and output growth suggest that the differences between the two types of forecaster are everywhere small so that there is no evidence that the two differ in terms of their herding behaviour.

5 Herding and forecast inconsistencies

Engelberg *et al.* (2009) and Clements (2009, 2010) compare the point predictions of the US SPF respondents with the histogram forecasts. They report inconsistencies between some of the pairs of forecasts. Of interest is whether the evidence of anti-herding we find (as reported in section 4) explains the inconsistencies between the different types of forecasts. We first update and explain the evidence on inconsistencies, and then provide an assessment of whether (anti-)herding offers an explanation.

To assess the evidence for inconsistencies, we use the non-parametric bounds approach of Engelberg *et al.* (2009) to determine whether each pair of histogram and point forecasts is consistent, in the following sense. We calculate whether the point forecast is consistent with a given measure of central tendency of the histogram (such as the mean) using only the information provided in the histogram: that is, without introducing any auxiliary assumptions about how the histogram relates to the underlying subjective probability distribution. We calculate upper and lower bounds on the moments from the histogram to determine whether the point prediction is consistent with that moment of the underlying subjective distribution: see e.g., Engelberg *et al.* (2009) for details.

Table 4 reports the percentages of point forecasts which are within, below and above the bounds calculated on the mean, median and mode. The results aggregate over all respondents and all time periods, but are presented separately for each quarter to allow for differences because of the length of the forecast horizon. The results broadly confirm those of earlier studies on shorter samples. Consider the forecasts made in the first-quarters of each year, corresponding to approximate 4-quarter (or year-ahead) forecasts. Around one quarter of the inflation forecast pairs, and one fifth of the output growth forecast pairs, are inconsistent when the point forecast is interpreted as the mean. A preponderance of the inconsistencies are in the direction of the point forecasts being more optimistic (point forecast exceeds the upper bound for output growth, is below the lower bound on the moment for inflation). The degree of inconsistency generally diminishes as the horizon shortens (going from the Q1 to Q4 surveys), but for inflation is still around one fifth for the third-quarter surveys (corresponding to an approximate half-year forecast horizon). The qualitative nature of these findings holds for the mode and median. Under the conservative approach of requiring only that the point forecast lies within the bounds on any one of the three measures of central tendency to deem the two mutually consistent, we still reject consistency 14% and 10% of the time for inflation and output growth, respectively, one-year ahead.

We investigate whether (anti-)herding is a possible explanation of the inconsistencies recorded in table 4. We base the test of Bernhardt *et al.* (2006) on the sum of two conditional probabilities:

1) the probability that the reported point prediction (F) exceeds the histogram moment (which we take to be the mean, μ) conditional on the point prediction exceeding the consensus,

$$\Pr\left(F_{\tau,h} > \mu_{\tau,h} \mid C_{\tau,h} < F_{\tau,h}\right) \qquad (\mathrm{CP}_1^*)$$

and, 2) the probability that the reported point prediction is less than the histogram mean conditional on the forecast being less than the consensus,

$$\Pr\left(F_{\tau,h} < \mu_{\tau,h} \mid C_{\tau,h} > F_{\tau,h}\right). \qquad (\mathrm{CP}_2^*)$$

Relative to CP_1 and CP_2 we have simply replaced the actual values by the histogram means. Unlike the bounds approach to assessing forecast inconsistencies, we now need point estimates of the histogram means. We consider two ways of calculating the histogram means: we estimate the means directly (by assuming the probability mass is uniform within each interval (equivalently, lies at the midpoint) as well as fitting a parametric distribution to the points on the distribution identified by the histogram. When there is a large difference in the probability mass attached to adjacent intervals, it might be thought desirable to attach higher probabilities to points in the lower interval near the boundary with the high probability interval: this is facilitated by fitting a parametric distribution. Rather than fitting a normal distribution with the assumption of symmetry, we follow Engelberg *et al.* (2009) amongst others and fit (unimodal) generalized beta distribution. This distribution uses two parameters to describe the shape of beliefs, and two more to give their support. Details of our approach are given in Appendix A. The findings are qualitatively similar for both approaches, suggesting the way of calculating the means is not driving the results. To save space we only report results for the parametric approach.

Notice that we suppose that if herding does take place it is based on a comparison of the point prediction with the consensus *point* prediction (not the consensus *mean* forecast). This assumption reflects the greater visibility of the point predictions (these are recorded in the survey) whereas the mean forecasts are implicit in the reported histograms.

Suppose that there are aggregate shocks. The probability that $F_{\tau,h} > \mu_{\tau,h}$ is less likely to be affected by unexpected aggregate shocks than the probability that $F_{\tau,h} > y_\tau$, because both the point prediction and the histogram mean will both fail to reflect the shock. However, if there are information rigidities which affect the two types of forecasts differently, then in response to a sequence of positive shocks (say) the point predictions would exceed the means, if it were the case that the point predictions are updated more frequently. Coibion and Gorodnichenko (2012) provide a recent discussion of information rigidities. By summing the two conditional probabilities CP_1^* and CP_2^* we obtain a test statistic which will not reject simply because of informational rigidities. To see this, suppose $\Pr\left(F_{\tau,h} > \mu_{\tau,h} \mid C_{\tau,h} < F_{\tau,h}\right) = \Pr\left(\mu_{\tau,h} < F_{\tau,h}\right)$ (i.e., no herding), but $\Pr\left(F_{\tau,h} > \mu_{\tau,h}\right) > \frac{1}{2}$, say, assuming a sequence of positive shocks and more frequent updating of $F_{\tau,h}$ relative to $\mu_{\tau,h}$. But then $\Pr\left(F_{\tau,h} < \mu_{\tau,h}\right) = 1 - \Pr\left(F_{\tau,h} > \mu_{\tau,h}\right)$ so the average probability will still be equal to $\frac{1}{2}$ under the null.

Suppose the point predictions are biased for reasons other than herding, such as the point prediction being optimal for an asymmetric loss function.[11] If we let $F_{\tau,h} = \mu_{\tau,h} + x_{\tau,h}$, then under some assumptions, $x_{\tau,h} = \phi_h \cdot \sqrt{V_{\tau-h}(y_\tau)}$, where ϕ_h is a constant which depends on the distribution function of the data and the loss function, and $V_{\tau-h}(y_\tau)$ is the conditional variance: see Patton and Timmermann (2007, Proposition 2). Clements (2010, 2013) investigate whether

[11] Granger (1969) and Zellner (1986) draw attention to the fact that optimal forecasts will be biased if the loss function is asymmetric. Contributions such as Elliott, Komunjer and Timmermann (2005), Elliott, Komunjer and Timmermann (2008), Patton and Timmermann (2007) and Lahiri and Liu (2009), inter alia, consider whether forecasts are rational once an allowance is made for asymmetric loss.

asymmetric loss accounts for the inconsistencies reported in table 4, but find little support for the contention. Our interest here is whether (anti-) herding explains the inconsistencies, and whether we can use the testing approach in section 2 to test this hypothesis in the presence of asymmetric loss. It turns out we can: it is straightforward to establish that the sum of the conditional probabilities CP_1^* and CP_2^* will be one half irrespective of whether loss is asymmetric. Under the null - that herding on the consensus point predictions does not explain the sign of the gap between the individuals' point predictions and mean forecasts - the conditional probability CP_1^* is simply $\Pr\left(F_{\tau,h} > \mu_{\tau,h}\right)$ which is equal to $\Pr\left(x_{\tau,h} > 0\right)$. Now, $\Pr\left(x_{\tau,h} > 0\right) \lesseqgtr \frac{1}{2}$ depending on whether $\phi_h \lesseqgtr 0$, where $\phi_h \neq 0$ when the costs of under- and over-prediction are different. However, under the null CP_2^* is $\Pr\left(F_{\tau,h} < \mu_{\tau,h}\right) = \Pr\left(x_{\tau,h} < 0\right) = 1 - \Pr\left(x_{\tau,h} > 0\right)$, so that $\frac{1}{2}(\mathrm{CP}_1^* + \mathrm{CP}_2^*) = \frac{1}{2}$ irrespective of the degree and direction of the asymmetry in the loss function.

6 Does herding explain forecast inconsistencies?

Table 5 shows that the null (that the S-type statistic equals a half) is rejected at the 5% level for the second and third quarter survey forecasts for both output and inflation, indicating that - at all but the shortest horizon - the discrepancies between the point predictions and means are systematically related to herding of the point predictions on the consensus. These results are for last quarter's consensus, which may be more reliable for the reasons given in section 4, but both sets of results are recorded and are largely the same.

Consider the constituent probabilities for inflation for the second-quarter survey forecasts. The probability $\mathrm{CP}_1^* = \Pr\left(F_{\tau,h} > \mu_{\tau,h} \mid C_{\tau,h} < F_{\tau,h}\right)$ is 0.37, suggesting the probability of reporting a pessimistic point prediction (relative to the forecast of the histogram mean) is less than a half. The probability $\mathrm{CP}_2^* = \Pr\left(F_{\tau,h} < \mu_{\tau,h} \mid C_{\tau,h} > F_{\tau,h}\right)$ is 0.77, suggesting optimistic point predictions are more likely than not when the point forecast is below the consensus (point) forecast. Given that we have shown that optimistic point predictions out-number pessimistic forecasts 4:1 for inflation (see table 4), these findings do not come as a surprise. They are consistent with (say) a series of supply-side shocks which reduce inflation, and with slower updating of the histogram forecasts relative to the point predictions. However, the fact that the average of the two probabilities is significantly different from one half suggests that this is not the whole story: we reject the hypothesis that respondents pay no heed to the consensus view when they determine their forecasts. The fact that the test statistic exceeds one half indicates 'anti-herding'. In our context, this suggests that negative gaps (where the gap is $F_{\tau,h} - \mu_{\tau,h}$) are exacerbated when the point prediction is low relative to the consensus.

For output growth the overall S-statistics are similar, and significantly in excess of one half for the longer horizon forecasts, but the CP_1^* probabilities are now in excess of one half, consistent with the point predictions tending to be more optimistic (i.e., higher) than the means. Nevertheless, the conclusion is the same as that for inflation: significant test outcomes indicate the inconsistencies are related to (anti-)herding behaviour.

7 Conclusions

Our findings present clear evidence that US-SPF participants file their forecasts aware of, and influenced by, the forecasts of others. The degree and nature of the imitative behaviour depends on the forecast horizon. Forecasters tend to herd at the shortest horizons (one quarter ahead) but tend to exaggerate their differences at longer forecast horizons. That herding behaviour may be horizon-specific is a novel finding, and suggests that such behaviour may contribute to the observed pattern of forecaster disagreement varying with the forecast horizon. We find no evidence that professional forecasters working in the financial services provider sector differ from those in the non-financial services sector in terms of the extent to which they herd.

We use the test for herding of Bernhardt *et al.* (2006), which was originally applied to assess the behaviour of professional financial analysts, but has recently been applied more widely, including to macroeconomic forecasting (see, Pierdzioch, Rülke and Stadtmann (2010) and Pierdzioch and Rülke (2012), *inter alia*). This approach often suggests anti-herding: that forecasters exaggerate their differences. For both output growth and inflation we reject the null of no herding in favour of an 'imitation effect' for the shortest horizon forecasts.

Having established that both types of herding behaviour are a feature of our set of annual calendar year forecasts, we then explore whether herding is able to explain a puzzle in the forecasting literature: that some forecasters probability distributions and point predictions of inflation and output growth are inconsistent. We adapt the Bernhardt *et al.* (2006) test to assess whether the probability of a forecaster issuing a point prediction which exceeds (say) the implied mean of a forecast probability distribution depends upon the consensus point predictions. Except at the shortest forecast horizon, there is evidence that discrepancies between point predictions and implied-mean forecasts are exacerbated by anti-herding behaviour, matching the results from the standard application of the test.

Table 1: Tests of whether the survey forecasters herd or exaggerate their differences

Survey Qtr.	Test statistic	Lower bound	Upper bound	CP_1	CP_2	No.
Inflation						
Previous quarter's consensus						
All	0.52	0.50	0.53	0.58	0.45	5084
2	0.55	0.53	0.58	0.58	0.53	1716
3	0.54	0.51	0.56	0.62	0.46	1633
4	0.47	0.45	0.50	0.54	0.40	1735
Current consensus						
All	0.68	0.67	0.70	0.72	0.65	6412
1	0.69	0.67	0.72	0.63	0.76	1617
2	0.70	0.68	0.73	0.75	0.65	1658
3	0.68	0.65	0.70	0.78	0.58	1548
4	0.67	0.64	0.69	0.73	0.60	1589
Output growth						
Previous quarter's consensus						
All	0.50	0.48	0.51	0.49	0.51	5127
2	0.58	0.56	0.61	0.62	0.55	1744
3	0.59	0.56	0.61	0.60	0.57	1639
4	0.33	0.31	0.36	0.27	0.40	1744
Current consensus						
All	0.64	0.63	0.65	0.63	0.64	6847
1	0.67	0.65	0.70	0.70	0.65	1679
2	0.66	0.63	0.68	0.69	0.63	1741
3	0.61	0.59	0.64	0.61	0.62	1630
4	0.61	0.59	0.63	0.54	0.68	1797

The test statistic is the test of of Bernhardt *et al.* (2006) defined in equation 1. The lower and upper bounds are for a 95% confidence interval. CP_1 and CP_2 are defined in section 2, and 'No.' denotes the number of forecasts.

Table 2: Tests by type of forecaster: Inflation

Survey Qtr.	Test statistic	Lower bound	Upper bound	CP_1	CP_2	No.
Financial forecasters						
Previous quarter's consensus						
All	0.60	0.56	0.63	0.67	0.52	1068
2	0.65	0.60	0.71	0.67	0.64	349
3	0.59	0.54	0.64	0.65	0.53	348
4	0.57	0.52	0.63	0.70	0.45	371
Current consensus						
All	0.70	0.68	0.73	0.74	0.67	1377
1	0.74	0.69	0.80	0.66	0.83	343
2	0.74	0.68	0.79	0.79	0.69	346
3	0.68	0.62	0.73	0.74	0.62	345
4	0.65	0.60	0.71	0.77	0.54	343
Non-financial forecasters						
Previous quarter's consensus						
All	0.59	0.56	0.61	0.67	0.50	1401
2	0.62	0.57	0.67	0.65	0.59	471
3	0.60	0.55	0.64	0.66	0.53	453
4	0.57	0.52	0.61	0.70	0.43	477
Current consensus						
All	0.69	0.67	0.72	0.75	0.64	1832
1	0.73	0.68	0.78	0.66	0.80	441
2	0.70	0.66	0.75	0.78	0.63	471
3	0.70	0.65	0.75	0.76	0.63	448
4	0.65	0.61	0.70	0.78	0.53	472

See notes to table 1.

Table 3: Tests by type of forecaster: Output Growth

Survey Qtr.	Test statistic	Lower bound	Upper bound	CP_1	CP_2	No.
Financial forecasters						
Previous quarter's consensus						
All	0.51	0.48	0.54	0.46	0.55	1108
2	0.51	0.46	0.56	0.59	0.43	368
3	0.62	0.57	0.68	0.58	0.66	359
4	0.39	0.34	0.44	0.28	0.50	381
Current consensus						
All	0.61	0.58	0.64	0.58	0.64	1458
1	0.63	0.57	0.68	0.63	0.62	358
2	0.68	0.62	0.73	0.75	0.60	368
3	0.58	0.52	0.63	0.51	0.65	357
4	0.56	0.51	0.61	0.44	0.69	375
Non-financial forecasters						
Previous quarter's consensus						
All	0.51	0.49	0.54	0.44	0.59	1464
2	0.54	0.50	0.59	0.55	0.54	496
3	0.62	0.57	0.66	0.55	0.68	471
4	0.40	0.35	0.44	0.26	0.53	497
Current consensus						
All	0.62	0.60	0.64	0.57	0.67	1932
1	0.66	0.61	0.71	0.69	0.63	468
2	0.64	0.59	0.68	0.64	0.63	496
3	0.62	0.57	0.66	0.52	0.71	471
4	0.57	0.52	0.61	0.42	0.72	497

See notes to table 1.

Table 4: Bounds violations: mean, mode, median and conservative

Survey	#	mean			median			mode			conservative		
		within	below	above	within	below	above	within	below	above	within	below	above
					Output growth								
1	937.0	81.0	7.4	11.6	81.9	9.3	8.9	84.1	9.2	6.7	91.2	4.3	4.5
2	1045.0	85.3	5.7	9.0	84.9	6.3	8.8	87.1	6.7	6.2	93.2	3.3	3.5
3	1000.0	86.0	5.9	8.1	86.8	6.1	7.1	87.4	6.7	5.9	92.9	3.1	4.0
4	1056.0	91.7	3.6	4.7	90.8	4.4	4.8	91.1	4.5	4.5	93.5	2.9	3.6
					Inflation								
1	1468.0	74.1	21.7	4.2	78.3	17.1	4.6	79.8	15.2	5.0	85.8	11.5	2.7
2	1539.0	75.8	19.7	4.5	80.2	15.4	4.4	81.7	13.8	4.5	86.5	10.3	3.1
3	1350.0	79.8	13.3	6.9	81.7	11.1	7.2	81.3	10.7	8.0	87.0	7.6	5.5
4	1103.0	87.4	9.7	2.9	87.9	8.5	3.6	87.8	8.5	3.7	91.9	5.7	2.4

For output growth the percentages reported in the tables are calculated for the surveys from 1981:Q3 to 2013:Q4. For inflation we use the surveys from 1968:Q4 to 2013:Q4.
The Q1 surveys of 1985 and 1986 are excluded as the Philadelphia Fed has documented possible problems with the forecast distributions in these surveys. The point forecasts of the growth rate are calculated using the actual data for the previous year from the RTDSM available in the quarter of the survey. The one exception is that the RTDSM for 1996Q1 is missing the value for 1995Q4. In constructing the year-on-year point forecast growth rates for the respondents to the 1996Q1 survey we use the previous-quarter forecasts (of 1995Q4).

There are missing observations for the histograms for a number of surveys, because respondents were mistakenly asked about the wrong year in those surveys. See the online documentation provided by the Philadelphia Feb: 'Documentation for the Philadelphia Fed's Survey of Professional Forecasters', http://www.phil.frb.org/econ/spf/. The problematic survey quarters are 1985Q1, 1986Q1, 1968Q4, 1969Q4, 1970Q4, 1971Q4, 1972Q3, 1972Q4, 1973Q4, 1975Q4, 1976Q4, 1977Q4, 1978Q4, 1979Q2, 1979Q3, 1979Q4. That these are predominantly Q4 surveys accounts for the smaller number of respondents to Q4 surveys in the table.

Table 5: Tests of whether herding is systematically related to forecast inconsistencies

Survey Qtr.	Test statistic	Lower bound	Upper bound	CP_1^*	CP_2^*	No.
Inflation						
Previous quarter's consensus						
All	0.54	0.52	0.55	0.39	0.69	3991
2	0.57	0.54	0.59	0.37	0.77	1539
3	0.55	0.52	0.58	0.42	0.68	1350
4	0.51	0.48	0.54	0.40	0.63	1102
Current consensus						
All	0.53	0.52	0.54	0.36	0.70	5232
1	0.57	0.54	0.59	0.34	0.79	1432
2	0.55	0.52	0.57	0.36	0.73	1485
3	0.52	0.49	0.55	0.39	0.65	1290
4	0.47	0.44	0.50	0.35	0.59	1025
Output growth						
Previous quarter's consensus						
All	0.54	0.52	0.55	0.57	0.50	3100
2	0.56	0.53	0.59	0.61	0.51	1045
3	0.54	0.50	0.57	0.57	0.50	1000
4	0.51	0.48	0.54	0.55	0.47	1055
Current consensus						
All	0.53	0.52	0.55	0.57	0.49	4017.00
1	0.55	0.52	0.59	0.60	0.51	935.00
2	0.54	0.51	0.57	0.59	0.49	1042.00
3	0.53	0.49	0.56	0.55	0.50	998.00
4	0.52	0.49	0.55	0.56	0.48	1042.00

See notes to table 1, except that the realized value is replaced by the forecast mean when the conditional probabilities and test statistic are calculated.

References

Andersen, T. G., and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accuate forecasts. *International Economic Review*, **39**, 885–905.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association*, **96**, 42–55.

Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modelling and forecasting realized volatility. *Econometrica*, **71**, 579–625.

Ashiya, M., and Doi, T. (2001). Herd behavior of Japanese economists. *Journal of Economic Behavior & Organization*, *46*(3), 343–346.

Balakrishnan, N., and Nevzorov, V. B. (2003). *A Primer on Statistical Distributions*: John Wiley & Sons, New Jersey.

Bernhardt, D., Campello, M., and Kutsoati, E. (2006). Who herds?. *Journal of Financial Economics*, *80*(3), 657–675.

Carroll, C. D. (2003). Macroeconomic expectations of households and professional forecasters. *Quarterly Journal of Economics*, **118**, 269–298.

Clements, M. P. (2009). Internal consistency of survey respondents' forecasts: Evidence based on the Survey of Professional Forecasters. In Castle, J. L., and Shephard, N. (eds.), *The Methodology and Practice of Econometrics. A Festschrift in Honour of David F. Hendry. Chapter 8*, pp. 206–226. Oxford: Oxford University Press.

Clements, M. P. (2010). Explanations of the Inconsistencies in Survey Respondents Forecasts. *European Economic Review*, **54(4)**, 536–549.

Clements, M. P. (2013). US inflation expectations and heterogeneous loss functions, 1968–2010. *Journal of Forecasting*, **33(1)**, 1–14.

Coibion, O., and Gorodnichenko, Y. (2012). What can survey forecasts tell us about information rigidities?. *Journal of Political Economy*, *120*(1), 116 – 159.

Croushore, D. (1993). Introducing: The Survey of Professional Forecasters. *Federal Reserve Bank of Philadelphia Business Review*, **"November"**, 3–13.

Croushore, D., and Stark, T. (2001). A real-time data set for macroeconomists. *Journal of Econometrics*, **105(1)**, 111–130.

Ehrbeck, T., and Waldmann, R. (1996). Why are professional forecasters biased? agency versus behavioral explanations. *The Quarterly Journal of Economics*, **111(1)**, 21–40.

Elliott, G., Komunjer, I., and Timmermann, A. (2005). Estimation and testing of forecast rationality under flexible loss. *Review of Economic Studies*, **72**, 1107–1125.

Elliott, G., Komunjer, I., and Timmermann, A. (2008). Biases in macroeconomic forecasts: Irrationality or asymmetric loss. *Journal of the European Economic Association*, **6**, 122–157.

Engelberg, J., Manski, C. F., and Williams, J. (2009). Comparing the point predictions and subjective probability distributions of professional forecasters. *Journal of Business and Economic Statistics*, **27(1)**, 30–41.

Granger, C. W. J. (1969). Prediction with a generalized cost of error function. *Operations Research Quarterly*, **20**, 199–207.

Lahiri, K., and Liu, F. (2009). On the use of density forecasts to identify asymmetry in forecasters' loss function. *Business avd Economic Statistics Section - JSM*, 2396–2408.

Lahiri, K., and Sheng, X. (2008). Evolution of forecast disagreement in a Bayesian learning model. *Journal of Econometrics*, **144(2)**, 325–340.

Lamont, O. A. (2002). Macroeconomic forecasts and microeconomic forecasters. *Journal of Economic Behavior & Organization*, *48*(3), 265–280.

Landefeld, J. S., Seskin, E. P., and Fraumeni, B. M. (2008). Taking the pulse of the economy. *Journal of Economic Perspectives*, **22**, 193–216.

Laster, D., Bennett, P., and Geoum, I. S. (1999). Rational bias in macroeconomic forecasts. *The Quarterly Journal of Economics*, **114(1)**, 293–318.

Ottaviani, M., and Sorensen, P. N. (2006). The strategy of professional forecasting. *Journal of Financial Economics*, **81**, 441–466.

Patton, A. J., and Timmermann, A. (2007). Testing forecast optimality under unknown loss. *Journal of the American Statistical Association*, **102**, 1172–1184.

Patton, A. J., and Timmermann, A. (2010). Why do forecasters disagree? lessons from the term structure of cross-sectional dispersion. *Journal of Monetary Economics*, **57(7)**, 803–820.

Pierdzioch, C., and Rülke, J.-C. (2012). Forecasting stock prices: Do forecasters herd?. *Economics Letters*, *116*(3), 326–329.

Pierdzioch, C., Rülke, J. C., and Stadtmann, G. (2010). New evidence of anti-herding of oil-price forecasters. *Energy Economics*, *32*(6), 1456–1459.

Zarnowitz, V. (1969). The new ASA-NBER Survey of Forecasts by Economic Statisticians. *The American Statistician*, **23, No. 1**, 12–16.

Zellner, A. (1986). Biased predictors, rationality and the evaluation of forecasts. *Economics*

Letters, **21**, 45–48.

8 Technical appendix: Fitting generalized beta distributions to the histograms

The SPF respondents assign probabilities to inflation / output growth falling in different intervals. The outer intervals are open: e.g., that inflation will be less that $x\%$. For the purpose of fitting the generalized beta distribution the outer intervals are closed by assuming they are equal to the width of the inner intervals.

The generalized beta CDF is given by:

$$\text{Beta}\,(t;a,b,l,r)=\left\{\begin{array}{cc} 0 & \text{if } t\leq l \\ \frac{1}{B(a,b)}\displaystyle\int_l^t \frac{(x-l)^{a-1}(r-x)^{b-1}}{(r-l)^{a+b-1}}dx & \text{if } l<t\leq r \\ 1 & \text{if } t>r \end{array}\right\}$$

where a and b determine the shape, and l and r the support, of the distribution, and where $B\,(a,b)=(\Gamma\,(a)\,\Gamma\,(b))\,/\Gamma\,(a+b)$, and $\Gamma\,(a)=\displaystyle\int_0^\infty x^{a-1}e^{-x}dx$ (see, e.g., Balakrishnan and Nevzorov (2003) for technical details). We can impose unimodality by restricting a, b such that $a>1$ and $b>1$.

The calculation of l and r depends on the distribution of probability across the histogram intervals. Suppose probability is only attached to interior intervals, and to illustrate, let $t_1,\ldots,t_{10}$ denote the right endpoints of the histogram intervals, so that $F\,(t_1)\,,\ldots,F\,(t_{10})$ are points on the individual's CDF. We then set l and r equal to the left and right endpoints of the intervals with positive probability. Suppose we have (say) $F(t_5)=0$, $F(t_6)=0.2$, $F(t_7)=0.7$, and $F(t_8)=1$. Then $l=t_5$, $r=t_8$. Then we minimize only over a and b:

$$\min_{a>1,b>1}\sum_{i=1}^{10}\left[\text{Beta}\,(t_i;a,b,t_5,t_8)-F\,(t_i)\right]^2.$$

If there is mass in either outer interval, then we need to make an assumption about l and r. Following Engelberg *et al.* (2009), the support parameters l and r are bounded by 'the most extreme values that have actually occurred in the United States since 1930': we use $(-0.12,\,0.20)$ and $(-0.13,\,0.19)$ for inflation and output growth respectively. If there is mass in the lower tail interval, then we allow the support to extend below the left endpoint of the lower interval, and l is a free parameter (similarly r if probability is assigned to the upper tail interval. For example, if $F(t_1)=0.2$, $F(t_2)=0.5$, $F(t_3)=0.7$, and $F(t_4)=1$, so there is a 20% chance that inflation will be less than t_1 (the upper value of the lower open-ended interval), we solve the minimization problem:

$$\min_{a>1,b>1,\widetilde{l_1}>l>l^*} \sum_{i=1}^{10} \left[\text{Beta}\left(t_i; a, b, l, t_4\right) - F\left(t_i\right)\right]^2$$

where $\widetilde{l_1}$ is the lower value of the (artificially-closed) open interval, and l^* is the lowest historical value of the variable. That is, the support includes (at a minimum) the range given by the bottom open interval.[12]

When $X \sim Beta\left(a, b, l, r\right)$, the first moments is given by:

$$EX = l + \frac{\left(r - l\right) a}{a + b}.$$

When probability is assigned to fewer bins, the histogram less clearly reveals the individual's underlying subjective distribution. Formally, we are unable to fit the generalized beta distribution when there are fewer than three bins with non-zero probabilities. For 1 and 2-bin histograms we follow Engelberg *et al.* (2009) and fit triangular distribution, which provide symmetric representations of the underlying distributions.

[12]This is closely related to, but not identical to, Engelberg *et al.* (2009, eqn. 3, p.38.) who estimate:

$$\min_{a>1,b>1,l>l^*} \sum_{i=1}^{10} \left[\text{Beta}\left(t_i; a, b, l, t_4\right) - F\left(t_i\right)\right]^2.$$