

Explanations of the Inconsistencies in Survey Respondents' Forecasts

Michael P. Clements*
Department of Economics,
University of Warwick,
UK. CV4 7AL.
m.p.clements@warwick.ac.uk.

December 31, 2008

Abstract

A comparison of the point forecasts and the probability distributions of inflation and output growth made by individual respondents to the US Survey of Professional Forecasters indicates that the point forecasts sometimes present a more optimistic outlook for the economy. We evaluate a number of possible explanations, ranging from the professional forecasters not being fully rational, to delayed updating in the presence of costs to information acquisition and processing, to fully rational forecasters with asymmetric loss functions. The evidence is generally not supportive of fully rational forecasters.

Journal of Economic Literature classification: C53, E32, E37

Keywords: Rationality, point forecasts, probability distributions.

*Computations were performed using code written in the Gauss Programming Language, and STATA. Helpful comments were received when the paper was presented at the European Central Bank, Frankfurt; the International Symposium of Forecasting, Nice; and Statistics Norway, Oslo. Special thanks are due to Daniel SgROI, Mark Stewart, and editor Gerard Pfann.

Explanations of the Inconsistencies in Survey Respondents' Forecasts

A comparison of the point forecasts and the probability distributions of inflation and output growth made by individual respondents to the US Survey of Professional Forecasters indicates that the point forecasts sometimes present a more optimistic outlook for the economy. We evaluate a number of possible explanations, ranging from the professional forecasters not being fully rational, to delayed updating in the presence of costs to information acquisition and processing, to fully rational forecasters with asymmetric loss functions. The evidence is generally not supportive of fully rational forecasters.

Journal of Economic Literature classification: C53, E32, E37

Keywords: Rationality, point forecasts, probability distributions.

1 Introduction

There is now a large literature addressing various aspects of the rationality of point forecasts of key macro aggregates, such as output and inflation (see Stekler (2002) for a recent review), and smaller but expanding literatures on the evaluation of other *types* of forecasts, such as probability distributions (e.g., Diebold, Gunther and Tay (1998), Diebold, Hahn and Tay (1999a)), interval and quantile forecasts (e.g., Granger, White and Kamstra (1989), Christoffersen (1998) and Giacomini and Komunjer (2005)), volatility forecasts (e.g., Andersen and Bollerslev (1998), Andersen, Bollerslev, Diebold and Labys (2003)), and probability forecasts (e.g., Clements and Harvey (2006)). The recent focus on more informative types of forecasts, such as probability distributions, has been welcomed by many, and in policy circles density forecasts are now routinely produced.¹

Recently, a number of authors have taken a different tack: rather than assessing whether a particular set of forecasts are unbiased, or make an efficient use of the available information, they have sought to assess forecaster rationality in terms of whether the different types of forecasts simultaneously made by individual forecasters are mutually consistent. The key papers are Engelberg, Manski and Williams (2007), who compare the point forecasts and histograms of the respondents to the US Survey of Professional Forecasters (SPF), and Clements (2009), who in addition assesses the evidence for consistency of the SPF respondents' histograms and forecast probabilities of declines in real output growth. Inconsistencies are found, and are generally in the direction of the point and probability forecasts indicating a rosier outlook than the histograms: the point forecasts of output growth and inflation are higher and lower, respectively, than implied by the histogram forecasts,² and the histogram probabilities of declines in output tend to overstate the directly-reported probabilities that respondents assign to such an event.

The aim of this paper is to explore a number of potential explanations of the tendency of some forecasters to produce forecasts which are more optimistic than measures of central tendency derived from their histograms. At the heart of our approach is the heterogeneity of expectations. But whereas agents have traditionally been observed to disagree about their point forecasts,³ we show that the professional forecasters we consider are also heterogeneous in terms of their ability or propensity to produce mutually consistent sets of forecasts. The heterogeneous agent model literature will be used to suggest a number of interesting avenues to explore.⁴ An advantage of

¹The Bank of England has for a number of years produced inflation 'fan charts' (see, e.g., Clements (2004) and Wallis (2003)) and a number of other central banks have followed suit.

²García and Manzanares (2007) report a similar tendency for the GDP growth and inflation forecasts of the ECB's Survey of Professional Forecasters, for the period 1999Q1 to 2006Q4.

³Zarnowitz and Lambros (1987) was an early contribution, considering the use of disagreement among forecasters as a proxy for (intra-personal) forecast uncertainty. See also Bomberger (1996), Mankiw, Reis and Wolfers (2003), D'Amico and Orphanides (2006) and Lahiri and Sheng (2008).

⁴Hommes (2006) provides an excellent review of heterogeneous agent models in economics and finance.

assessing rationality in terms of whether the different types of forecasts made by the respondent are consistent is that both types of forecasts are conditioned on the same information set: when the survey respondent fills in their histogram forecast and their point forecast, presumably they do so at the same time and with access to the same information. A problem with the traditional approach to assessing rationality is that it is not always clear what the forecast is conditioned upon.⁵

The explanations we consider can be grouped into four sets. The first is the possibility that the apparent inconsistencies may result from the difficulties inherent in deriving measures of central tendency from the reported histograms, as well as from lack of knowledge of what it is that the point forecasts are. That is, the two sets of forecasts do not match because they are not forecasting the same object. The earlier work on this topic and additional calculations that we report suggest that this is not the explanation. Secondly, we consider the possibility that the complexity of the task leads to inconsistent forecasts. Thirdly, the respondents do not update their forecasts in a timely fashion, leading to inconsistencies creeping in. Finally, we allow the possibility that the two sets of forecasters are consistent once we allow the forecasters to have more general loss functions. We assess the support for these hypotheses from the forecast data.

To date the inconsistencies have been documented in the cited papers, but remain largely unexplained. Engelberg *et al.* (2007) consider whether the tendency of survey respondents to round the probabilities of their histogram forecasts might be responsible, but find that their results are qualitatively unchanged if they allow for rounding. Clements (2009) considers the possibility that respondents' point forecast loss functions are asymmetric (whilst the reported histograms accurately reflect the individuals' true beliefs). Then, the forecaster may rationally report as a point forecast a quantile of their probability distribution other than the central tendency (see, e.g., Elliott, Komunjer and Timmermann (2008) and Patton and Timmermann (2007)). As individuals may have loss functions with different degrees of asymmetry, Clements (2009) tests for rationality allowing for asymmetric loss separately for each individual, based on whether the differences between the histogram means and point forecasts vary systematically with variables in the information set (other than the conditional variance or standard deviation). We will revisit this issue using panel regressions, as we find that forecast uncertainty appears to play an important role, as would be predicted by the asymmetric loss explanation: forecast uncertainty will drive a wedge between the optimal point forecast and the conditional mean. However, there are other explanations that suggest a role for uncertainty, and the evidence is not fully supportive of the asymmetry hypothesis.

To pre-empt our empirical findings, we find evidence in favour of the computational complexity

⁵As Keane and Runkle (1990, p.715) remark, 'Tests of forecast rationality depend upon correct assumptions about what the forecasters tried to predict and what they know when they made their predictions'. The 'tried to predict' is a criticism of the practice of evaluating forecasts against actual values that incorporate benchmark revisions, which were unforeseeable at the time the forecast was made. Most of our analysis does not make use of actual values, but when we do we use an early vintage release.

and delayed updating explanations. Thus our findings are broadly inkeeping with the literature on bounded rationality and learning (see, e.g., Evans and Honkapohja (2001), Mankiw and Reis (2002), and Carroll (2003)): not all the forecasters continually update their forecasts as new information arrives, and the inconsistencies that arise are likely due to the use of simple forecasting rules rather than the use of a fully-integrated forecasting system that produces forecast distributions and point forecasts in a mutually consistent fashion. There is evidence that the point forecasts are updated more rapidly than the histograms, and this is borne out by our study of the relative forecast accuracy of the two types of forecasts. We find that the point forecasts are more accurate on a squared-error loss criterion, such as (root) mean squared forecast error.

A set of possible issues that we do not address is that forecasters may face economic incentives to act strategically in the sense of balancing accuracy against conflicting aims, such as convincing the market that they are well-informed, or of attracting media attention (see, e.g., Ehrbeck and Waldmann (1996), Laster, Bennett and Geoum (1999) and Ottaviani and Sorensen (2006)). It is possible that these factors may impinge differently on the point forecasts and histograms. Whilst the anonymity of the SPF respondents might be expected to rule out some of these strategic motives, one might also argue that the respondents are likely to report the same forecasts to the SPF as they make public, so that these issues remain pertinent.

The plan of the paper is as follows. Section 2 briefly describes the SPF survey, and section 3 records the evidence for inconsistencies between the point forecasts and measures of central tendency of the histograms for inflation and output growth. Section 4 considers the possibility that respondents are ‘boundedly rational’, in the sense that they are more likely to produce inconsistent forecasts if they are more uncertain about the future, and the greater the complexity of the calculations they face. Section 5 considers the possibility that forecasters react differently to the arrival of new information when they update point and histogram forecasts, and that the efficiency of the two types of forecasts may differ. Section 6 presents panel regressions that investigate possible determinants of the difference between the point forecast and the histogram, motivated by asymmetric loss, although the results are amenable to alternative interpretations, as indicated by the discussion in section 4. Section 7 compares the point forecasts and histogram means in terms of forecast accuracy, and indicates that rather than speaking of a tendency of the point forecasts to optimism we ought instead to label the histograms as tending to be too pessimistic, at least in terms of first moments. Section 8 concludes.

2 The Survey of Professional Forecasters (SPF)

The Survey of Professional Forecasters (SPF) is a quarterly survey of professional macroeconomic forecasters that elicits information on the outlook for the US economy. The respondents provide

point forecasts for a number of macro variables, as well as reporting histograms for output growth and inflation. The SPF began in 1968 as the ASA-NBER Survey of Forecasts by Economic Statisticians, administered by the American Statistical Association (ASA) and the National Bureau of Economic Research (NBER). Since June 1990 it has been run by the Philadelphia Fed, renamed as the SPF: see Zarnowitz (1969), Zarnowitz and Braun (1993) and Croushore (1993). This is a unique resource in that it provides histogram and point forecasts at an individual level from the fourth quarter of 1968 to the present day. Moreover, because it is a survey of professional forecasters, authors such as Keane and Runkle (1990) have argued that one can reasonably assume that the reported forecasts reflect the forecasters' expectations, which might not be true when ordinary individuals and firms are surveyed.

We use data from 1968:4 to 2006:4 for inflation, and from 1981:3 onwards for output growth, as prior to 1981:3 the histograms for output growth referred to nominal output, and point forecasts for real GDP (GNP) were not recorded. For these surveys, we have individual respondents' point forecasts for the levels of output and the GNP/GDP deflator in the current year, as well as for the previous quarter, for the current quarter, and the next four quarters. We construct forecasts of the annual growth rates as follows. The current year annual value is constructed by summing the forecasts of the quarters and actual values as appropriate, and then dividing by the previous year's value. The actual values are taken from the Real Time Data Set for Macroeconomists (RTDSM) maintained by the Federal Reserve Bank of Philadelphia (see Croushore and Stark (2001)). The RTDSMs contain the values of output that would have been available at the time the forecast was made, as subsequent revisions, base-year and other definitional changes that occurred after the reference date are omitted. Thus, for a forecast made in 2001:Q1, the 2000 value of output would be taken from the 2001:Q1 RTDSM, and the forecast for 2001 would be calculated by summing the forecast of the current quarter (2001:Q1) and the forecasts of the next three quarters (2001:Q2 to 2001:Q4). For a forecast made in 2001:Q4, the value in 2000 is taken from the 2001:Q4 RTDSM, and the current year forecast consists of the current quarter forecast of 2001:Q4 and the actual values of the first three quarters of the year taken from the same RTDSM. As of the 1981:3 survey, respondents annual forecasts of the current year were also recorded. For both variables from the 1981:3 survey onwards the actual annual level of the variable in the previous year is provided to the survey respondents.

3 Consistency of the point forecasts and histograms

Establishing whether the point forecasts and probability distributions are consistent is complicated by it being unclear whether the point forecasts should be interpreted as the means, modes or even medians of the probability distributions. Engelberg *et al.* (2007) calculate non-parametric bounds

for the three measures of central tendency from the histograms, and obtain similar results, in the sense that roughly similar proportions of the pairs of point forecasts and histograms are found to be inconsistent whether we assume that the point forecasts are the means, the medians, or the modes of the underlying subjective distributions. Our analysis of consistency using bounds replicates that of Engelberg *et al.* (2007), but extends the sample. Engelberg *et al.* (2007) restrict their start period to 1992:1, in part because it is not clear whether prior to this period respondents were always provided with the previous year's level of output / deflator. There is therefore uncertainty over the previous year's value of output that the forecaster had in mind when reporting their current year forecast level. We use all the available surveys and construct the previous years' levels of output and the deflator using the RTDSMs.

The bounds approach allows us calculate the lower and upper values (the 'bounds') of each of the three measures of central tendency that are consistent with a histogram when we do not wish to make any assumption about the distribution of probability mass within the histogram bins. We obtain an upper and lower bound on the measure of central tendency, rather than a point estimate, and the point forecast can then be compared to the bounds to see whether it is consistent with the interpretation that it is the mean, median or mode (or any number of these three) of the underlying distribution reported in the form of the histogram. For the mean, the lower (upper) bound is calculated by assuming that all the probability lies at the lower (upper) limit of the histogram bin. For the median, the bounds are given by the lower and upper values of the histogram bin which contains 50% of the cumulative probability, and the mode is given by the bin with the maximum probability.⁶

Point estimates of moments can be calculated from histograms by assuming that the probability mass is uniform within a bin (e.g., Diebold, Tay and Wallis (1999b) make this assumption in the context of calculating probability integral transforms), as well as by fitting parametric distributions such as the normal (e.g., Giordani and Söderlind (2003, p. 1044)) or the unimodal generalized beta distribution (Engelberg *et al.* (2007)).

Table 1 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, although they are presented separately for each quarter to control for differences that might arise because of the length of the forecast horizon. The results are broadly similar, in that similar proportions of point forecasts are found to be favourable for each measure of central tendency. We conclude that respondents' point forecasts are more optimistic than their histograms to similar degrees for both these variables. The degree of inconsistency diminishes as the horizon shortens

⁶See Engelberg *et al.* (2007) for details. We simplify by considering only histograms with two or more contiguous bins with non-zero probabilities. When two adjacent bins share the maximum probability, the bounds on the mode double to take in both intervals.

(going from the Q1 to Q4 surveys) but the tendency toward optimism persists.

A problem with this is that there is no reason to assume that all the respondents report the same measure, or even that an individual reports the same measure through time. As we do not know what it is that is being reported, we report evidence of inconsistency which is loaded in favour of finding that the point forecasts and histograms are consistent. For each pair of point forecast and histogram, we require only that the point forecast lies within the bounds on any one of the three measures of central tendency in order for the two types of forecast to be ruled mutually consistent. A forecast is ruled favourable only if it exceeds the upper bounds on each of the three measures (in the case of output growth), and is unfavourable if it lies below the lowest of the three lower bounds. Even in this most conservative case we find that approximately 15% of the inflation forecasts made in the first quarter surveys are inconsistent, and the favourable inflation point forecasts outnumber the unfavourable approximately two to one. A similar tendency to optimism is also still apparent for output growth in this conservative scenario.

In the following, we will define inconsistent forecasts as those for which the point forecast is outside the bounds on the mean. Figures 1 and 2 present histograms of the distribution of respondents by the percentage of their forecasts which are above or below the bounds. For both inflation and output it is apparent that the vast majority of respondents fall in the first few cells, indicating that the inconsistencies we observe are not attributable to a small number of errant respondents who habitually report inconsistent forecasts. In section 7 we will discuss the distribution of inconsistencies over time, but it is worth remarking that the inconsistencies are not primarily due to a small number of time periods.

4 Shape of underlying probability distributions and computational complexity

Recall that the bounds approach is required because of the ‘grouping’ or ‘discretization’ of the histogram. Interpreting the point forecast as the mean of the respondent’s underlying subjective distribution, a ‘rational-expectations’ forecaster would produce a point forecast consistent with this distribution, and therefore consistent with the histogram bound, irrespective of the shape of the underlying distribution. However, one might suppose that a boundedly-rational forecaster would produce mutually consistent forecasts when a relatively large overall probability is assigned to the outcome falling in a narrowly defined range but would be less likely to do so when the outcome is perceived as being more uncertain.⁷

⁷So, for example, one would expect the two to be more likely to be consistent in the extreme case of all the probability mass being assigned to a single interval, compared to when positive probability is attached to all the bins.

A testable implication of this conjecture would be that inconsistent forecasts should be associated with more dispersed probability distributions. We calculate measures of dispersion for all the reported histograms, and report averages of those measures conditional on whether the histograms correspond to the point forecast lying within the bounds on the histogram mean, below the bounds, or above the bounds. The measures of dispersion are the variance and the inter-quartile range (IQ).⁸ We take as the averages of these measures the median, to lessen dependence on a small number of extreme values.⁹ Table 2 bears out our hypothesis: the median variances and IQs for the optimistic output and inflation forecasts (labelled ‘above’ and ‘below’, respectively) are markedly higher than the within. However, at least for inflation the pessimistic forecasts are similar to the consistent forecasts in terms of the degree of uncertainty, whilst for output the pessimistic forecasts are generally characterised by more uncertainty. So for output growth greater uncertainty is associated with more optimistic point forecasts, but for inflation greater uncertainty is associated with bound violations in either direction.

An alternative yet closely related explanation is illustrated in table 3. Table 3 reports the number of violations of the bounds according to the number of bins to which the respondent attached positive probabilities. As the maximum number of permitted bins has changed over the period, we report results separately for two sub-periods for output growth, and for three sub-periods for inflation (see notes to table). Except for the first sub-period for output, it is apparent that the proportion of optimistic point forecasts increases with the number of bins used by the respondent. Note that the majority tend to use only a relatively small number of bins. For example, aggregating over the sub-periods, for both variables around 60% of the total number of histograms have positive probabilities attached to only 2 or 3 bins. The table also shows that forecasts using low numbers of bins tend to be those made in the third and fourth quarters of the year, as expected, because there is less uncertainty surrounding the shorter horizon forecasts.

Because the number of non-zero bins (henceforth, b) is positively correlated with forecast uncertainty, the correlation between b and the number of bounds violations may simply reflect the dependence of bounds violations on the shape of the underlying probability distribution. Alternatively, the correlation between violations and uncertainty evident in table 2 may itself be attributable to the greater computational demands of producing mutually consistent histograms and point forecasts for larger b .

⁸The variance is calculated assuming the probability mass lies at the mid-point of each bin, and applying Sheppard’s correction, and the IQ is calculated from linear interpolation (assuming the probability mass is uniform within a bin).

⁹The results reported in the table are based on histograms for which there are two or more bins with non-zero probability, and for which the non-zero bins are contiguous, but neither of these assumptions is crucial. The results are qualitatively unchanged if, for example, we require at least three non-zero bins and allow multi-modal distributions. Requiring at least three non-zero bins has the greatest effect on the shorter horizon forecasts - about a half of the Q4 survey histograms make use of only two bins, reflecting less uncertainty one-step ahead.

Can we disentangle the ‘computational costs’ and ‘distribution shape’ explanations of bounds violations? To check whether there is a correlation between inconsistency and uncertainty holding computational complexity (b) fixed, for b equal to one of $\{2, 3, 4\}$ we calculate the median variance and IQ of the histograms separately depending on whether the point forecast is within the bounds, below the bounds, or above the bounds. Table 4 records the results for small values of b , namely $b = 2, 3, 4$.¹⁰ Apart from the first subsample for inflation, for the other periods for both variables the median histogram variance is higher for the more optimistic forecasts (‘below’ for inflation, ‘above’ for output growth) for $b = 3$ and $b = 4$. So controlling for computational complexity we still find a positive correlation between bounds violations and forecast uncertainty.

An alternative approach is to try and exploit the changes in the degree of discretization of the histogram that occurred over the period for both output and inflation, as documented in the notes to table 3. Essentially, the bin widths were 2 percentage points between 1981:3 and 1991:4, but only 1 percentage point subsequently (1992:1 to 2006:4), and in the earlier period in the case of the inflation forecasts (1968:4 to 1981:2). If the degree of computational complexity is assumed to depend solely on the number of bins assigned non-zero probabilities, and given on average that uncertainty will be higher for a given b for a coarser discretization, we can infer the following directly from table 3. For output, comparing the two periods for either $b = 2$ or 3 , it is evident that the optimistic proportion is similar, and for $b = 4$ or 5 optimism is greater in the second period. Qualitatively the same results hold for inflation. This would suggest that the shape of the underlying probability distributions is not the explanation: computational complexity is key. For a given degree of complexity (given by b) the tendency towards optimism does not increase with the finer discretization (going from 1981:3 - 1991:4 to 1992:1 - 2006:4). The main problem with this is the assumption that the degree of complexity does not depend on the degree of discretization for a given b .

The evidence presented in this section does not clearly discriminate between the two explanations. This is unsurprising given the difficulty of unravelling the two factors. Both explanations share the drawback of not being able to explain the tendency of the point forecasts to relative optimism (as opposed to the two sets of forecasts simply being inconsistent).

5 Delayed updating

A possible explanation for the inconsistency of the point forecasts and histograms is that forecasters react differently to the arrival of new information when they come to update point and histogram forecasts. This might occur if the costs of information acquisition and processing were to differ for

¹⁰As is evident from table 3, for values of b in excess of 4 the number of respondents tails off.

the two types of forecasts.¹¹ One might suppose that a histogram is more costly to produce, and that agents may not revise their histograms as frequently as their point forecasts. Point forecasts that are optimistic relative to the histogram means would then arise if over the period under study the shocks to the economy were predominantly favourable, such that successive point forecasts (of output growth, say) were revised up over time compared to the histogram mean forecasts. In this section we investigate the possibility that point forecasts and histogram means are revised differently as new information arises.¹² An implication of the less rapid updating of the histograms relative to the point forecasts is that the point forecasts should be more accurate. We explore forecast accuracy in section 7.

As well as assessing how the two sets of forecasts are revised in response to new information, we also consider whether an efficient use is made of past information. If so, then revisions to ‘fixed-event’ forecasts should not be systematically related to any information available at the time the initial forecast was made (see, e.g., Nordhaus (1987), Clements (1995)). Revisions to fixed-event forecasts are the changes between successive forecasts of a given target (say, the year-on-year rate of inflation in year t). We allow for information sets comprising macro-economic variables known at the time the initial forecast was made.

We analyse these two issues using panels of forecasters. The panels are unbalanced because few respondents are ever-present given the duration of the surveys (1968 onwards for inflation; 1981 onwards for real output). We include all individuals who reported full sets of point forecasts and histograms for five or more years. The following two regressions were estimated for both point forecasts and histogram means:

$$x_{it,q} = \alpha_{1,q}x_{it,q-1} + \alpha_{2,q}C_{t,q} + u_{it,q} \quad (1)$$

and:

$$x_{it,q} - x_{it,q-1} = \beta_q Z_{t,q-1} + u'_{it,q} \quad (2)$$

where $t \in [1982, 1983, \dots, 2006]$ for output growth and $t \in [1969, 1983, \dots, 2006]$ for inflation, q indexes the survey quarter, and i the respondent. Hence when $q = 2$, $x_{it,q}$ denotes a forecast of the annual rate of output growth or inflation made in the second quarter of that year, and $Z_{t,q-1}$ in (2) contains macro-variables known at the time of the first-quarter survey of year t (which will

¹¹That latest information may not be incorporated in expectations is an implication of the sticky-information model: Mankiw and Reis (2002), Mankiw *et al.* (2003) and Carroll (2003). However, one might argue that such a model is unlikely to apply to professional forecasters - Carroll (2003) takes the view that professional forecasts embody the latest news.

¹²This may also arise within a Bayesian learning setting, whereby individual i updates their prior forecast (say, their first forecast of y_t , made at $t - p$; $y_{i,t|t-p}$) as new information accrues to produce a posterior forecast (say, $y_{i,t|t-p+1}$). Kandel and Pearson (1995), Kandel and Zilberfarb (1999) and Lahiri and Sheng (2008) consider the possibility that individuals interpret the new information differently.

relate to first estimates of their values in the fourth quarter of the previous year).

We estimate (1) and (2) separately for point forecasts and histogram mean forecasts, and for each $q = \{2, 3, 4\}$. This allows for the possibility that the rate at which ‘news’ is incorporated may depend on the forecast horizon, in the case of (1), and that the efficiency of the forecasts may depend on the horizon, in the case of (2). Survey data on exchange rate expectations suggests agents use different forecasting strategies for different horizons: specifically, there is a tendency to adopt technical trading rules or ‘chartist’ forecasting rules for short-horizons, but to rely on mean-reverting ‘fundamentalist’ trading rules for long-horizon forecasts (see, e.g., Frankel and Froot (1987, 1990), and the survey article by Hommes (2006)). It seems plausible to allow that professional forecasters of macro-variables, such as real output growth and inflation, may also adopt simple forecasting strategies that may also be horizon-dependent.¹³

Consider regression (1). $c_{t,q}$ is the ‘consensus’ or average of the point forecasts of all those who responded to the quarter q survey in year t . The consensus forecast is assumed to incorporate all the latest available information or ‘news’. We are assuming the consensus forecast is known when the individual makes their forecast. Literally, this requires that an individual knows the forecast being made by all other respondents. This may not be too unreasonable - as professional forecasters, the SPF respondents are likely to contribute point forecasts to analysts reports and other news media on a regular basis, and so will be in the public domain. (Note we use the consensus of the point forecasts in the regression for the individual histogram means, rather than the consensus mean forecast, for precisely this reason). The consensus can also be viewed as a proxy for a forecast informed by the latest information on monthly indicators, federal announcements, etc. Of interest in (1) is the value of α_2 , and whether this differs between the point forecasts and histogram means, as well as α_1 , which captures the tendency to anchor on a prior position (here, last quarter’s forecast).

In (2), $\beta = 0$ implies that revisions are not correlated with information in the agents’ information sets. The macro-variables we include are the unemployment rate and the 3-month Treasury bill rate, as well as output growth and the rate of inflation (GNP/GDP deflator), all taken from the RTDSMs. We construct series consisting of the latest available value at each point in time (so, for example, $Z_{t,q-1}$ is the value of the macro-variable in the RTDSM available in quarter $q - 1$ of year t : this will be the first-release value for quarter $q - 2$).¹⁴

¹³Especially as simple forecasting rules, such as ‘next year’s inflation rate will be the same as the current rate’ may prove hard to beat’: see Atkeson and Ohanian (2001), for example. See also Hendry (2002) for a theoretical justification for using simple forecasting ‘rules of thumb’.

¹⁴Output growth and inflation are constructed as quarterly changes. The Treasury bill and unemployment rate are entered as levels.

Note that the 1996:1 vintage value of output and the deflator for 1995:4 is missing. In the test regressions this is replaced by the 1996:1 vintage value of 1995:3 (relative to 1995:2).

Tables 5 and 6 report the results of estimating (1) and (2). Both are estimated using the fixed effects estimator (FE) and the random effects GLS estimator (RE-GLS).¹⁵ Note that for (2) there are no individual-specific explanatory variables, and the use of pooled OLS (with an appropriate estimator of the coefficient covariance matrix) gave very similar results to those reported in the table.¹⁶ In each case we include all four macro-variables and record the p -value of a test of joint insignificance.

The estimates in table 5 indicate that the point forecasts for both inflation and output growth exhibit a similar pattern: the degree of anchoring (α_1) diminishes as the forecast horizon shortens, as greater weight is placed on the consensus view. However there are marked differences between the point forecasts and the mean forecasts. The histogram mean forecasts for output growth follow a broadly similar pattern, in that α_1 decreases and α_2 increases as the horizon shortens, but the magnitudes of the effects are such that the forecasts are more anchored on the previous forecast, and are less responsive to new information, than the equivalent point forecasts. For example, α_1 is over twice as large for the Q2 survey mean forecasts of output growth as for the point forecasts, while the value of α_2 is around 0.60 compared to 0.85 for the point forecasts. In the case of inflation, the estimated value of α_1 is around one third for the shortest horizon (Q4 survey forecasts), compared to 0.05 for the point forecasts, and the sensitivity to new information is markedly lower.

Next, the results in table 6 indicate that generally the revision to the point forecasts and histogram means are systematically related to macro-variables known at the time the initial forecast was made. There are exceptions: the long-horizon point forecasts of output growth appear to be efficient, as do the Q3 histogram mean output growth forecasts. The results reported in table 6 on the efficiency of the revisions to the forecasts do not suggest clear differences by type of forecast that might explain the inconsistencies between the two types of forecasts. However they do cast doubt on rationality in terms of the efficient use of public information: revisions are predictable.

6 Asymmetric loss functions

In this section we consider the possibility that the inconsistencies stem from respondents reporting point forecasts that are optimal for asymmetric loss functions. That is, respondents are rational forecasters once we allow for the appropriate loss function. Our motivation is the recent literature on asymmetric loss (e.g., Elliott *et al.* (2008) and Patton and Timmermann (2007)) and in particular

¹⁵Both models are estimated in STATA using ‘panel robust’ estimation of the covariance matrix of the estimated parameters. If we write the composite error in (1) (and similarly for (2)) as $u_{it,q} = \delta_{i,q} + v_{it,q}$, then RE-GLS requires that the $\{v_{it,q}\}$ are serially independent. Because we run separate regressions for each forecast horizon (i.e., for the first-quarter surveys, the second quarter, etc) this may be a reasonable assumption.

¹⁶Note that OLS is a consistent estimator of the RE model even if autocorrelation in $\{v_{it,q}\}$ induces correlation in the composite errors $\{u_{it,q}\}$, but is inconsistent for the FE model in these circumstances.

Elliott *et al.* (2008) who find much lower rates of rejection of the rationality of the SPF respondents real output growth and inflation forecasts once an allowance is made for asymmetric loss. We pair the two types of forecasts (histogram mean and point forecast) by individual and report panel regressions of whether the discrepancy between the two is systematically related to a number of explanatory variables, including forecast uncertainty. This is motivated by the recognition that systematic differences between histogram means and point forecasts may be consistent with optimal forecasting under asymmetric loss. However, we have already discussed in section 4 alternative explanations of a possible dependence between the difference between the histogram means and point forecasts and forecast uncertainty, and so the evidence for asymmetric loss will be critically assessed.

For the asymmetric loss explanation we need to assume that the histograms accurately reflect the individuals' true (subjective) beliefs, in the sense that the histograms are not an intentionally 'biased' representation of the individuals' probability assessments. By way of contrast, individuals report point forecasts that are optimal for asymmetric loss functions, in the sense that greater costs are attached to (say) under-predictions compared to over-predictions.¹⁷ Clements (2009) considers whether asymmetric loss can explain the tendency to report relatively optimistic output growth forecasts. He estimates separate regressions for each individual (who made a minimum number of returns), whereas we will pool over individuals, and consider both the output growth and inflation forecasts. We will also contrast the results we obtain using the estimated histogram means with a panel-logit approach that only requires the calculation of non-parametric bounds on the histogram mean.

Using the results in Patton and Timmermann (2007), we can show that, under relatively weak restrictions on the form of the loss function and the data generating process, the optimal forecast is given by:

$$f_{t+h,t} = E_t(y_{t+h}) + \phi_h \cdot \sqrt{V_t(y_{t+h})}$$

where $E_t(y_{t+h}) \equiv E(y_{t+h} | \Omega_t)$, $V_t(y_{t+h}) \equiv Var(y_{t+h} | \Omega_t)$, with the data generating process $y_{t+h} | \Omega_t \sim D(E_t[y_{t+h}], V_t[y_{t+h}])$ for some constant distribution function D , and where ϕ_h is a constant that depends only on the form of D and the forecast loss function. $f_{t+h,t}$ is the directly reported point forecast. $\phi_h < 0$ when over-predictions ($e_{t+h,t} < 0$) are penalised more heavily than under-predictions, and vice versa. The deviation between the optimal point forecast and the conditional mean depends on the conditional standard deviation. Intuitively, the more costly over-predictions relative to under-predictions, say, and the more likely both over- and under-predictions

¹⁷The literature on the properties of optimal forecasts under asymmetric loss includes, *inter alia*, Granger (1969), Zellner (1986), Christoffersen and Diebold (1997), Elliott and Timmermann (2004), Elliott, Komunjer and Timmermann (2005), Patton and Timmermann (2007), and Elliott *et al.* (2008), and is used to motivate this section.

(because the more uncertain the outlook), then the more the forecaster will aim to under-predict on average. Under the assumptions we have made, the bias of a rational forecaster should depend on the forecast standard deviation but should not be systematically related to other variables known at time t :

$$\begin{aligned} E(y_{t+h} - f_{t+h,t} | \Omega_t) &= E\left(y_{t+h} - \left(E_t(y_{t+h}) + \phi_h \cdot \sqrt{V_t(y_{t+h})}\right) | \Omega_t\right) \\ &= -\phi_h \cdot \sqrt{V_t(y_{t+h})} \end{aligned}$$

This motivates the suggestion of Pesaran and Weale (2006) to test for rational expectations with asymmetric losses by running a regression such as:

$$e_{t+h,t} \equiv y_{t+h} - f_{t+h,t} = \tilde{\zeta}_1 \sqrt{V_t(y_{t+h})} + \tilde{\zeta}_2' \mathbf{Z}_t + \tilde{\epsilon}_{t+h} \quad (3)$$

where under the null we would expect to find $\tilde{\zeta}_2 = \mathbf{0}$, but $\tilde{\zeta}_1 \neq 0$ if loss is asymmetric. Alternatively, the same implications hold for a regression in which the dependent variable is not the forecast error but the difference between the histogram mean and the point forecast, say

$$E_t(y_{t+h}) - f_{t+h,t} = \zeta_1 \sqrt{V_t(y_{t+h})} + \zeta_2' \mathbf{Z}_t + \epsilon_{t+h} \quad (4)$$

In general, (4) might be preferred to (3) for the following reason. Suppose y_{t+h} and $E_t(y_{t+h})$ differ by a random term, $\xi_{t+h} = y_{t+h} - E_t(y_{t+h})$. In this interpretation, ξ_{t+h} is a measurement error, such that the dependent variable in (3) is a noisy proxy for the dependent variable $E_t(Y_{t+h}) - f_{t+h,t}$ in (4). Standard analysis suggests that even if $E(\xi_{t+h} | \Omega_t) = 0$, so that ξ_{t+h} is uncorrelated with any variables that might be included as explanatory variables, then inference based on (3) will be less precise than tests based on (4). But when in addition $E(\xi_{t+h}, \mathbf{X}_t) \neq \mathbf{0}$, where $\mathbf{X}_t = \left[\sqrt{V_t(y_{t+h})} \quad \mathbf{Z}_t \right]'$, then inference based on (3) is invalid.¹⁸ Clements (2009) provides further

¹⁸Consider the simplest case where the only explanatory variable is $\sqrt{V_t(y_{t+h})}$, and the population regression is given by:

$$E_t(Y_{t+h}) - f_{t+h,t} = \delta_1 \sqrt{V_t(Y_{t+h})} + \varepsilon_{t+h}$$

where in fact $\delta_1 = 0$, so that loss is quadratic. Inference is based on this regression but with the dependent variable measured with error, $e_{t+h,t} = y_{t+h} - f_{t+h,t} = E_t(Y_{t+h}) - f_{t+h,t} + \xi_{t+h}$, so that the actual regression the investigator runs is:

$$e_{t+h,t} = \zeta_1 \sqrt{V_t(y_{t+h})} + \epsilon_{t+h}$$

where $\epsilon_{t+h} = \varepsilon_{t+h} + \xi_{t+h}$. Then $\hat{\zeta}_1 \xrightarrow{p} \delta_1 + Cov\left(\sqrt{V_t(y_{t+h})}, \varepsilon_{t+h} + \xi_{t+h}\right) \times Var\left(\sqrt{V_t(y_{t+h})}\right)^{-1}$. Therefore, $\hat{\zeta}_1 \xrightarrow{p} Cov\left(\sqrt{V_t(y_{t+h})}, \xi_{t+h}\right) \times Var\left(\sqrt{V_t(y_{t+h})}\right)^{-1}$. If ξ_{t+h} is positively (negatively) correlated with the conditional standard deviation, then we may conclude that δ_1 is positive (negative).

discussion of (4) relative to (3) when we allow for individual heterogeneity¹⁹: we will simply report tests based on (4) as the dependent variable in this regression is the quantity of direct interest in our context.

When we allow heterogeneous information sets and individual-specific ϕ_h values, ϕ_{hi} , then $f_{t+h,t,i} = E_{t,i}(y_{t+h}) + \phi_{hi}\sqrt{V_{t,i}(y_{t+h})}$, where the conditional mean and variance are defined as expectations with respect to the individual's information set $\Omega_{t,i}$. In general, this suggests that ζ_1 in (4) may differ over individuals, and we test the assumption that the coefficients of the panel regression are the same over individuals by allowing both random effects and random coefficients.

The results of the panel regressions are given in table 7. We restrict the explanatory variables to the standard deviation, an intercept, and either the value of output growth or inflation known at the time the forecast was made (taken from the RTDSMs), in order that we can more easily interpret the estimated coefficients. The forecast standard deviation is the square root of the variance calculated assuming the probability mass lies at the mid-point of each bin, and applying Sheppard's correction. Separate regressions are run for forecasts made in the four quarters of the year, as the theory suggests the coefficient on the forecast standard deviation may depend on the forecast horizon. We begin by reporting results for the random-effects GLS estimator, as the Hausman specification tests do not reject the assumptions that support the use of this more efficient estimator relative to the fixed-effects estimator. (The Hausman tests are calculated assuming the random-effects estimator is fully efficient). For inflation, we find that the standard deviation is positive and (last period's) inflation rate is negative for the Q1 survey forecasts, and that these terms are individually and jointly significant. At shorter horizons neither term is statistically significant. Interpreted in the context of asymmetric (point forecast) loss functions, the positive effect of the forecast standard deviation would suggest higher costs to over-predicting relative to under-prediction. For the forecasts of output, actual output is significant for the Q1 forecasts, and remains so for the Q2 forecasts. The sign on the standard deviation is now negative, so in terms of asymmetric loss, has the interpretation that it pays to be optimistic, but is not statistically significant from zero. For both the output and inflation forecasts the negative coefficient on the

¹⁹Essentially the use of (4) allows tests of the asymmetry of the loss function without requiring that the forecasts make full use of all available information. Consider the standard way of testing for asymmetry. This regresses the forecast error (constructed from the outcomes and point forecasts) on variables known at the time the forecast was made, and the conditional standard deviation of the forecasts. Suppose that, relative to an individual's information set, there are a series of negative shocks to output growth over the sample period, so that the individual's point forecasts tend to be too favourable - this could be taken as evidence that the individual has asymmetric loss such that over-predictions are less costly than under-predictions. Using (4), the conditional mean would control for the negative shocks - it would be higher than warranted based on an information set that includes the shocks, but under the asymmetry hypothesis the deviation between the conditional mean and the point forecast should only depend on the conditional standard deviation of the forecast. Under quadratic loss this deviation should not differ systematically from zero and should not be related to any variables known at the time the forecasts are made.

actual value indicates that higher output growth (inflation) is systematically related with next period's point forecast exceeding the histogram mean, at least for the Q1 survey forecasts. In table 6 we rejected the efficiency of both types of forecast in separate regressions. Our results here indicate that the difference between the two varies systematically with lagged output growth (inflation), at least for the Q1 survey forecasts.

The random-coefficients estimator treats the parameter vector as a different realization of a stochastic process for each individual respondent. This may be warranted to the extent that the degree of asymmetry of the loss function is not the same over individuals. The coefficient estimates reported in table 7 are the means of the distributions. In addition, we report a test of parameter constancy, which for both variables and all four horizons rejects the assumption that the coefficients are the same across the individuals. Even so, the picture that emerges is similar to that assuming random effects - we reject the joint null of insignificance of both the explanatory variables (other than the constant) for the Q1 forecasts, and the signs of the average coefficients agree with the random-effects estimates.

Finally, we also wish to check whether the results are qualitatively unchanged if we change the assumptions behind the calculation of the histogram means. We could make a different assumption about the distributions that underlie the reported histograms, or we could simply use the non-parametric bounds approach. We choose to do the latter, and define a binary dependent variable which is unity when the forecasts are more optimistic than the bounds (i.e., exceed the upper bound, fall below the lower bound, for output and inflation, respectively).²⁰ The cost of using the bounds approach is a loss of information because large mean-point forecast discrepancies are treated on a par with those which are only just inconsistent with the bounds. Against this, we can be confident that the observed discrepancies are a characteristic of the point forecasts and underlying subjective distributions and not just our chosen method of calculating moments from histograms.

The final set of results in table 7 are for the random-effects panel-logit estimator. The results confirm the finding that the forecast standard deviation and the actual value in the previous period are significant for the Q1 forecasts for both variables. In addition we find that both variables remain significant for the Q2 output growth forecasts, and that the forecast standard deviation matters for inflation for all quarters. The signs on the standard deviations are consistent with the signs we found earlier given the definitions of the dependent variables.

We also report a specification test of whether the panel (individual) dimension is statistically significant, that is, whether the individual-level variance component matters. For output, we reject the likelihood ratio test that the panel and pooled-OLS estimators are equivalent, but for inflation

²⁰It may seem natural to allow three categories - 'optimistic', 'pessimistic' and 'consistent with the bounds', but have not done this because there are few pessimistic forecasts.

we find the opposite (except for the Q2 forecasts).

In summary, the different models and panel estimators all indicate a role for the forecast standard deviation and last period's actual value in explaining the discrepancy (or the inconsistency, in the case of the panel-logit regressions) between the mean and point forecasts, especially at the longest horizon forecasts, which are those made in the first quarter of the year. In terms of the discussion of asymmetric loss, the relevance of the forecast standard deviation would point towards asymmetry, whilst the significance of the lagged actual value is inconsistent with rational behaviour by the forecaster for the class of loss functions we have allowed. The lagged values of output growth and inflation should not be systematically related to the difference between the respondent's histogram mean and point forecast. The significance of the forecast standard deviation in these regressions is consistent with the explanations of distribution shape and computational complexity of section 4, but it is only in conjunction with the hypothesis of asymmetric loss that this factor explains the direction of the inconsistencies (i.e., the tendency to over-optimism of the point forecasts). In the next section we consider forecast accuracy, which as we will show is not supportive of the asymmetric loss hypothesis.

7 Accuracy of point forecasts and histogram means

Hitherto we have not considered which of the two types of forecast are the more accurate. Although the notions of forecast accuracy and the consistency of the different types of forecasts are conceptually distinct, there are good reasons to consider accuracy. It is natural to ask whether the apparent 'optimism' of the point forecasts is well founded, especially as the relative accuracy of the two types of forecasts bears on the credibility of two of the explanations we have offered for the inconsistencies. The delayed updating explanation of section 5 indicates that the point forecasts should be more accurate, whilst the discussion of asymmetric loss in section 6 predicts the reverse.

One of the problems with measuring forecast accuracy is that it is unclear what should be used as outturns to calculate forecast errors. Respondents may be seeking to forecast the first announcements of the data, or the second, or some subsequent revision. A number of authors have chosen to use the second data release (e.g., Romer and Romer (2000) and Patton and Timmermann (2007)) and we report results for this choice of outturns. One might justify the use of a relatively early vintage if revisions are 'news' in the sense of being unpredictable at the time the initial estimates was released,²¹ as well as when there are 'benchmark revisions', such that the latest estimates are compiled using different accounting procedures or methodological practices.

Table 8 reports the accuracy (measured by mean squared forecast error - MSFE) of the point

²¹Mankiw and Shapiro (1986) found that revisions to real output added news - see Croushore (2006) for a full discussion of data vintage revisions.

forecasts and histogram means for a number of sub-periods, and also compares the relative accuracy of the two sets of forecasts when they are inconsistent (in the sense that the point forecasts lie outside the non-parametric bounds on the histogram means). If we consider all the forecasts for both variables, then the point forecasts are markedly more accurate than the histogram means calculated in the conventional way (assuming all the probability mass is located at the midpoints of the bins) or via a normal approximation (following Giordani and Söderlind (2003)), for all the sub-periods. Although the differences in MSFE are often large (for output growth, for example, the MSFE of the point forecasts is only 53% of that of the histogram means for the period 1981:3 to 1991:4, and only 72% for the period 1992:1 to 2006:4), we are assuming that the methods we use to calculate the histogram means accurately reflect the means of the respondents' probability distributions. To counter the possibility that our finding in favour of the point forecasts in terms of accuracy is an artifact of the way the histogram means are calculated, we next focus on those pairs of forecasts when the means and point forecasts are inconsistent, either because the point forecast exceeds the upper bound, or falls below the lower bound. For these two sets of forecasts, we undertake a comparison where the histogram mean is calculated so as to minimize the squared forecast error, subject to it lying within the bounds.²² The MSFEs of this most-favourable scenario for the histogram means are given in the final column of the table. For inflation, the point forecasts are more accurate than the histogram mean when they are more optimistic (rows headed by 'point < bound') for the second and third subperiods, but not for the first. For output growth, the 'optimistic' ('point > bound') point forecasts are more accurate for the first subperiod and just about the same in the second period.

This would suggest that the point forecasts are more accurate indicators of first moments than the histograms. Engelberg *et al.* (2007) view the discrepancies between the point forecasts and histograms as suggesting 'that point predictions may have a systematic, favourable bias', and infer from this that measures of central tendency should be calculated from 'probabilistic expectations'. The evidence presented here indicates instead that first moments derived from survey respondents histograms have a tendency toward pessimism, and that directly reported point forecasts are more accurate. The greater accuracy of the point forecasts is consistent with the point forecasts incorporating new information more quickly than the histograms, as found in section 5. Figures 3 and 4 portray the distribution of inconsistencies over time, and provide further support of the delayed updating of the histograms explanation. Figure 4 for inflation shows that the majority of the pairs of forecasts for which the point forecast exceeded the histogram mean forecast occur in the earlier period (from table 8, 80% (192 out of 241) of the 'pessimistic' point forecasts occur in

²²Thus the histogram mean equals the actual when the actual lies within the bounds, giving a zero error; equals the upper bound when the actual lies above the upper bound; and equals the lower bound when the actual lies below the lower bound.

the 68:4 to 81:2 period). This historical period was marked by high and variable rates of inflation, and a tendency for the mean histogram forecasts to lag developments would account for the greater forecast accuracy on MSFE of the point forecasts over this period (as shown in table 8). Figure 3 also provides evidence that ‘optimistic’ output growth point forecasts were more concentrated in the nineties, the longest recorded period of expansion of the US economy²³, which is again consistent with delayed updating of the histogram means.

The relative forecast accuracy findings also bear on the credibility of the asymmetric loss explanation of the observed inconsistencies. This is because the expected squared error of a point forecast must be at least as large as the expected squared error of the conditional mean; the conditional mean minimizes the expected squared error amongst all possible forecasts. The expected squared error of the point forecast will equal that of the conditional mean when loss is quadratic, when the optimal point forecast is the conditional mean, but for asymmetric loss the squared-error loss of the optimal point forecast should exceed that of the conditional mean. Our finding that point forecasts are more accurate under quadratic loss (MSFE) counts against the asymmetric loss argument.

8 Conclusions

We have sought explanations for the finding that respondents to the Survey of Professional Forecasters habitually report point forecasts of annual output growth and inflation which are higher and lower, respectively, than the mean forecasts of their forecast histograms for these two variables. This occurs most frequently for the forecasts made in the first quarter - approximately 20% of first-quarter output growth and inflation forecasts lie above and below, respectively, the bounds on the corresponding histogram means. We show that this finding is not driven by an incorrect assumption that respondents’ reported point forecasts correspond to their means - approximately 15% of the point forecasts of inflation made in the first quarter surveys are not compatible with the histograms whether we interpret them as either the mean or the mode or the median of the underlying subjective probability distributions. Moreover, we find that these inconsistencies are not attributable to a small number of errant respondents.

We find that the degree of forecast uncertainty, as measured by the dispersion of the histogram, is systematically related to the tendency for the point forecasts to present a rosier outlook than the histograms. We have considered a number of possible reasons for this correlation. The first supposes that the respondents do not use fully-integrated forecasting systems to produce mutually consistent point forecasts and probability distributions, and are more likely to make mistakes when

²³The economy was in an expansionary phase from the trough in the first quarter of 1991 to the peak in the fourth quarter of 2001, according to the NBER (<http://www.nber.org/cycles/cyclesmain.html>).

they perceive the future to be more uncertain (flatter distributions) or the calculations required of them are more complex. These two explanations are hard to distinguish, as a more dispersed underlying distribution will be correlated with non-zero probabilities being assigned to a greater number of histogram bins, and the use of a greater number of bins may decrease the likelihood of the respondent calculating mutually consistent point forecasts and histogram means. Comparing across histograms before and after a change in the bin widths (allowing greater dispersion for a given level of computational complexity) provides some tentative evidence that computational complexity, rather than the shape of the underlying probability distributions, is the relevant factor, although the evidence is somewhat mixed.

However, the recent literature on testing for rationality allowing for asymmetric loss also suggests that forecast uncertainty will affect the relationship between the conditional mean and the point forecasts. This explanation restores the notion of the rational expectations forecaster. If the histograms accurately reflect the individuals' true (subjective) beliefs, then individuals will report point forecasts that under or over-predict relative to their mean forecast if they have asymmetric loss functions, and the extent to which they do so will depend on their perceived forecast uncertainty. Our results are generally not supportive of the asymmetric loss explanation, for two reasons. Firstly, although we find a role for the forecast standard deviation in our panel regressions, we also find that last period's actual value helps predict the difference between the mean and point forecast. The significance of the lagged actual value is inconsistent with rational behaviour by the forecaster for the class of loss functions we have allowed. Secondly, we find that point forecasts are more accurate (assuming a squared-error loss function) than the histogram means. If the respondents have asymmetric loss functions, then the expected squared error of a point forecast must exceed the expected squared error of the conditional mean. Our relative forecast accuracy findings count against the asymmetric loss explanation. The discrepancies between the point forecasts and histograms have been viewed as evidence 'that point predictions may have a systematic, favourable bias' (Engelberg *et al.* (2007)), but this appears to be unwarranted. Rather, our findings indicate that first moments derived from survey respondents histograms have a tendency toward pessimism relative to the outcomes. The relative accuracy of the point forecasts would also tend to lend support to our maintained assumption that the point forecasts are estimates of the mean.

An alternative explanation for the inconsistency of the point forecasts and histograms that receives some support is that forecasters react differently to the arrival of new information when they come to update point and histogram forecasts, as might be expected if the costs of information acquisition and processing differ for the two types of forecasts. The longer-horizon histogram mean forecasts are not updated to incorporate the latest information at same rate as the point forecasts, and are more anchored to the past view. The greater accuracy of the point forecasts is consistent with the more timely updating of the point forecasts.

References

- Andersen, T. G., and Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, **39**, 885–905.
- Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modelling and forecasting realized volatility. *Econometrica*, **71**, 579–625.
- Atkeson, A., and Ohanian, L. (2001). Are Phillips curves useful for forecasting inflation?. *Federal Reserve Bank of Minneapolis Quarterly Review*, **25**, 2–11. (1).
- Bomberger, W. A. (1996). Disagreement as a measure of uncertainty. *Journal of Money, Credit and Banking*, **28**, 381–392.
- Carroll, C. D. (2003). Macroeconomic expectations of households and professional forecasters. *Quarterly Journal of Economics*, **118**, 269–298.
- Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, **39**, 841–862.
- Christoffersen, P. F., and Diebold, F. X. (1997). Optimal prediction under asymmetric loss. *Econometric Theory*, **13**, 808–817.
- Clements, M. P. (1995). Rationality and the role of judgement in macroeconomic forecasting. *Economic Journal*, **105**, 410–420.
- Clements, M. P. (2004). Evaluating the Bank of England density forecasts of inflation. *Economic Journal*, **114**, 844 – 866.
- 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, Chapter 8*: Oxford University Press.
- Clements, M. P., and Harvey, D. I. (2006). Forecast encompassing tests and probability forecasts. Working paper, Department of Economics, University of Warwick.
- Croushore, D. (1993). Introducing: The Survey of Professional Forecasters. *Federal Reserve Bank of Philadelphia Business Review*, **November/December**, 3–13.
- Croushore, D. (2006). Forecasting with real-time macroeconomic data. In Elliott, G., Granger, C., and Timmermann, A. (eds.), *Handbook of Economic Forecasting, Volume 1. Handbook of Economics 24*, pp. 961–982: Elsevier, North-Holland.
- Croushore, D., and Stark, T. (2001). A real-time data set for macroeconomists. *Journal of Econometrics*, **105**, 111–130.
- D’Amico, S., and Orphanides, A. (2006). Uncertainty and disagreement in economic forecasting. mimeo, Board of Governors of the Federal Reserve System, Washington, D.C.

- Diebold, F. X., Gunther, T. A., and Tay, A. S. (1998). Evaluating density forecasts: With applications to financial risk management. *International Economic Review*, **39**, 863–883.
- Diebold, F. X., Hahn, J. Y., and Tay, A. S. (1999a). Multivariate density forecast evaluation and calibration in financial risk management: High frequency returns on foreign exchange. *Review of Economics and Statistics*, **81**, 661–673.
- Diebold, F. X., Tay, A. S., and Wallis, K. F. (1999b). Evaluating density forecasts of inflation: The Survey of Professional Forecasters. In Engle, R. F., and White, H. (eds.), *Cointegration, Causality and Forecasting: A Festschrift in Honour of Clive Granger*, pp. 76–90. Oxford: Oxford University Press.
- 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.
- Elliott, G., and Timmermann, A. (2004). Optimal forecast combinations under general loss functions and forecast error distributions. *Journal of Econometrics*, **122**, 47–79.
- Engelberg, J., Manski, C. F., and Williams, J. (2007). Comparing the point predictions and subjective probability distributions of professional forecasters. *Journal of Business and Economic Statistics*. Forthcoming.
- Evans, G. W., and Honkapohja, S. (2001). *Learning and Expectations in Macroeconomics*. Princeton: Princeton University Press.
- Frankel, J. A., and Froot, K. (1987). Using survey data to test standard propositions regarding exchange rate expectations. *American Economic Review*, **77**, 133–153.
- Frankel, J. A., and Froot, K. (1990). The rationality of the foreign exchange rate. chartists, fundamentalists, and trading in the foreign exchange market. *American Economic Review*, *80*(2), 181–185.
- García, J. A., and Manzanares, A. (2007). Reporting biases and survey results: evidence from European professional forecasters. ECB Working Paper No. 836, European Central Bank, Frankfurt.
- Giacomini, R., and Komunjer, I. (2005). Evaluation and Combination of Conditional Quantile Forecasts. *Journal of Business and Economic Statistics*, **23**, 416 – 431.
- Giordani, P., and Söderlind, P. (2003). Inflation forecast uncertainty. *European Economic Review*, **74**, 1037–1060.

- Granger, C. W. J. (1969). Prediction with a generalized cost of error function. *Operations Research Quarterly*, **20**, 199–207.
- Granger, C. W. J., White, H., and Kamstra, M. (1989). Interval forecasting: An analysis based upon ARCH-quantile estimators. *Journal of Econometrics*, **40**, 87–96.
- Hendry, D. F. (2002). Forecast failure, expectations formation, and the Lucas critique. *Econometrics of Policy Evaluation, Special Issue of Annales d’Economie et de Statistique*, **66-67**, 21–40.
- Hommes, C. (2006). Heterogeneous agent models (HAMs) in economics and finance. In Judd, K. L., and Tesfatsion, L. (eds.), *Handbook of Computational Economics, Volume 2, Agent-based Computational Economics*. Amsterdam: North Holland.
- Kandel, E., and Pearson, N. (1995). Differential interpretation of information and trade in speculative markets. *Journal of Political Economy*, **103**, 831–872.
- Kandel, E., and Zilberfarb, B. Z. (1999). Differential interpretation of information in inflation forecasts. *The Review of Economics and Statistics*, **81**, 217–226.
- Keane, M. P., and Runkle, D. E. (1990). Testing the rationality of price forecasts: new evidence from panel data. *American Economic Review*, **80**, 714–735.
- Lahiri, K., and Sheng, X. (2008). Evolution of forecast disagreement in a Bayesian learning model. *Journal of Econometrics*, **144**, 325–340.
- Laster, D., Bennett, P., and Geoum, I. S. (1999). Rational bias in macroeconomic forecasts. *The Quarterly Journal of Economics*, **114(1)**, 293–318.
- Mankiw, N. G., and Reis, R. (2002). Sticky information versus sticky prices: a proposal to replace the new keynesian phillips curve. *Quarterly Journal of Economics*, **117**, 1295–1328.
- Mankiw, N. G., Reis, R., and Wolfers, J. (2003). Disagreement about inflation expectations. mimeo, National Bureau of Economic Research, Cambridge MA.
- Mankiw, N. G., and Shapiro, M. D. (1986). News or noise: An analysis of GNP revisions. *Survey of Current Business (May 1986)*, US Department of Commerce, Bureau of Economic Analysis, 20–25.
- Nordhaus, W. D. (1987). Forecasting efficiency: Concepts and applications. *Review of Economics and Statistics*, **69**, 667–674.
- 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.
- Pesaran, M. H., and Weale, M. (2006). Survey expectations. In Elliott, G., Granger, C., and Timmermann, A. (eds.), *Handbook of Economic Forecasting, Volume 1. Handbook of Economics*

- 24, pp. 715–776: Elsevier, North-Holland.
- Romer, C. D., and Romer, D. H. (2000). Federal Reserve information and the behaviour of interest rates. *American Economic Review*, **90**, 429–457.
- Stekler, H. O. (2002). The rationality and efficiency of individuals' forecasts. In Clements, M. P., and Hendry, D. F. (eds.), *A Companion to Economic Forecasting*, pp. 222–240: Oxford: Blackwells.
- Wallis, K. F. (2003). Chi-squared tests of interval and density forecasts, and the Bank of England's fan charts. *International Journal of Forecasting*, **19**, 165–176.
- Zarnowitz, V. (1969). The new ASA-NBER Survey of Forecasts by Economic Statisticians. *The American Statistician*, **23**, No. 1, 12–16.
- Zarnowitz, V., and Braun, P. (1993). Twenty-two years of the NBER-ASA quarterly economic outlook surveys: aspects and comparisons of forecasting performance. In Stock, J., and Watson, M. (eds.), *Business Cycles, Indicators, and Forecasting*, pp. 11–84: Chicago: University of Chicago Press and NBER.
- Zarnowitz, V., and Lambros, L. A. (1987). Consensus and uncertainty in economic prediction. *Journal of Political Economy*, **95**, 591–621.
- Zellner, A. (1986). Biased predictors, rationality and the evaluation of forecasts. *Economics Letters*, **21**, 45–48.

Table 1: Bounds violations: mean, mode, median and conservative

| Q | # | mean | | | median | | | mode | | | conservative | | |
|---------------|------|--------|-------|-------|--------|-------|-------|--------|-------|-------|--------------|-------|-------|
| | | within | below | above | within | below | above | within | below | above | within | below | above |
| Output growth | | | | | | | | | | | | | |
| 1 | 547 | 75.5 | 4.6 | 19.9 | 71.5 | 5.9 | 22.7 | 81.4 | 6.0 | 12.6 | 89.0 | 2.6 | 7.9 |
| 2 | 586 | 80.7 | 3.8 | 15.5 | 76.1 | 3.8 | 20.1 | 85.0 | 4.1 | 10.9 | 91.1 | 2.4 | 6.5 |
| 3 | 523 | 81.8 | 4.2 | 14.0 | 81.1 | 4.0 | 14.9 | 86.0 | 4.6 | 9.4 | 91.6 | 2.1 | 6.1 |
| 4 | 424 | 87.7 | 3.8 | 8.5 | 81.8 | 5.4 | 12.7 | 84.9 | 5.2 | 9.9 | 91.8 | 3.1 | 5.2 |
| Inflation | | | | | | | | | | | | | |
| 1 | 1045 | 72.0 | 21.4 | 6.6 | 69.2 | 17.5 | 13.3 | 75.2 | 14.9 | 9.9 | 83.9 | 11.2 | 4.6 |
| 2 | 1074 | 75.0 | 19.7 | 5.2 | 73.2 | 14.3 | 12.5 | 78.7 | 12.8 | 8.6 | 86.6 | 9.3 | 4.1 |
| 3 | 893 | 76.3 | 13.7 | 10.1 | 69.9 | 11.1 | 19.0 | 75.1 | 11.0 | 13.9 | 84.1 | 7.7 | 8.2 |
| 4 | 541 | 84.1 | 11.6 | 4.3 | 78.2 | 10.2 | 11.6 | 82.8 | 10.5 | 6.8 | 89.6 | 6.8 | 3.5 |

Output growth. The figures are based on surveys from 1981:3 to 2006:4. 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).

Inflation. Based on surveys 68:4 to 06:4. 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 Fed: ‘Documentation for the Philadelphia Fed’s Survey of Professional Forecasters’, <http://www.phil.frb.org/econ/spf/>. The problematic survey quarters are 1985.1, 1986.1, 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, 1979.4. That these are predominantly Q4 surveys accounts for the smaller number of respondents to Q4 surveys in the table.

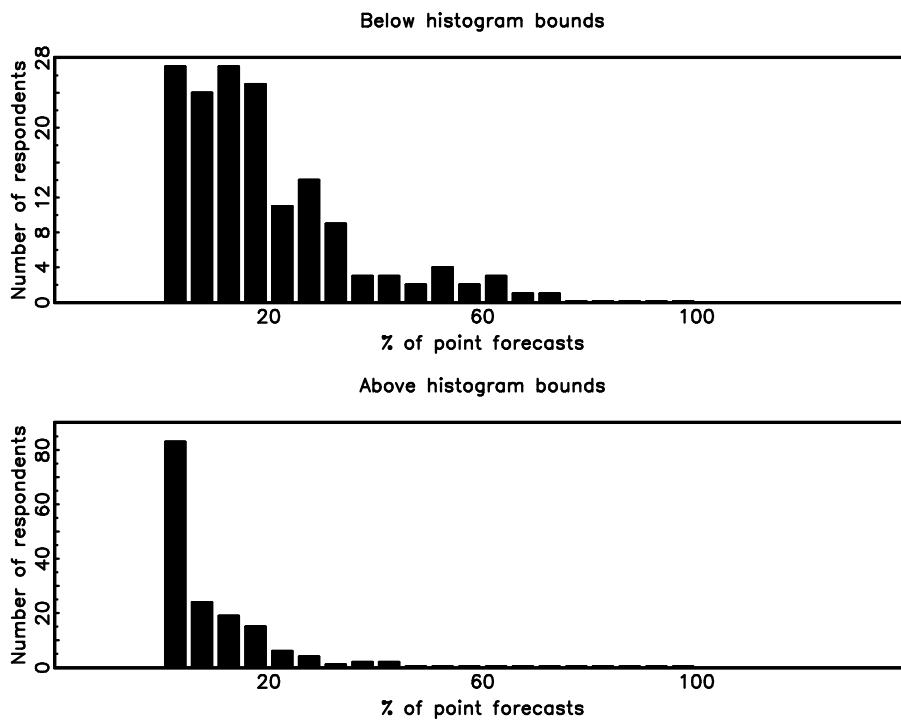


Figure 1: Inflation. The number of respondents with percentages of their point forecasts below (top panel) and above (bottom panel) their bounds, where the bars are less than 5%, 5 to 10% etc.

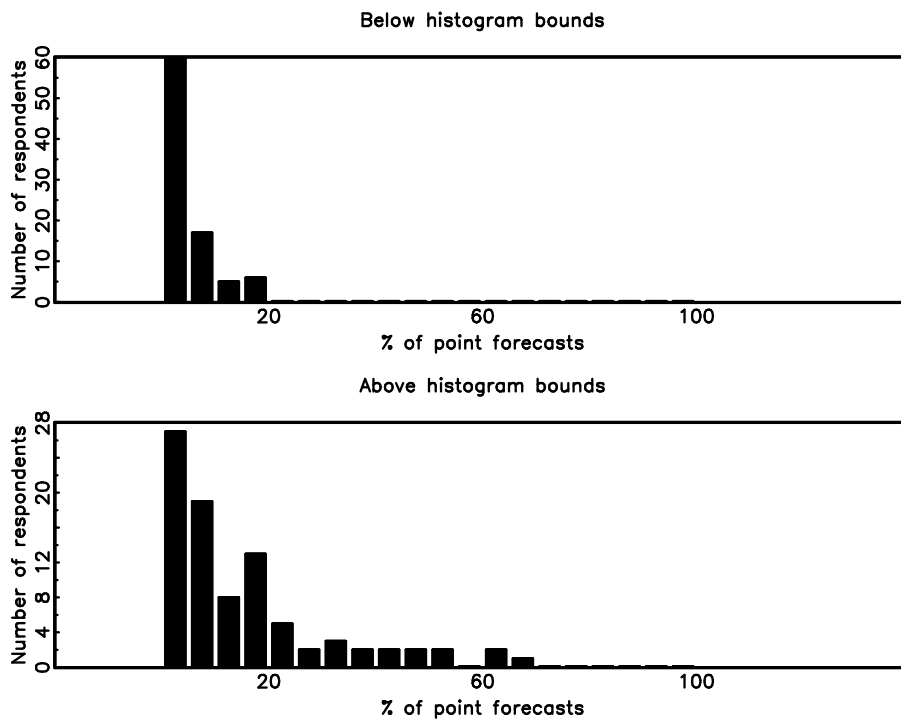


Figure 2: Output Growth. The number of respondents with percentages of their point forecasts below (top panel) and above (bottom panel) their bounds, where the bars are less than 5%, 5 to 10% etc.

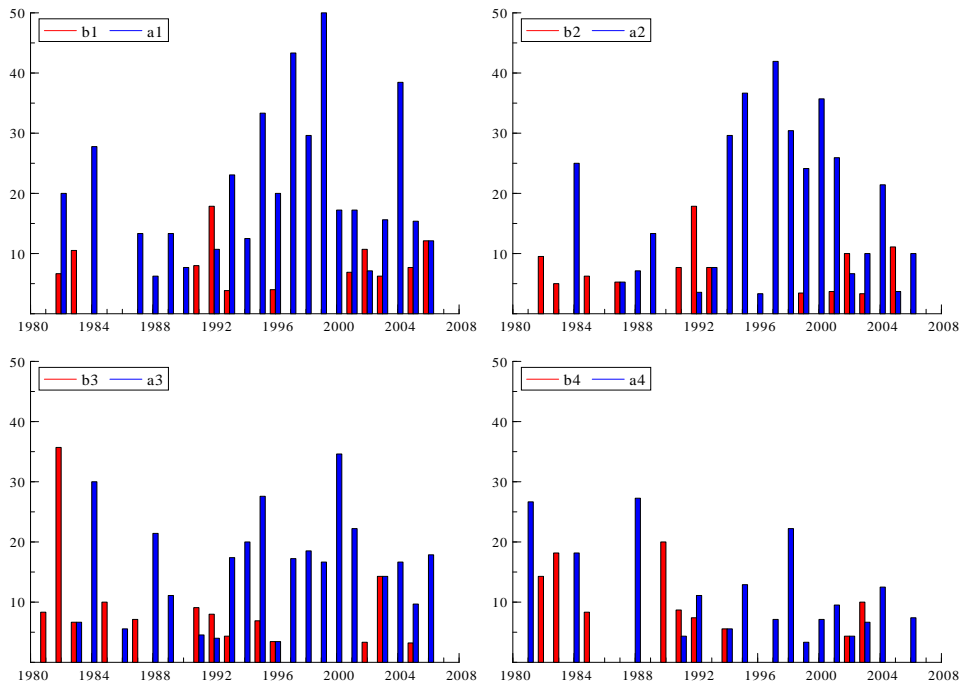


Figure 3: Output growth. For each survey quarter, the two bars report the percentage of individuals whose point forecasts were below their mean bounds ('b'), and the percentage above the bounds ('a'). The top left panel is for forecasts made in the first quarter, the top right for second quarter forecasts etc.

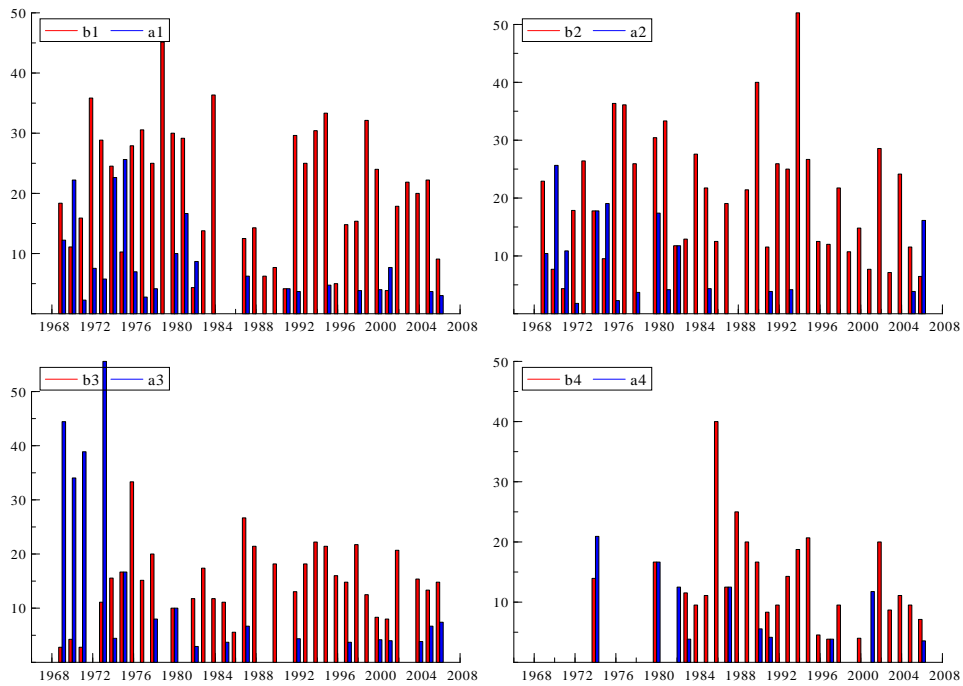


Figure 4: Inflation. For each survey quarter, the two bars report the percentage of individuals whose point forecasts were below their mean bounds ('b'), and the percentage above the bounds ('a'). The top left panel is for forecasts made in the first quarter, the top right for second quarter forecasts etc.

Table 2: Bounds violations and measures of uncertainty

| Survey quarter | forecasts | % point forecasts | | | point forecasts: | | | point forecasts: | | |
|-------------------|-----------|-----------------------|-------|-------|------------------|-------------------|-------|---------------------|-------|-------|
| | | within mean bounds | below | above | within median | below variance | above | within median IQ | below | above |
| Output growth | | | | | | | | | | |
| Q1 | 547 | 75.5 | 4.6 | 19.9 | 0.66 | 0.76 | 1.02 | 1.18 | 1.20 | 1.40 |
| Q2 | 586 | 80.7 | 3.8 | 15.5 | 0.54 | 0.67 | 0.78 | 1.12 | 1.18 | 1.27 |
| Q3 | 523 | 81.8 | 4.2 | 14.0 | 0.34 | 0.88 | 0.77 | 0.99 | 1.42 | 1.30 |
| Q4 | 424 | 87.7 | 3.8 | 8.5 | 0.17 | 0.44 | 0.40 | 0.89 | 1.27 | 1.05 |
| Inflation | | | | | | | | | | |
| Q1 | 1045 | 72.0 | 21.4 | 6.6 | 0.43 | 0.78 | 0.43 | 1.07 | 1.25 | 1.07 |
| Q2 | 1074 | 75.0 | 19.7 | 5.2 | 0.38 | 0.69 | 0.54 | 1.06 | 1.20 | 1.12 |
| Q3 | 893 | 76.3 | 13.7 | 10. | 0.33 | 0.58 | 0.33 | 1.01 | 1.18 | 1.00 |
| Q4 | 541 | 84.1 | 11.6 | 4.3 | 0.18 | 0.38 | 0.43 | 0.96 | 1.06 | 1.06 |

See notes to Table 1.

Table 3: Violations of bounds and the number of non-zero histogram bins

| No. bins | No. forecasts | % within | % below | % above | Q1 | Q2 | Q3 | Q4 |
|-----------------------------|---------------|----------|---------|---------|-----|-----|-----|-----|
| Inflation, 68:4 to 81:2 | | | | | | | | |
| 2 | 199 | 76.88 | 9.55 | 13.57 | 46 | 74 | 62 | 17 |
| 3 | 537 | 70.20 | 14.15 | 15.64 | 180 | 181 | 149 | 27 |
| 4 | 311 | 63.67 | 23.79 | 12.54 | 128 | 105 | 65 | 13 |
| 5 | 160 | 66.25 | 23.75 | 10.00 | 61 | 63 | 31 | 5 |
| 6 | 64 | 54.69 | 35.94 | 9.38 | 32 | 15 | 17 | 0 |
| 7 | 47 | 51.06 | 42.55 | 6.38 | 23 | 14 | 8 | 2 |
| 8 | 30 | 40.00 | 43.33 | 16.67 | 16 | 9 | 5 | 0 |
| 9 | 23 | 34.78 | 43.48 | 21.74 | 11 | 11 | 1 | 0 |
| 10 | 10 | 50.00 | 30.00 | 20.00 | 7 | 1 | 1 | 1 |
| 11 | 4 | 100.00 | 0.00 | 0.00 | 2 | 1 | 1 | 0 |
| 12 | 3 | 33.33 | 33.33 | 33.33 | 2 | 0 | 1 | 0 |
| 13 | . | . | . | . | . | . | . | . |
| 14 | 3 | 33.33 | 66.67 | 0.00 | 1 | 2 | 0 | 0 |
| 15 | 7 | 28.57 | 28.57 | 42.86 | 2 | 3 | 2 | 0 |
| Inflation, 81:3 to 91:4 | | | | | | | | |
| 2 | 280 | 88.6 | 8.9 | 2.5 | 39 | 68 | 74 | 99 |
| 3 | 273 | 86.08 | 10.99 | 2.93 | 52 | 87 | 84 | 50 |
| 4 | 92 | 78.26 | 18.48 | 3.26 | 31 | 29 | 19 | 13 |
| 5 | 37 | 78.4 | 18.9 | 2.7 | 13 | 11 | 9 | 4 |
| 6 | 38 | 52.63 | 47.37 | 0.00 | 15 | 12 | 7 | 4 |
| Inflation, 92:1 to 2006:4 | | | | | | | | |
| 2 | 367 | 88.28 | 10.35 | 1.36 | 53 | 73 | 95 | 146 |
| 3 | 496 | 86.09 | 11.69 | 2.22 | 135 | 129 | 138 | 94 |
| 4 | 270 | 81.48 | 16.30 | 2.22 | 88 | 80 | 59 | 43 |
| 5 | 139 | 74.10 | 23.74 | 2.16 | 49 | 46 | 33 | 11 |
| 6 | 70 | 65.71 | 34.29 | 0.00 | 23 | 31 | 11 | 5 |
| 7 | 44 | 47.73 | 52.27 | 0.00 | 22 | 10 | 10 | 2 |
| 8 | 14 | 35.71 | 42.86 | 21.43 | 2 | 9 | 3 | 0 |
| 9 | 15 | 53.33 | 46.67 | 0.00 | 5 | 6 | 2 | 2 |
| 10 | 20 | 50.00 | 50.00 | 0.00 | 7 | 4 | 6 | 3 |
| Output Growth, 81:3 to 91:4 | | | | | | | | |
| 2 | 200 | 89.00 | 4.00 | 7.00 | 26 | 57 | 64 | 53 |
| 3 | 202 | 86.14 | 5.94 | 7.92 | 55 | 64 | 45 | 38 |
| 4 | 78 | 91.03 | 2.56 | 6.41 | 26 | 30 | 15 | 7 |
| 5 | 36 | 88.89 | 2.78 | 8.33 | 12 | 9 | 8 | 7 |
| 6 | 36 | 63.89 | 22.22 | 13.89 | 15 | 9 | 7 | 5 |
| Output Growth, 92:1 to 06:4 | | | | | | | | |
| 2 | 312 | 89.10 | 3.85 | 7.05 | 28 | 54 | 84 | 146 |
| 3 | 455 | 88.35 | 2.42 | 9.23 | 105 | 117 | 138 | 95 |
| 4 | 282 | 77.66 | 3.19 | 19.15 | 91 | 84 | 66 | 41 |
| 5 | 211 | 72.99 | 4.27 | 22.75 | 79 | 73 | 41 | 18 |
| 6 | 92 | 58.70 | 4.35 | 36.96 | 41 | 28 | 18 | 5 |
| 7 | 71 | 63.38 | 8.45 | 28.17 | 31 | 22 | 14 | 4 |
| 8 | 37 | 72.97 | 0.00 | 27.03 | 14 | 16 | 7 | 0 |
| 9 | 19 | 47.37 | 10.53 | 42.11 | 8 | 7 | 3 | 1 |
| 10 | 49 | 40.82 | 2.04 | 57.14 | 16 | 16 | 13 | 4 |

Output Growth Between 1981:3 and 1991:4 respondents attached probabilities to up to 6 bins, each with width of 2% points. From 1992:1 onwards there were 10 bins with widths of 1% point.

Inflation Between 68:4 and 81:2 there were 15 bins with width of 1 %, between 81:4 and 91:4, 6 bins with width of 2, and from 92:1 onwards 10 bins with a width of 1.

Table 4: Median variance estimates for bounds violations conditional on the number of bins

| No. bins | point forecast | | |
|----------------|-----------------|-------|-------|
| | within | below | above |
| | median variance | | |
| Inflation | | | |
| 68:4 to 81:2 | | | |
| 2 | 0.14 | 0.17 | 0.17 |
| 3 | 0.29 | 0.31 | 0.29 |
| 4 | 0.57 | 0.58 | 0.58 |
| 81:3 to 91:4 | | | |
| 2 | 0.45 | 0.54 | 0.54 |
| 3 | 0.86 | 1.30 | 0.61 |
| 4 | 2.05 | 2.29 | 2.30 |
| 92:1 to 2006:4 | | | |
| 2 | 0.14 | 0.17 | 0.02 |
| 3 | 0.22 | 0.29 | 0.22 |
| 4 | 0.46 | 0.57 | 0.58 |
| Output Growth | | | |
| 81:3 to 91:4 | | | |
| 2 | 0.45 | 0.60 | 0.54 |
| 3 | 0.86 | 1.05 | 1.40 |
| 4 | 1.86 | 1.30 | 2.30 |
| 92:4 to 06:4 | | | |
| 2 | 0.12 | 0.17 | 0.14 |
| 3 | 0.22 | 0.29 | 0.29 |
| 4 | 0.46 | 0.42 | 0.58 |

Table 5: Point forecasts, histogram means and ‘news’

| Dep. variable | | Point forecasts | | | | Histogram means | | | |
|---------------|------------|-----------------|------|----------------|------|-----------------|------|----------------|------|
| | | Fixed effects | | Random effects | | Fixed effects | | Random effects | |
| | | Coeff. | t | Coeff. | t | Coeff. | t | Coeff. | t |
| Output growth | | | | | | | | | |
| Q4 | α_1 | 0.067 | 2.8 | 0.069 | 3.1 | 0.130 | 1.6 | 0.142 | 1.8 |
| | α_2 | 0.934 | 33.6 | 0.933 | 42.0 | 0.890 | 13.0 | 0.886 | 13.0 |
| Q3 | α_1 | 0.069 | 3.1 | 0.094 | 3.6 | 0.186 | 2.7 | 0.224 | 3.1 |
| | α_2 | 0.956 | 29.6 | 0.934 | 25.7 | 0.726 | 15.9 | 0.693 | 13.9 |
| Q2 | α_1 | 0.138 | 2.8 | 0.167 | 3.2 | 0.332 | 4.7 | 0.385 | 6.4 |
| | α_2 | 0.852 | 22.1 | 0.831 | 21.1 | 0.629 | 7.8 | 0.592 | 8.5 |
| Inflation | | | | | | | | | |
| Q4 | α_1 | 0.047 | 1.4 | 0.062 | 1.8 | 0.314 | 2.3 | 0.335 | 2.4 |
| | α_2 | 0.963 | 25.8 | 0.943 | 24.7 | 0.637 | 4.9 | 0.621 | 4.5 |
| Q3 | α_1 | 0.078 | 2.5 | 0.094 | 3.1 | 0.190 | 2.5 | 0.211 | 2.5 |
| | α_2 | 0.917 | 34.8 | 0.901 | 33.9 | 0.867 | 8.3 | 0.817 | 8.1 |
| Q2 | α_1 | 0.127 | 2.7 | 0.152 | 3.4 | 0.198 | 5.6 | 0.264 | 5.6 |
| | α_2 | 0.862 | 19.0 | 0.835 | 19.5 | 0.748 | 17.1 | 0.700 | 12.4 |

Note: We report t -statistics of the significance of the coefficients although the small-sample distribution is unknown for the Random effects estimator.

Table 6: Point forecasts, histogram means and past information

| Dep. variable | Point forecasts | | Histogram means | |
|---------------|-----------------|----------------|-----------------|----------------|
| | Fixed effects | Random effects | Fixed effects | Random effects |
| | Output growth | | | |
| Q4 | 0.000 | 0.000 | 0.001 | 0.000 |
| Q3 | 0.011 | 0.002 | 0.168 | 0.181 |
| Q2 | 0.229 | 0.151 | 0.002 | 0.001 |
| | Inflation | | | |
| Q4 | 0.000 | 0.000 | 0.014 | 0.000 |
| Q3 | 0.000 | 0.000 | 0.035 | 0.099 |
| Q2 | 0.025 | 0.000 | 0.039 | 0.001 |

The entries in the table are the p -values of tests that the four macro-variables are jointly insignificant.

Table 7: Panel regressions explaining the discrepancy between the histogram mean and point forecasts, and the probability of the point forecast being more ‘optimistic’ than the bound on the histogram mean (‘random-effects logistic’)

| | | Random-effects GLS | | | Random-coefficients | | | Random-effects logistic | | |
|---------------|----------|--------------------|----------|-------|----------------------------|----------|-------|----------------------------------|----------|-------|
| | | Coeff. | Std. Err | z | Coeff. | Std. Err | z | Coeff. | Std. Err | z |
| Inflation | | | | | | | | | | |
| Q1 | σ | .449 | .119 | 3.77 | .302 | .192 | 1.58 | 1.73 | .377 | 4.61 |
| | Inf | -.048 | .009 | -5.48 | -.059 | .027 | -2.15 | -.277 | .118 | -2.33 |
| | Con | -.007 | .064 | -0.10 | .0811 | .126 | 0.64 | -2.48 | .379 | -6.55 |
| | | Wald p -value | | | Constancy; Wald p -value | | | Panel vs pooled; Wald p -value | | |
| | | 0.000 | | | 0.006; 0.045 | | | 0.272; 0.000 | | |
| Q2 | σ | .113 | .109 | 1.04 | .131 | .200 | 0.66 | 1.49 | .450 | 3.32 |
| | Inf | -.015 | .025 | -0.60 | -.041 | .042 | -0.98 | -.123 | .126 | -0.98 |
| | Con | .137 | .077 | 1.77 | .154 | .137 | 1.12 | -2.50 | .497 | -5.03 |
| | | Wald p -value | | | Constancy; Wald p -value | | | Panel vs pooled; Wald p -value | | |
| | | 0.571 | | | 0.000; 0.609 | | | 0.007; 0.004 | | |
| Q3 | σ | .255 | .148 | 1.73 | .099 | .185 | 0.54 | 2.79 | .615 | 4.53 |
| | Inf | -.005 | .030 | -0.16 | .005 | .038 | 0.15 | -.354 | .163 | -2.17 |
| | Con | -.015 | .089 | -0.16 | .003 | .139 | 0.02 | -2.89 | .564 | -5.12 |
| | | Wald p -value | | | Constancy; Wald p -value | | | Panel vs pooled; Wald p -value | | |
| | | 0.142 | | | 0.000; 0.841 | | | 0.241; 0.000 | | |
| Q4 | σ | .109 | .067 | 1.62 | .052 | .384 | 0.14 | 1.29 | .546 | 2.36 |
| | Inf | -.032 | .022 | -1.49 | -.060 | .056 | -1.07 | -.288 | .196 | -1.47 |
| | Con | .092 | .061 | 1.52 | .151 | .216 | 0.70 | -2.24 | .456 | -4.90 |
| | | Wald p -value | | | Constancy; Wald p -value | | | Panel vs pooled; Wald p -value | | |
| | | 0.117 | | | 0.000; 0.543 | | | 1.00; 0.046 | | |
| Output growth | | | | | | | | | | |
| Q1 | σ | -.169 | .122 | -1.38 | -.202 | .184 | -1.10 | 1.85 | .544 | 3.40 |
| | Out | -.067 | .017 | -4.02 | -.058 | .018 | -3.16 | .330 | .091 | 3.61 |
| | Con | .186 | .115 | 1.62 | .208 | .153 | 1.36 | -4.70 | .786 | -5.99 |
| | | Wald p -value | | | Constancy; Wald p -value | | | Panel vs pooled; Wald p -value | | |
| | | 0.000 | | | 0.000; 0.007 | | | 0.000; 0.000 | | |
| Q2 | σ | -.133 | .103 | -1.29 | -.166 | .188 | -0.88 | 1.28 | .586 | 2.18 |
| | Out | -.047 | .019 | -2.41 | -.044 | .016 | -2.81 | .394 | .110 | 3.57 |
| | Con | .075 | .120 | 0.63 | .142 | .138 | 1.03 | -4.81 | .818 | -5.88 |
| | | Wald p -value | | | Constancy; Wald p -value | | | Panel vs pooled; Wald p -value | | |
| | | 0.029 | | | 0.000; 0.017 | | | 0.000; 0.000 | | |
| Q3 | σ | -.035 | .073 | -0.47 | .156 | .200 | 0.78 | 1.14 | .646 | 1.76 |
| | Out | -.057 | .039 | -1.44 | -.037 | .034 | -1.09 | .350 | .118 | 2.98 |
| | Con | .047 | .124 | 0.38 | -.098 | .1149 | -0.85 | -3.97 | .643 | -6.18 |
| | | Wald p -value | | | Constancy; Wald p -value | | | Panel vs pooled; Wald p -value | | |
| | | 0.347 | | | 0.000; 0.514 | | | 0.083; 0.004 | | |
| Q4 | σ | -.121 | .092 | -1.31 | .372 | .215 | 1.73 | -.163 | 1.028 | -0.16 |
| | Out | .016 | .028 | 0.58 | .014 | .029 | 0.50 | .073 | .170 | 0.43 |
| | Con | -.072 | .142 | -0.51 | -.276 | .157 | -1.76 | -3.46 | .935 | -3.70 |
| | | Wald p -value | | | Constancy; Wald p -value | | | Panel vs pooled; Wald p -value | | |
| | | 0.105 | | | 0.000; 0.222 | | | 0.022; 0.902 | | |

The columns headed ‘ z ’ are the ‘ t -statistics’, where z denotes that the small-sample distributions are unknown. The ‘Wald p -values’ are for the tests of the joint insignificance of the regressors other than the constant (‘Con’); where σ is the standard deviation of histogram; and Inf and Out are the actual quarterly inflation rates and output growth (at an annual rate) in the previous quarter. ‘Constancy’ is the p -value of a test that parameter vector is the same over all individuals. ‘Panel vs pooled’ is the p -value of a likelihood ratio test that the panel-dimension is unimportant in the panel-logit regression, in the sense that the pooled estimator and panel estimator are equivalent.

Table 8: Accuracy of point forecasts and histogram means

| | No. of forecasts | Point forecast | Histogram mean | Histogram mean Normal approx. | Histogram mean 'bounds' |
|---------------|------------------|----------------|----------------|-------------------------------|-------------------------|
| Inflation | | | | | |
| 68:4 to 81:2 | | | | | |
| All | 1410 | 1.37 | 1.61 | 1.66 | |
| point > bound | 192 | 1.99 | 3.32 | 3.48 | 2.28 |
| point < bound | 285 | 1.72 | 2.08 | 2.07 | 1.30 |
| 81:3 to 91:4 | | | | | |
| All | 783 | 0.40 | 0.87 | 0.76 | |
| point > bound | 21 | 1.46 | 1.07 | 1.46 | 0.13 |
| point < bound | 99 | 0.96 | 3.54 | 3.01 | 1.18 |
| 92:1 to 06:4 | | | | | |
| All | 1507 | 0.12 | 0.27 | 0.25 | |
| point > bound | 28 | 0.11 | 0.42 | 0.50 | 0.07 |
| point < bound | 243 | 0.16 | 0.89 | 0.76 | 0.38 |
| Output growth | | | | | |
| 81:3 to 91:4 | | | | | |
| All | 607 | 0.58 | 1.08 | 1.08 | |
| point > bound | 43 | 0.69 | 4.27 | 4.12 | 1.82 |
| point > bound | 33 | 0.56 | 3.12 | 3.01 | 0.99 |
| 92:1 to 06:4 | | | | | |
| All | 1606 | 0.39 | 0.55 | 0.54 | |
| point > bound | 267 | 0.61 | 1.31 | 1.20 | 0.60 |
| point < bound | 54 | 0.40 | 0.52 | 0.61 | 0.18 |

The actual values used to calculate the forecast errors are the second-release real-time data.