Uncertainty and disagreement in economic prediction: the Bank of England Survey of External Forecasters

Gianna Boero, Jeremy Smith and Kenneth F. Wallis*

Department of Economics
University of Warwick
Coventry CV4 7AL, UK

April 2006

Abstract This paper introduces a new source of survey data, namely the Bank of England Survey of External Forecasters. The survey collects point and density forecasts of inflation and GDP growth, and hence offers the opportunity of constructing direct measures of uncertainty. We present a simple statistical framework in which to define and interrelate measures of uncertainty and disagreement. The resulting measures are compared with other direct measures of uncertainty, nationally and internationally. A major, sustained reduction in inflation uncertainty followed the 1997 granting of operational independence to the Bank of England to pursue a monetary policy of inflation targeting.

Keywords Forecast surveys; point forecasts; density forecasts; uncertainty; disagreement

JEL classification C42, C53, E37

Acknowledgements We are grateful to Colin Ellis, Debbie Lockhart and Rebecca Capuano (Bank of England) for assembling and helping to clean the survey dataset, to Paul Soderlind (University of St Gallen) for providing extended data from one of his articles, and to Jagjit Chadha for helpful comments following the presentation of some preliminary material in the third author’s Sir Richard Stone lectures, October 2005.

*Corresponding author [K.F.Wallis@warwick.ac.uk]
1. INTRODUCTION

Uncertainty plays an important role in many areas of economic behaviour. Three leading examples concern the effects of demand uncertainty on firms’ investment decisions, the effects of income uncertainty on household consumption and saving, and the effects of various kinds of uncertainty on monetary policy-making. In most of these areas the literature contains not only theoretical models but also empirical analyses, where a basic requirement is the construction of relevant measures of uncertainty. Direct measures are relatively rare; moreover typically “there is no consensus about the appropriate way to proxy uncertainty in an empirical formulation,” as Carruth et al. (2000, p.133) conclude in their review of the literature on investment under uncertainty. There are exceptions, however, in particular in surveys that elicit expectations about future economic outcomes in the form of probability distributions – so-called density forecasts. Our purpose in this paper is to introduce a new source of survey data, relating to the United Kingdom, namely the Bank of England Survey of External Forecasters.

The best-known series of density forecasts in macroeconomics dates from 1968, when the American Statistical Association and the National Bureau of Economic Research jointly initiated a quarterly survey of macroeconomic forecasters in the United States, known as the ASA-NBER survey; Zarnowitz (1969) describes its original objectives. In 1990 the Federal Reserve Bank of Philadelphia assumed responsibility for the survey, and changed its name to the Survey of Professional Forecasters (SPF). Survey respondents are asked not only to report their point forecasts of several variables but also to attach a probability to each of a number of pre-assigned intervals, or bins, into which output growth and inflation, this year and next year, might fall. In this way, respondents provide density forecasts of these two variables, in the form of histograms. The probabilities are then averaged over respondents to obtain the mean or combined density forecasts, again in the form of histograms, and these are included in the quarterly publication of forecasts. Similar questions about inflation have been included in the Bank of England Survey of External Forecasters (SEF) since 1996, and about GDP growth since 1998; the combined density forecasts likewise feature in the Bank’s quarterly Inflation Report.

The individual SPF responses, suitably anonymised, and with a short delay, have been made available for research purposes, in accordance with one of the survey’s original
objectives, to construct an archive that would support scientific study. The resulting literature includes studies of the relationship between direct measures of uncertainty based on density forecasts, and disagreement among point forecasts, by Zarnowitz and Lambros (1987), updated and extended by Giordani and Soderlind (2003); the relationship between inflation uncertainty and real interest rates (Lahiri, Teigland and Zaporowski, 1988; Batchelor and Dua, 1996), the latter authors also comparing the direct measure of inflation uncertainty with various proxies used in empirical studies; and the relationship between uncertainty and the equity premium puzzle in the consumption-based asset pricing model (Giordani and Soderlind, 2006). Recently the Bank of England has likewise agreed to make available the anonymised individual SEF responses, and a first analysis of these is presented in this paper. We look forward to a similarly wide range of applications based on UK data in due course, although the available time series are as yet relatively short.

Further studies of uncertainty that use direct measures based on density forecasts reported by survey respondents can be found at the micro level, usually as a result of adding a special question on a particular occasion to an ongoing large-scale sample survey of households or firms. Thus Guiso, Jappelli and Terlizzese (1992) test for the presence of precautionary saving using a measure of uncertainty based on households’ density forecasts of their earnings, and Guiso and Parigi (1999) investigate the effects of uncertainty on investment decisions using a measure based on firms’ density forecasts of the demand for their product. In both cases the data source is a one-off question included in an annual Bank of Italy survey. Similar studies of household income expectations can be found in the United States (Dominitz and Manski, 1997) and the Netherlands (Das and Donkers, 1999). These articles contain a good discussion of the reliability of individual subjective distributions, facing the objection that people have no incentive to answer the questions carefully, nor indeed the requisite ability; see also the survey by Manski (2004). Nevertheless Batchelor (2005) argues that models from the earnings forecasting literature are of little relevance to macroeconomic forecasters. The SEF respondents are professional economists, regularly providing forecasts and analysis to economic and financial decision makers, and users rely on their professionalism and experience in supplying careful answers. The anonymity of the survey implies that there are no returns to variety-seeking or branding behaviour, although we find occasional mistakes, as discussed below, which suggests the use of statistical methods that are robust to the potentially excessive influence of outliers.
The remainder of this paper is organised as follows. Section 2 provides general information about the survey and the dataset of individual responses. Section 3 presents some possible direct measures of uncertainty and disagreement, and the results of implementing them in this dataset. Comparisons of these measures with other direct measures of uncertainty, nationally and internationally, are presented in Section 4. Section 5 concludes.

2. THE BANK OF ENGLAND SURVEY OF EXTERNAL FORECASTERS

Every quarter, the Bank of England asks a group of external forecasters for their views on inflation, output growth, interest rates and the sterling exchange rate. The institutions covered in the sample include City firms, academic institutions and private consultancies, and are predominantly based in London. The sample changes from time to time as old respondents leave or new survey members are included, and not every institution provides a forecast to the Bank every quarter. There is no record of the response rate, and only one institution is ever-present. The survey is carried out just before the quarterly Inflation Report, and a summary of the results is reported in a box. Each chart or table includes a note of the number of responses on which it is based, and it can be seen that not all respondents answer every question. The Inflation Report is published each February, May, August and November, and we refer to individual surveys by these dates, although the questionnaires are sent out during the preceding month.

Initially the Bank asked the forecasters only about their central projections for RPIX inflation. From February 1996 the forecasters were also asked about the probabilities they attach to various possible RPIX inflation outcomes, while the definition of inflation switched to the CPI with the February 2004 survey, following the change in the Bank’s official targeted measure in December 2003. Questions about point and density forecasts of GDP growth have appeared since February 1998, and about point forecasts of the official interest rate and the sterling exchange rate index since November 1999. Given our focus on density forecasts these last two variables are not considered in the present paper. The density forecast questions divide the range of each variable into a number of bins, which has varied between four and six. For annual inflation (percent), initially five bins were specified.
(<1, 1-2.5, 2.5-4, 4-5.5, >5.5), from February 1998 reduced to four (<1.5, 1.5-2.5, 2.5-3.5, >3.5), finally from May 1999 increased to six (<1.5, 1.5-2, 2.5-3, 3-3.5, >3.5), each change reducing the width of the interior intervals. For GDP growth, six bins (<0, 0-1, 1-2, 2-3, 3-4, >4) were specified until May 2003, whereupon they were reduced to four by combining each pair of outer bins (<1, 1-2, 2-3, >3). Our dataset of individual responses covers 39 surveys, beginning with the May 1996 survey and continuing to November 2005. The number of forecasters providing at least one density forecast in a given survey ranges between 23 and 38, and is mostly in the upper twenties.

Each quarterly survey since February 1998 asks for forecasts at three future points in time: the fourth quarter (Q4) of the current year; the fourth quarter of the following year; and the corresponding quarter two years ahead. (For the first seven surveys in our sample, which relate only to inflation, only the first two questions appear.) This structure eventually delivers nine successive forecasts of a given Q4 outcome, which can be treated as “fixed-event” forecasts, with the date of the forecast preceding the date of the outcome by 8, 7, …, 1, 0 quarters. Given that the survey goes out early in the quarter, when no data on current-quarter inflation and growth are available, we treat these as $h$-step-ahead forecasts with horizon $h$ equal to 9, 8, …, 2, 1 quarters successively. For time series of “fixed-horizon” forecasts, the third question delivers a quarterly series of nine-quarter-ahead forecasts, but the first two questions give only one observation per year at intermediate horizons, $h = 4$ and 8 in February, $h = 3$ and 7 in May, and so on. This end-year focus is clearly more familiar to forecasters, since there are usually a few respondents who provide probability distributions in answer to the first two questions but not to the third question.

In over five thousand individual densities, there are nine in which the reported probabilities, with allowance for rounding, do not sum to 100%. (This is a much lower relative frequency of such mistakes than that reported by Zarnowitz and Lambros, 1987, for the SPF.) Three contain fairly obvious typographical errors, which we have corrected. In the remaining six cases the total probability differs from 100% by no more than 10%; here we have amended one or two entries, with reference to the respondent’s other densities reported at the same time and densities reported for the same outcome in adjoining surveys. These investigations uncovered a further case in which it is clear that the probabilities have been entered in reverse order, and this too has been corrected. Details of these ten adjustments to the dataset are available to other users on request.
3. MEASURES OF UNCERTAINTY AND DISAGREEMENT

The published SEF results include the distribution of point forecasts, in the form of a histogram, and the aggregate or mean density forecasts, in tabular form. We first present a simple statistical framework in which to introduce, and interrelate, direct measures of uncertainty and disagreement, respectively identified as measures of the dispersion of the aggregate and individual forecast densities and the distribution of point forecasts. Results are then presented, based in turn on the survey sequence and the fixed-event and fixed-horizon schemes discussed above; the standard error of our preferred measure of collective uncertainty is also reported.

3.1 The statistical framework

Denote \( n \) individual density forecasts of a random variable \( Y \) at some future time as \( f_i(y) \), \( i = 1, \ldots, n \). In the SEF these are expressed numerically, but the statistical framework also accommodates density forecasts that are expressed analytically, for example, via the normal or two-piece normal distributions. For economy of notation time subscripts and references to the information sets on which the forecasts are conditioned are suppressed. The published average or aggregate density forecast is then

\[
f_A(y) = \frac{1}{n} \sum_{i=1}^{n} f_i(y). \tag{1}
\]

From a statistical point of view, this is an example of a finite mixture distribution (Wallis, 2005). The moments about the origin of \( f_A(y) \) are given as the corresponding equally-weighted combinations of the moments about the origin of the individual densities. We assume that the individual point forecasts are the means of the individual forecast densities and so denote these means as \( \hat{y}_i \); the individual variances are \( \sigma_i^2 \). Then the mean of the aggregate density is

\[
\mu_1' = \frac{1}{n} \sum_{i=1}^{n} \hat{y}_i = \hat{y}_A, \tag{2}
\]

namely the average point forecast, and the second moment about the origin is

\[
\mu_2' = \frac{1}{n} \sum_{i=1}^{n} \left( \hat{y}_i^2 + \sigma_i^2 \right).
\]

Hence the variance of \( f_A \) is...
\[ \sigma_A^2 = \mu_A^2 - \mu_i^2 = \frac{1}{n} \sum_{i=1}^{n} \sigma_i^2 + \frac{1}{n} \sum_{i=1}^{n} (\hat{y}_i - \hat{y}_A)^2. \]  

This expression decomposes a measure of aggregate uncertainty, \( \sigma^2_A \), into the average individual uncertainty or variance, plus a measure of the dispersion of, or disagreement between, the individual point forecasts. It shows that the disagreement measure, often suggested as a proxy variable for uncertainty in the absence of a direct measure, is indeed a component of the direct measure based on the aggregate density forecast.

For a measure of collective uncertainty, the general approach in the SPF literature is to choose a measure of average individual uncertainty, rather than a measure based on the aggregate distribution, since the former is free of the influence of disagreement about point forecasts. This is implicit in the work of Zarnowitz and Lambros (1987), who use the average individual standard deviation and the standard deviation of point forecasts as their respective measures of uncertainty and disagreement. The use of standard deviations rather than variances breaks equation (3), although Zarnowitz and Lambros seem unaware of this decomposition of the variance of the aggregate distribution. Giordani and Soderlind (2003) follow suit, explicitly arguing against the use of the aggregate distribution, whose standard deviation they nevertheless calculate for comparative purposes. In general it is more informative to present standard deviations, not variances, since their units coincide with the units of the variable under consideration, percentage points of inflation or growth. However our preferred measure of collective uncertainty is then defined as the square root of the first term on the right-hand side of equation (3), and not as the average individual standard deviation. Batchelor and Dua (1996) call the preferred measure the root mean subjective variance (RMSV), and note that it is more consistent with an assumption of quadratic utility than the average standard deviation.

Several practical considerations arise in calculating these measures. The first question is how to estimate means and variances from histograms with rather few bins, between four and six in the SEF data, with the first and last being open-ended. A simple approach, referred to by several authors as providing “crude” estimates, is to apply standard formulae assuming that the reported probabilities are concentrated at the mid-points of the respective intervals, and that the open intervals have an assumed finite width, often equal to that of the interior intervals. An alternative assumption is that the reported probabilities are spread uniformly
across each bin, which has no effect on the calculated mean but increases the variance by one-twelfth of the squared bin width. Such a correction would apply equally to the aggregate variance and the average individual variance appearing in equation (3). However Giordani and Soderlind (2003) note that the SPF histograms “often look fairly bell shaped” and so find it more reasonable to assume that more of the probability mass within an interval is located closer to the overall mean, which leads them to estimate means and variances by fitting normal distributions to the histograms. In the SEF data the aggregate distributions often look quite normal, and although the individual densities differ considerably in their dispersion, they are invariably unimodal, with the mode at an interior interval (except for GDP growth since May 2003, when the reduction to four bins results in some J-shaped individual distributions). We therefore consider the same statistical approach.

For a histogram with \( k \) intervals, denote the upper boundaries of the first \( k - 1 \) intervals as \( b_j, j = 1, \ldots, k - 1 \), and the cumulative probabilities at these boundaries as \( \theta_j, j = 1, \ldots, k - 1 \). Assume that the first and last intervals have non-zero probability, so that \( \theta_1 > 0 \) and \( \theta_{k-1} < 1 \). If the distribution were normal, then the standardised boundaries would be the corresponding quantiles of the standard normal distribution, that is,

\[
\theta_j = \Phi\left(\frac{b_j - \mu}{\sigma}\right), \quad j = 1, \ldots, k - 1,
\]

where \( \Phi(\cdot) \) is the standard normal distribution function. This suggests estimating the mean and standard deviation by the corresponding non-linear least squares regression of the cumulative probabilities on the interval boundaries. No assumption about the open-ended intervals is necessary. On the other hand there may be insufficient degrees of freedom if \( k \) is small and/or some intervals have zero probability. In this respect the GDP density forecast histograms since May 2003, which have \( k = 4 \), are particularly at risk.

For some purposes we wish to calculate the three elements of equation (3) independently, but in such a way that the equation holds. In practice this requires the use of the crude variance estimates, with the point forecasts taken to be the corresponding crude means. For other purposes we may prefer to obtain the “best” estimates of each separate component, by relaxing the restriction, whereupon variance estimates based on normal approximations come into consideration. Moreover measures of disagreement based on reported point forecasts, rather than density forecast means, are relevant to the question of the
performance of such indirect measures of uncertainty in the absence of direct measures, and in the data differences between point forecasts and density forecast means are observed. Also there are usually some respondents who provide point forecasts but no density forecasts, and we may not wish to exclude them from measures of disagreement. We consider two approaches to measuring disagreement: first, the usual variance or standard deviation of the point forecasts can be calculated; second, following Giordani and Soderlind (2003), a quasi-standard deviation may be calculated as half the difference between the 16th and 84th percentiles of the point forecasts, appropriately interpolated. For a normal distribution this calculation delivers the standard deviation, otherwise it is more robust to outliers. Finally we note that estimating individual variances by either the crude method or via normal approximations presents difficulties when the distribution is highly concentrated, with non-zero probability assigned to at most two intervals. In the absence of individual variance estimates, a “quasi-direct” estimate of the average individual variance can nevertheless be obtained from equation (3), as the difference between aggregate uncertainty and disagreement, using whichever of the possible measures of these are preferred.

3.2 Uncertainty and disagreement, question-by-question

Each Inflation Report presents aggregate results for the current survey, but only very occasionally includes consideration of these in relation to previous surveys. We first present time series of survey results by question, in effect time series of Inflation Report boxes. Crude mean and variance estimates are used, so that the separately calculated direct measures of uncertainty and disagreement satisfy equation (3). These are presented in standard deviation form for each of the three questions and two variables in the six panels of Figure 1.

It is seen that the uncertainty measures obtained from the responses to the first question show a strong seasonal pattern, which is less pronounced in question 2 and absent in question 3. This reflects the variation in forecast horizon across the four surveys each year, which is successively 4, 3, 2, 1, 4, 3, 2, 1, … quarters for question 1, and 8, 7, 6, 5, 8, 7, 6, 5, … quarters for question 2, but constant at 9 quarters for question 3. The mean square forecast error for relevant time series models falls much less between 8- and 5-quarter-ahead forecasts than it does between 4- and 1-quarter ahead forecasts. There is no similar pattern in any of the disagreement (dispersion of mean forecasts) series. The declining pattern is emphasised when we consider fixed-event forecasting schemes below, and is in turn removed
when we consider time series of fixed-horizon forecasts. (The lowest panels of Figure 1 are already of this form, with \( h = 9 \).)

There are two changes in the general level of the uncertainty series that catch the eye. First, with respect to inflation, is the reduction in uncertainty in the early part of the sample period, until the May 1999 survey, after which the general level is approximately constant. There may have been some initial learning about how to answer the new question on the part of the forecasters, but the major cause is undoubtedly the increasing confidence in, and credibility of, the new monetary policy arrangements. The operational independence of the Bank of England was announced within days of the change of government in the 1 May 1997 general election, and the newly-established Monetary Policy Committee (MPC) first met on 5-6 June 1997. The enabling legislation gradually made its way through Parliament, and the Bank of England Act came into force on 1 June 1998. Market-based measures of inflation expectations fell sharply following the announcement of the new monetary policy arrangements, and continued to fall, reaching a level slightly under the Bank’s official target of 2.5% in late 1998. When the Survey of External Forecasters first asked for forecasts two years ahead, in February 1998, the median point forecast of inflation at that horizon was already close to the target. Our measures of uncertainty provide further evidence of the rapid gains in monetary policy credibility over this period.

The second prominent reduction in uncertainty, with respect to GDP growth in May 2003, has a more prosaic explanation. It coincides with the reduction in the number of histogram bins from six to four noted above, which was achieved by combining the outer, open-ended bins with their nearest neighbours, leaving only two interior bins. Whereas the crude calculation of variances is generally insensitive to assumptions about the width of the open-ended bins, because the associated probabilities are small, on this occasion this condition does not apply. In the aggregate density, the last two bins for question 3 in February 2003 have a combined probability of 26%, while in May 2003 the single, newly combined upper bin has a probability of 22%. In some of the individual densities, corresponding probabilities of 40% and above are observed. Placing weights of this order of magnitude on assumed mid-points of 3.5 and 4.5, separately, in February, then on 3.5 in total in May, results in the observed reduction. Using an alternative assumption about the width of the open-ended intervals, that they are twice as wide as the interior intervals, or estimating variances via fitted normal distributions, removes this anomaly, as shown below. With the
MPC’s forecasts, as published in the *Inflation Report*, showing non-negligible probabilities of GDP growth exceeding 4%, it is hard to understand the reasons for the change in the basis of the SEF histograms.

The uncertainty measures shown in Figure 1, namely the crude estimates of the aggregate standard deviation, $\sigma$, and the RMSV, are little different from one another; equivalently the disagreement measure, based on the estimated means of the density forecasts, is relatively small. Apart from the seasonal patterns and the two specific changes in level discussed above, the uncertainty measures appear to fluctuate randomly around a constant level over most of the sample period. In the first seven surveys the two inflation disagreement measures are generally higher than their subsequent levels, but from February 1998, when the third inflation question and the GDP questions first appear, none of the disagreement measures show any systematic variation.

### 3.3 Fixed-event forecasts

We use the fixed-event scheme to study the effect of forecast horizon on measures of uncertainty and disagreement. Except for the first two years of our dataset, there are nine successive forecasts of each Q4 outcome, as noted above. For example, for the 2001Q4 outcomes, the first (nine-step-ahead) forecasts are given in response to question 3 in the November 1999 survey, then 8, 7, 6, and 5-step-ahead forecasts are given in response to question 2 in the four successive surveys in 2000, finally 4, 3, 2 and 1-step-ahead forecasts in the four 2001 surveys. We reassemble comparable information from the three panels of Figure 1, and present it in Figure 2, as separate trajectories for the nine available Q4 inflation outcomes, 1997-2005, and the seven available Q4 growth outcomes, 1999-2005. The measure of uncertainty (upper panels) is the aggregate standard deviation $\sigma_A$, estimated by fitting normal distributions to the histograms. The measure of disagreement (lower panels) is based on the reported point forecasts, that is, the data that might be available in the absence of a direct measure of uncertainty based on density forecasts, and that might be used to construct a proxy for it. The particular measure we choose is the quasi-standard deviation defined above, which robustifies each series on two or three occasions.

Looking first at the inflation panels of Figure 2, we note that the earliest trajectory, for the 1997Q4 outcome, starts, with our dataset, in May 1996, at $h = 7$, while the next two start
at $h = 8$, since there was no question 3 in the November 1996 and 1997 surveys. Beginning with the uncertainty measures, the first two trajectories show substantially higher levels of inflation uncertainty, with the second, 1998Q4, trajectory falling dramatically between November 1997 and February 1998. The next two show less dramatic reductions from their initial levels, and from May 1999 onwards the plots are remarkably close to one another. This is an alternative presentation of the rapid gains in monetary policy credibility described above. The five more recent trajectories are very similar, and show a steady decline in uncertainty as the forecast horizon is reduced, as described for various models in forecasting textbooks. For these five, the standard deviation is reduced from approximately 0.6 percentage points at $h = 9$ to 0.35 at $h = 1$.

Forecasting textbooks give little guidance on disagreement measures, however, and there are some similarities and some differences between the two inflation panels. The first two disagreement trajectories also start high and then decline, although the earliest has a peak in the second survey in 1997, when there may have been little disagreement about the result of the forthcoming general election, but widely dispersed views about its consequences for inflation. In all nine cases, there is more disagreement at the beginning ($h = 9$ or slightly less) than at the end ($h = 1$), although in no case is the decline monotonic.

The GDP growth forecasts begin in 1998 and do not exhibit the early reductions in uncertainty associated in the inflation forecasts with the gaining of monetary policy credibility. Relatively steady reductions in uncertainty are observed across the fixed-event forecasts. Uncertainty in growth is at a higher level than in inflation, however, and its dispersion across the various fixed-event forecasts is greater than in the more recent fixed-event inflation forecasts. Disagreement in the growth forecasts is similarly higher and noisier than in the inflation forecasts, although the same general tendency, of a reduction in disagreement as the forecast horizon reduces, is observed.

### 3.4 Fixed-horizon series

Fixed-horizon time series are free of the effects of varying lead time but, as noted above, quarterly series are available only for $h = 9$; at each intermediate horizon, $h = 1, \ldots, 8$, only an annual series is available, for each variable. Nevertheless this is the most appropriate format
in which to compare the SEF measures with direct measures from other sources, which is the subject of the next section.

Correlation analysis of the quarterly constant-horizon series (at \( h = 9 \)) shows that disagreement is not a good proxy for collective uncertainty, \( RMSV \). For inflation we focus on the period since May 1999, when the histograms have six bins and are more informative: the correlation between the robust disagreement measure and \( RMSV \) is 0.26. In an attempt to obtain comparable sample sizes at shorter horizons we deseasonalise the data from questions 1 and 2 by dummy variable regression, to correct for variation in forecast horizon within each series. The correlations between disagreement and collective uncertainty are then similar to that reported above, namely 0.28 for the adjusted question 1 series and 0.19 for question 2, confirming that disagreement is not a good proxy for uncertainty. For GDP growth forecasts, over the complete period for which they are available, the correlation between disagreement and collective uncertainty at \( h = 9 \) is 0.10; the equivalent coefficients for the adjusted questions 1 and 2 series are respectively 0.17 and 0.09, yielding the same conclusion.

3.5 Disagreement about uncertainty

The individual forecast histograms represent a considerable spread of opinion, as noted above, indicating a lack of agreement about forecast uncertainty among the sample of forecasters. This is illustrated by the example in the upper panel of Figure 3, which shows the 21 individual inflation histograms at \( h = 9 \) reported in the November 2005 survey. These are ordered by their individual dispersion, which reduces as we move from the front to the back of the array. Some histograms towards the front are distorted by the replacement of the open-ended intervals by closed intervals of the same width as the interior intervals: clearly the first forecast density shown has tails extending well beyond the range used in this illustration. Towards the back of the array there are a few responses with non-zero probability in only two interior intervals.

To summarise the impact of this sample variation on our collective uncertainty measure, \( RMSV \), we calculate its standard error under the conventional random sampling assumption. We denote the average individual variance, the first term on the right-hand side of equation (3), as \( \sigma^2 \).
\[
\overline{\sigma^2} = \frac{1}{n} \sum_{i=1}^{n} \sigma_i^2 ,
\]
then the RMSV is its square root,
\[
RMSV = \sqrt{\overline{\sigma^2}} .
\]
Using the standard expression for the variance of the sample mean, and taking account of the square root transformation, the standard error of RMSV is then obtained as
\[
\text{se}(RMSV) = \sqrt{\frac{\sum (\sigma_i^2 - \overline{\sigma^2})^2}{4n(n-1)\sigma^2}} .
\]
Again using an inflation example, we plot a band of two standard errors on either side of the RMSV estimate in the lower panel of Figure 3, which refers to the quarterly constant-horizon series at \( h = 9 \), taken from the lower left panel of Figure 1. It is seen that taking account of sampling variation in this way strengthens the conclusion that inflation uncertainty has fluctuated randomly around a constant level once credibility of the new monetary policy regime had been achieved.

4. COMPARISONS WITH MEASURES FROM OTHER SOURCES

4.1 The Monetary Policy Committee’s forecasts
Each quarterly Inflation Report includes the famous fan charts for inflation and growth, representing the MPC’s density forecasts up to two years ahead. (GDP growth forecasts first appeared in November 1997; the horizon for both variables was extended to three years in August 2004.) Numerical information used in the construction of the distributions described by the fan charts is published on the Bank of England website (http://www.bankofengland.co.uk/publications/inflationreport/irprobab.htm). It includes an uncertainty measure defined as the “input” standard deviation of the distribution. We treat this measure, agreed by the MPC, as the average individual uncertainty across the nine-member committee, and compare it to the average individual uncertainty, RMSV, in the SEF. Comparisons are made at three forecast horizons, \( h = 1,5,9 \), since these forecasts, current-quarter, one-year-ahead and two-years-ahead, appear to be the ones that receive most
attention from the MPC, with intermediate forecasts then obtained by interpolation. Results
for inflation and growth are shown in Figure 4.

The overall level of external forecasters’ uncertainty is higher than that of the MPC at
shorter horizons, while the time series plots at $h = 9$ intermingle, for both variables. The
general practice of the MPC is first to calibrate the fan charts with reference to forecast errors
over the previous ten years, and then to discuss whether current circumstances call for
adjustments. Some recent examples of adjustments for which the MPC has given reasons are
visible in Figure 4. First, in the top left panel, the increase in November 2001 was attributed
to “volatile short-term movements in inflation from month to month”, despite the fact that the
fan charts relate to average inflation over the three months of each quarter; it has no
counterpart in the lower panels. Second, in February 2003, all panels of the figure show an
increase that reflected the uncertainties associated with the war in Iraq, while it was judged
that the exceptional uncertainty had dispersed by August 2003. Finally, the exceptional step
in the top right panel of Figure 3 in August 2005, increasing the standard deviation of the
GDP forecast at $h = 1$ to almost the same value as that at $h = 5$, followed recognition of the
uncertainty attached to recently-published national accounts data and the impact of data
revisions on forecast errors. Perhaps the external forecasters were already aware of this
effect; overall, none of the above adjustments are reflected in the SEF uncertainty measures.
Noticeable more generally at the longer horizons is the tendency for the SEF plots to remain
constant, after the initial reductions, while the MPC plots gradually decline, as the calibration
window moves away from the earlier, more turbulent experience. The correlations between
the two series at $h = 9$ are 0.70 for the inflation forecasts and 0.49 for GDP growth.

Independent evaluations of the MPC’s inflation density forecasts at one and five-
quarter horizons against inflation outcomes up to the end of 2003 find that forecast
uncertainty was overestimated (Clements, 2004; Wallis, 2004). The MPC had correctly
reduced the standard deviation of the fan chart at $h = 5$ from the high levels set initially, but
had underestimated the required reduction, while the increase in November 2001 at $h = 1$ was
a step in the wrong direction. The SEF uncertainty measures shown in Figure 4 suggest that
the same conclusion will emerge more strongly when these forecasts face the same tests.
4.2 The US Survey of Professional Forecasters

An international comparison can be based on the two existing survey datasets that include density forecasts, namely the SEF and the SPF. For this purpose we use the measures of uncertainty and disagreement for the SPF data reported by Giordani and Soderlind (2003, Fig.5), kindly updated by the second-named author. In particular, the uncertainty measure is the standard deviation of the aggregate histogram, based on a normal approximation, and the disagreement measure is the robust, quasi-standard deviation defined in Section 3.1 above. The SPF questions have an end-year focus, corresponding to questions 1 and 2 of the SEF. Like many other SPF researchers, Giordani and Soderlind concentrate on the first-quarter surveys, hence results are available at horizons $h = 4$ and $h = 8$. Comparisons of the two surveys, based on the same measures for inflation and growth at these two horizons, are shown in the four panels of Figure 5.

The results for inflation at both horizons show that, after the initial credibility-gaining period, UK uncertainty and disagreement are well below the US levels. They provide further evidence that “inflation targeting has helped to confer tangible benefits,” as argued by the Governor of the Bank of England in his 2005 Mais lecture (King, 2005). He cites the falling volatility of expected UK nominal short rates over this period, compared to its relative stability in the US, as evidence that inflation expectations are better anchored under inflation targeting, and our evidence makes the same point more directly. The US inflation uncertainty series are nevertheless remarkably stable, possibly reflecting the credibility of the Greenspan policy regime. We expect to see little evidence of short-lived turbulent episodes in an annual series, unless they coincide with the timing of observations. The most dramatic such episode in this period was the financial crisis of August-September 1998 – the Russian bond default and the near-collapse of Long-Term Capital Management (LTCM), a US hedge fund – and the Federal Reserve’s aggressive policy response appears to have stabilised short-term inflation expectations by the turn of the year. The local peaks in GDP growth uncertainty in both countries in 1999Q1, which are more prominent in the forecasts for 2000, are harder to associate with the financial crisis: in three of the four cases shown there is also a local peak in disagreement, which is a component of this measure of aggregate uncertainty.

The difference in the two countries’ monetary policy regimes appears to have no effect on output forecasts, since their GDP growth uncertainty measures are broadly similar. At the longer horizon, the disagreement measure is again rather smaller in the UK, although
at the shorter horizon the picture is mixed. Here there is a prominent peak in disagreement among US forecasters in early 2002, accompanied by a less prominent peak in uncertainty. These indicate a wide dispersion of views, and increased uncertainty, about the short-term effects on the real economy of the terrorist attacks of 11 September 2001.

4.3 Option-implied probability density functions

A source of direct measures of uncertainty about future asset prices is the prices of contracts traded in options markets. Given a sufficiently active market, option-implied probability density functions – in effect, density forecasts – can be derived under an assumption of risk-neutrality, and measures of their dispersion reflect the uncertainty of option market participants about future equity index and interest rate outcomes. These are estimated on a daily basis by the Monetary Instruments and Markets Division of the Bank of England, using methods described by Clews, Panigirtzoglou and Proudman (2000). The published standard deviations of the implied probability density functions provide a general indicator of financial market uncertainty. The standard deviation of the implied pdf of the short-term interest rate, six months ahead, 1997-99, is used for illustrative purposes by Clews et al. (2000, Chart 9).

For comparative purposes the corresponding series for the short-term interest rate, three, six and twelve months ahead, are plotted in Figure 6, together with the “question 1” RMSV measure of collective inflation uncertainty taken from the top left panel of Figure 1. The options-based series is noisy, day-by-day, and the inflation uncertainty measure exhibits the seasonal pattern noted above as the forecast horizon varies between one and four quarters, nevertheless some broad conclusions are clear. After the initial credibility-gaining period, the general level of SEF inflation uncertainty is closest to that of the six-month horizon interest rate uncertainty, but the lead-time effect is much more pronounced in the interest rate uncertainty series. The three-month-ahead series generally lies below the inflation uncertainty series, including its troughs which correspond to the comparable one-quarter horizon. And the twelve-month interest rate series is well above the inflation uncertainty series, including its peaks which mostly correspond to the four-quarter horizon. The amplitude of fluctuations in the individual interest rate series is also much greater, and their most important movements are not reflected in the inflation uncertainty series, nor vice versa. Clews et al. (2000) draw attention to the large rise in financial market uncertainty during the period of turbulence in late summer and autumn 1998, but this has little effect on forecasters’ inflation uncertainty, as noted for the annual series in the preceding section. Similarly, the
terrorist attacks of 11 September 2001 appear to have increased financial market uncertainty but not inflation uncertainty. The fall in inflation uncertainty associated with the change of monetary policy regime in 1997 gathered pace during 1998, whereas the interest rate options-based series suggest that the financial markets saw little impact on interest rate uncertainty of the change in arrangements for setting the official rate.

5. CONCLUSION

Access to the individual responses to the Bank of England Survey of External Forecasters allows the construction of direct measures of forecasters’ uncertainty and disagreement, since for the last ten years the survey has included questions about the probabilities of possible future outcomes as well as questions about their expected values. In this respect the survey matches the long-running Survey of Professional Forecasters in the United States, and its availability offers the prospect of a range of studies to extend a literature which has hitherto been entirely based on the SPF data. This paper makes a start on that agenda, by constructing relevant measures and studying their behaviour, in relation to the macroeconomic environment, to the SPF, and to other domestic measures.

An aggregate measure of uncertainty defined as a measure of dispersion of an aggregate probability distribution includes the effect of disagreement among point forecasts, hence the published aggregate density forecasts are inadequate indicators of uncertainty, unless an allowance for disagreement is made. A preferable measure of collective uncertainty can be constructed from the individual density forecasts as an aggregation of individual measures of dispersion. Sometimes the competing measures of collective uncertainty may be close to one another, because disagreement is relatively small, although our study, limited to two variables, offers little guidance on the more general circumstances in which this might apply. With respect to a third variable, options-implied interest rate uncertainty measures are seen to exhibit rather different behaviour.

The main economic event of the last ten years reflected in these measures is the new Labour government’s adoption in 1997 of a monetary policy of inflation targeting by a central bank granted operational independence for the purpose. This was followed by a
reduction in our measures of inflation uncertainty, to a level which is lower than that in the
United States and which has been remarkably undisturbed since the turn of the century.
Movements in financial market indicators have no counterparts in inflation forecast
uncertainty, whose lack of movement emphasises the secure anchoring of inflation
expectations by a credible monetary policy.

REFERENCES


Carruth, A., Dickerson, A. and Henley, A. (2000). What do we know about investment under


Clews, R., Panigirtzoglou, N. and Proudman, J. (2000). Recent developments in extracting


Review*, 47, 1037-1059.

distributions? Implications for the equity premium puzzle. *Journal of Economic
Dynamics and Control*, in press.

Economics*, 114, 185-227.


Figure 1. Measures of uncertainty and disagreement
Figure 2. Fixed-event forecast measures

Inflation uncertainty

Horizon

GDP growth uncertainty

Horizon

Inflation disagreement

Horizon

GDP growth disagreement

Horizon


Figure 3. Disagreement about inflation uncertainty

Individual histograms two years ahead, November 2005

RMSV ±2 standard errors: Question 3
Figure 4. Uncertainty, MPC and SEF, 1996Q4-2005Q4

Inflation uncertainty: $h=1$

GDP growth uncertainty: $h=1$

Inflation uncertainty: $h=5$

GDP growth uncertainty: $h=5$

Inflation uncertainty: $h=9$

GDP growth uncertainty: $h=9$
Figure 5. Measures of uncertainty and disagreement, US and UK

Inflation: last quarter current year

GDP growth: last quarter current year

Inflation: last quarter following year

GDP growth: last quarter following year
Figure 6. SEF inflation uncertainty and option-implied interest rate uncertainty