HERE IS THE NEWS: FORECAST REVISIONS IN THE BANK OF ENGLAND SURVEY OF EXTERNAL FORECASTERS

Gianna Boero, Jeremy Smith and Kenneth F. Wallis*

This article analyses the forecasts of inflation and GDP growth supplied by the individual respondents to the Bank of England's quarterly Survey of External Forecasters, 1996–2005, using a recently released dataset. This comprises a conventional incomplete panel dataset, with an additional dimension arising from the collection of repeated forecasts for the last quarter of each year. This fixed-event forecast structure allows study of the forecast revision process, its weak and strong efficiency, and its relation to macroeconomic news. The collection of density forecasts as well as point forecasts allows study of the revision process of forecast uncertainty.

Keywords: Forecast surveys; point forecasts; density forecasts; fixed-event forecasts; forecast revisions

JEL Classifications: C42; C53; E37

1. Introduction

Forecast early and forecast often might be a motto for many macroeconomic forecasters who publish a sequence of forecasts for each one of a series of future outcomes. The point and density forecasts of inflation and GDP growth published in each quarterly issue of the *National Institute Economic Review*, for example, are focused on the fourth-quarter or end-year outcome in the current year and the following year. The forecasts are revised in response to new information from one *Review* to the next, and eventually form a sequence of eight successive forecasts before the respective outcome is known. For a given outcome, these are referred to as a sequence of 'fixed-event' forecasts.

The quarterly Survey of External Forecasters (SEF) conducted by the Bank of England since 1996 has included questions with the same current-year and following-year fourth-quarter focus, until the questionnaire was redesigned in May 2006. The Bank has recently agreed to make available the individual survey responses, suitably anonymised, for research purposes, and this article studies the revision process in these forecasts. A particular feature of the survey is its inclusion of questions about the probabilities of possible future outcomes as well as questions about their expected values – density forecasts as well as point forecasts. In this respect the survey matches the well-established Survey of Professional Forecasters in the United States, and its availability offers the prospect of a

range of studies to extend a literature which is at present entirely based on the SPF data. Our previous work on this dataset derives measures of uncertainty and disagreement (Boero, Smith and Wallis, 2007a), and analyses the performance of the forecasts of inflation and GDP growth (2007b).

Concepts of forecast efficiency for fixed-event forecasts were introduced by Nordhaus (1987). Observing that recent forecasts will provide a good explanation of the current forecast of the same future outcome, attention focuses on the change or revision in the current forecast. A sequence of fixed-event forecasts is said to be weakly efficient if forecast revisions are independent of one another. This is relatively easy to test in principle, although in practice the available sequence of revisions may be rather short; Clements (1997) shows how pooling different sequences can overcome this difficulty. Forecasts are revised in response to the latest news, but it is only the surprise component of the news that is relevant, because a previous forecast has already incorporated the predictable component of the news, if it is efficient. This leads to the concept of strong efficiency, that forecast revisions are independent of the information used to construct the previous forecast. This is harder to test, since the forecaster's complete information set is not known to the researcher. Typically a few candidate variables are tested to see whether they can explain the revision: if they can, the rejection of

strong efficiency is conclusive, but if not, rejection may simply await some other series being tried.

The remainder of this article is organised as follows. Section 2 provides general information about the survey and the dataset of individual responses. Section 3 presents the statistical framework used for analysis of the forecast revision process and its properties. The data are described as a three-dimensional panel, following Davies and Lahiri (1995), because in addition to the usual two dimensions of panel data – individuals and time periods – for relevant time periods there are multiple forecasts, which are the subject of the present article. Section 4 contains the results of our empirical analysis, and Section 5 concludes.

2. The Bank of England Survey of External Forecasters

Every quarter for the past decade or so, the Bank of England has asked a group of external forecasters for their views on some key macroeconomic indicators. The survey provides useful information on expectations outside the Bank about future economic developments and the likely achievement of the Bank's inflation target. The institutions covered in the survey 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. The survey is carried out just before each quarterly *Inflation Report*, which then presents a summary of the aggregate results, usually 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 RPIX inflation, seeking their central projections or point forecasts together with the probabilities they attach to various possible inflation outcomes – a density forecast in the form of a histogram. 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. In the present article we study the point and density forecasts of inflation and GDP growth.

The dataset of individual SEF responses made available by the Bank and used in this article and its predecessors begins with the May 1996 survey and continues to November 2005. Each respondent has an identification number, so that their individual responses can be tracked over time and their point and density forecasts can be matched. The total number of respondents appearing in the dataset is 48, one of whom stayed for only two years, while only one is ever-present. To avoid complications caused by long gaps in the data, and to maintain degrees of freedom at a reasonable level, most of our analyses that refer to individual forecasters are conducted on a subsample of nineteen 'regular respondents'. These are respondents who each provided more than 70 per cent of the total possible responses to the inflation and GDP growth questions over time.

Each quarterly survey from February 1998 to February 2006 asks for forecasts of inflation and GDP growth 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. The inflation section of a recent questionnaire is shown in figure 1. (The first seven surveys in the dataset, which relate only to inflation, contained only the first two questions.) This structure eventually delivers nine successive forecasts of a given Q4 outcome, which we treat as a sequence of 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. In principle, eleven sequences of fixed-event forecasts of inflation, targeted on the Q4 outcomes of 1996–2006, are available in our sample. With a starting date of May 1996, there are three forecasts of 1996Q4 inflation, seven of 1997Q4 and eight of Q4 of the next two years. The two-year-ahead question was introduced in 1998, so there are then nine forecasts for Q4 of 2000–5 and finally five of 2006Q4, the present sample ending with the November 2005 survey. In practice the total sample size of 85 is available for the survey average and the single ever-present individual noted above, but all other individual forecasters' sequences have some gaps. For GDP growth there are nine sequences, with a total sample size of 71 forecasts. In May 2006 all three questions were switched to a constant-horizon format, focusing on the corresponding

---

Figure 1. Bank of England questionnaire, May 2005 survey, inflation question

YOUR COMPANY NAME (please complete): _____

PROBABILITY DISTRIBUTION OF 12-MONTH CPI INFLATION OVER THE MEDIUM TERM

Please indicate the percentage probabilities you would attach to the various possible outcomes in 2005 Q4, 2006 Q4 and 2007 Q2. The probabilities of these alternative forecasts should of course add up to 100, as indicated.

PROBABILITY OF 12-MONTH CPI INFLATION FALLING IN THE FOLLOWING RANGES			
	2005 Q4	2006 Q4	2007 Q2
<1.0%			
1.0% to 1.5%			
1.5% to 2.0%			
2.0% to 2.5%			
2.5% to 3.0%			
$\geq$ 3.0%			
TOTAL	100%	100%	100%

CENTRAL PROJECTION FOR 12-MONTH CPI INFLATION		
2005 Q4	2006 Q4	2007 Q2

quarter one, two and three years ahead, so fixed-event forecast analysis of the SEF has come to a halt.

3. The statistical framework

We adopt the notational convention of the three-dimensional panel introduced by Davies and Lahiri (1995) and denote as F_{ith} the h-step-ahead forecast of the outcome at time t made by individual i. The data are sorted first by individual, $i = 1,...,N$, with $N = 48$, then by target period t, from 1996Q4 to 2006Q4 (fourth quarters only, for the present purpose), and finally by horizon, $h = 9,8,...,1$ (except in the initial and final years, as noted above). With these definitions and the time series structure described in the preceding paragraph, we are in effect dating the information set on which F_{ith} is based at $t–h$, and the forecast is elicited by the survey carried out in quarter $t–h+1$. A practical detail is that the data, having been supplied sorted by survey date, have to be re-sorted by forecast target date.

Studies of forecast performance usually begin by noting that, in principle, forecasts should be evaluated in a specific decision context, in terms of the gains and losses that resulted from using the forecasts to solve a sequence of decision problems. In practice, however, macroeconomic forecasts are published for general use, and researchers have little knowledge of users'

particular decision contexts, so their evaluations are based on the statistical performance of the forecasts, using the standard squared error loss function. This underpins Nordhaus's concepts of forecast efficiency discussed above, for example, and our analyses presented below. Denoting the actual outcome for the variable of interest as A_t, this loss function leads to the definition of the rational expectation or efficient forecast as the expected value of A_t conditional on information available at time $t–h$, and the null hypothesis of the efficiency of the forecast F is then written $F_{ith}=E(A_t|I_{t–h})$. For notational convenience we suppress continual reference to the information set $I_{t–h}$, without intending to understate its importance.

Forecast errors

Davies and Lahiri (1995) decompose the forecast error $A–F$, which can be calculated once the outcome is known, as

$$A_t – F_{ith} = \phi_i + \lambda_{th} + \varepsilon_{ith}. \tag{1}$$

The second component λ_{th} is common to all individual forecast errors – it has no i subscript – and it represents the cumulative effect on A_t of uncorrelated period-by-period aggregate shocks since the forecast was made. With a slight change to Davies and Lahiri's notation we

write this as

$$\lambda_{th} = \sum_{j=0}^{h-1} \theta_j u_{t-j}, \quad h = 1,...,9. \tag{2}$$

Maintaining their assumption of the homoskedasticity of u, the θ-coefficients represent the dependence of the contribution of current economic surprises to the future outcome and hence to the error in previous forecasts of it on the length of time between the surprise and the outcome. The first and third components of the forecast error are specific to individual forecasters, separating a possible systematic effect or individual bias ϕ_i from an idiosyncratic non-autocorrelated error, ε_{ith}. This might reflect individual sentiment or 'animal spirits', and mishandling of or inadequacies in the initial information set.

Point forecast revisions

The availability of a sequence of nine forecasts of the same fourth-quarter outcome for the years 2000–5 (and shorter sequences for Q4 of other years) allows us to study the efficiency of the forecast revision process, and to extract estimates of the period-by-period aggregate shocks. Taking first differences with respect to h of equations (1) and (2) gives the forecast revision as

$$R_{ith} = F_{ith} - F_{i,t,h+1} = \theta_h u_{t-h} - \varepsilon_{ith} + \varepsilon_{i,t,h+1}, \quad h = 1,...,8. \tag{3}$$

It is the change in the forecast of A_t between times $t–h–1$ and $t–h$, with $\theta_h u_{t-h}$ representing the contribution of the 'news' component of the information available at time $t–h$. Note that the presence of the idiosyncratic error implies that the revision (3) has a moving average structure. The first-order autocorrelation coefficient is negative, with its absolute magnitude depending on the relative variances of the $\theta_h u_{t-h}$ and ε_{ith} terms. The relevant test of weak efficiency is whether the revision series, each of maximum possible length eight, exhibit autocorrelation of order greater than one. This represents a departure from Nordhaus's concept of weak efficiency, which in this framework may be more relevant to survey average forecasts, in which the influence of individual idiosyncratic errors is expected to be much reduced.

The strong efficiency condition is that the revisions are uncorrelated with variables known at the time of the earlier forecast, and it can be tested by regressing the revision on candidate variables from this information set, as noted above. The weak efficiency tests are also set up in a regression framework, which facilitates the pooling of the different fixed-event sequences and the calculation of appropriate HAC (heteroskedasticity and autocorrelation consistent) coefficient standard errors. An advantage of testing efficiency in the context of forecast revisions rather than forecast errors is that data on actual outcomes are not needed, which sidesteps the common problem of data revisions and the question of whether to use preliminary, 'real-time' or final, 'historical' outcome data to calculate GDP growth forecast errors. This problem does not arise with inflation, which is never revised after first publication.

Averaging forecast revisions over the $N_t(< N)$ responding individuals tends to remove the influence of idiosyncratic errors, as noted above, and hence provides a sequence of estimates of the perceived impact of the aggregate quarterly shocks:

$$\hat{\theta}_h \hat{u}_{t-h} = \frac{1}{N_t} \sum_{i=1}^{N_t} (F_{ith} - F_{i,t,h+1}), \quad h = 1,...,8. \tag{4}$$

These are available for values of the time index t that correspond to the fourth quarter of each year, hence successive sequences overlap, giving two estimates of the impact of current news: one is based on the resulting revisions in the forecasts for the fourth quarter of the current year, the other on the revisions in the forecasts for the fourth quarter of the following year. Rather than undertake further averaging we keep these separate, expecting revisions in response to a given shock to be smaller at longer horizons. More generally, in time series of the contribution of news extracted from the SEF forecasts, periodic heteroskedasticity might be expected.

Density forecast revisions

There is no generally accepted framework for analysing revisions to density forecasts, although it is a subject of much current research. Meanwhile we restrict attention to measures of location and scale, rather than the complete probability distribution. A measure of location of the distribution, such as the mean, is an implicit point forecast, and the properties of its revision process can be studied using the methods discussed above. The National Institute's density forecasts published in this *Review* have the form of a normal distribution, centred on the point forecast (with variance calibrated with reference to past forecast errors), so in this case there is no difference between the implicit and explicit point

forecasts. In the survey dataset, however, differences between point forecasts and density forecast means are observed, as documented and discussed in our previous work (2007a,b), hence it is of interest to replicate analysis of the point forecast revision process on the density forecast means. As in those studies, given the forecast histograms, we use 'crude' estimates of means and variances obtained by applying standard formulae, assuming that the reported probabilities are concentrated at the mid-points of the respective intervals, and replacing the first and last open-ended intervals by closed intervals whose width is twice that of the interior intervals. In the present article this last assumption is varied in one instance, discussed below.

Measures of scale or dispersion of the density forecast are summary measures of forecast uncertainty, and again it is of interest to study their revision process. For an optimal forecast in a stationary environment the forecast error variance increases monotonically with forecast horizon (with perhaps small departures at near horizons due to parameter estimation error). This is the source of the shape of the Bank of England's fan chart, and we expect to see the same shape in reverse, with the uncertainty of a fixed-event forecast reducing as the event draws nigh. Moreover it is generally convex, with uncertainty reducing more rapidly the closer the event. In an unstable world, however, with structural breaks and policy changes, none of this could be taken for granted.

4. Empirical results

We discuss the pooled individual results from the panel data structure described in the previous section together with, where possible, separate individual results for the nineteen regular respondents described in Section 2. We also include separate results for the survey average forecasts as published in the *Inflation Report*, these representing the 'public face' of the Bank's survey.

4.1 Weak efficiency

The results of the autocorrelation tests on the inflation point forecast revisions are very much in line with the statistical model summarised in equation (3). Pooling all respondents gives a highly significant negative first-order autocorrelation coefficient and a second-order coefficient that is not significantly different from zero at the 5 per cent level. The survey average forecast revisions have a smaller, and insignificant, negative first-order autocorrelation coefficient, reflecting the

averaging of the idiosyncratic errors, and again an insignificant second-order effect. Missing data have a larger impact on the available sample size for autocorrelation tests, and maintaining the same cut-off of 70 per cent of the total possible regression sample leaves only eight of the regular respondents who can be tested individually. Their results are all in accordance with the model: there are two cases of a first-order coefficient that is significant at the 5 per cent level, both negative, and no significant second-order coefficients. The results for the revisions to the implicit point forecasts, namely the density forecast means, are almost identical – the same pattern of signs and (in)significance of autocorrelation coefficients – the closest to a rejection of weak efficiency being one regular respondent whose second-order autocorrelation coefficient is significantly different from zero at the 10 per cent level. Departures from zero first-order autocorrelation towards negative coefficients in a few cases indicate the importance of the inclusion of the idiosyncratic errors in the Davies-Lahiri model, also the existence of individual heterogeneity in their magnitude.

The results for GDP growth forecast revisions have the same general form, although we observe greater individual heterogeneity in the details, as in our previous article (Boero, Smith and Wallis, 2007b). In the pooled, average and individual results for point forecast revisions and density forecast mean revisions any significant first-order autocorrelation is negative, except for a significant positive first-order autocorrelation coefficient in the survey average density forecast mean revisions, which is at variance with the model. The weak efficiency test, on the second-order autocorrelation coefficient, is now failed in one case, that of an individual respondent, on the point forecast revisions – the same respondent who came close to rejection on the inflation density forecast means, a result that they repeat for the GDP growth density forecast means.

4.2 Strong efficiency

Turning to tests of strong efficiency, the proposition is that, if the news at time $t–h–1$ is efficiently incorporated into the forecast $F_{i,t,h+1}$, then it should be uncorrelated with the revision R_{ith}, which on the other hand is a response to the new information at time $t–h$. We consider a simple 'own-variable' information context, defining the news as the most recently observed change in the target variable, paying attention to the National Statistics release dates of inflation and national accounts data, in relation to the survey schedule. Thus the test

variable is the latest change in inflation or GDP growth available at time $t–h–1$, and we consider its possible relation with R_{ith} in two regression settings. The first is a simple regression of the revision on the earlier news; the second augments this regression by adding the later news which, by hypothesis, is correlated with the current revision. This allows for possible autocorrelation in the test variable, and also validates its selection for an appropriate test. The augmented regression also includes the error in the Bank of England's forecast of the target variable observed at time $t–h$ and, for inflation, a dummy variable for the switch to the CPI in the February 2004 survey, which saw a substantial downward revision to the November 2003 forecasts, based on RPIX, for all but one of our regular respondents.

The aggregate results for the point forecasts of inflation, considering both the survey average as published in the *Inflation Report* and the pooled individual responses, do not reject the strong efficiency hypothesis. In neither the simple regressions nor the augmented regressions is the coefficient on the test variable significantly different from zero at the 5 per cent level. In the augmented regressions both proxies for the news in the most recent inflation data are highly significant, validating the choice of test variable. At the individual level, the strong efficiency hypothesis is not rejected in either regression for seventeen of the nineteen regular respondents. In the two remaining cases, one has significant coefficients on the test variable at the 5 per cent level in both the simple and augmented regressions, while the significance drops from the 5 per cent to the 10 per cent level for the simple regression in the other case.

For the inflation density forecast means, the results for the survey average forecast are the same as above. The pooled individual forecast results, however, show a significant rejection of strong efficiency in the simple regression test, albeit not in the augmented regression. To study this at the individual level we first consider the corresponding pooled regression results for the nineteen regular respondents. This reduces the sample size of the all-respondents pooled regression by one third, but the coefficients and their standard errors are virtually unchanged. Turning next to the nineteen individual regressions, we find that there are four respondents with rather large negative coefficients in the simple regression test which appear, arithmetically, to be mainly responsible for the significant negative coefficient in the pooled regression. Given variations in their standard errors, one of these coefficients is significantly different from zero at the 1 per cent level,

and two at the 10 per cent level, while the fourth is not significant. This leads us to de-emphasise the rejection of strong efficiency in the pooled simple regression test. In the individual augmented regression tests there is again one rejection, at the 5 per cent level, by the same respondent who fails both strong efficiency tests on the point forecasts. On the other hand the respondent who fails the simple regression test on the density forecast means passes the strong efficiency tests on the point forecasts, suggesting that the implied point forecasts we have calculated from this respondent's density forecasts, namely their means, may not be the appropriate choice for this individual.

Our tests of the strong efficiency of the GDP growth forecast revisions are based on real-time data, and sidestep the question of whether this, or the final revised growth number, is the forecaster's target. For the point forecasts, the simple regression tests of the survey average and the pooled individual forecast revisions do not reject the null hypothesis of strong efficiency, while at the individual level, there are two rejections at the 5 per cent level among the nineteen regular respondents. In the augmented regression tests, the survey average forecast again passes the test, but in the pooled individual regression the coefficient on the test variable is significantly different from zero at the 5 per cent level. To pin down this result we follow the same strategy as in the previous paragraph. Excluding all but the regular respondents from the pooled individual regression replicates the rejection of strong efficiency, the coefficient on the test variable increasing its magnitude. The individual regression results suggest that this is due to the same two individual respondents, who have large and significant coefficients on the test variable, while the null hypothesis is not rejected for the remaining seventeen regular respondents. In one of the two rejections the exceptionally large coefficient on the test variable is almost matched by a similar coefficient of opposite sign on the change in GDP that forms part of the current information set, suggesting that our interpretation of information flows may not match this forecaster's use of the data. As an alternative test variable we consider the change in the index of industrial production, which is available with a much shorter publication lag, and is often used as a leading indicator of GDP growth. This does indeed result in non-rejection of strong efficiency for the forecaster in question, although four other individual respondents now have significant coefficients on the test variable.

For the GDP growth density forecast mean revisions, the

survey average and pooled individual forecasts pass both the simple and augmented regression tests of strong efficiency. Among the regularly responding individuals, however, the two cases of rejection in the previous paragraph are joined by two more in the augmented regression tests, and by a fifth in the simple regression test. In this case the use of the alternative test variable based on the industrial production index leads to fewer rejections of strong efficiency, although we recall our opening observation that non-rejection of strong efficiency is not final, but may merely await the discovery of the variable that was used inefficiently in constructing the original forecast.

4.3 Estimates of the quarterly aggregate shocks

Time series of estimates of the impact of aggregate shocks are constructed as in equation (4), equivalently as revisions to the survey average point forecast. As noted above, the structure of the SEF questionnaire implies that current news impacts on the revision of two forecasts, one for the fourth quarter of a given year (question 1) and the other for the fourth quarter of the following year (question 2). These are plotted separately in each figure, for the inflation forecasts in figure 2 and the GDP growth forecasts in figure 3. The labels on the horizontal axes give the date of the survey that collected the revised forecast, indicating the timing of the news that may have influenced the revision. In the notation of equation (4), successive points on the 'question 1' plots

relate to horizons $h = 4,3,2,1,4,3,2,1,...$, with the target date for the forecasts in each subsequence being Q4 of that year; on the 'question 2' plots the successive horizons are $h = 8,7,6,5,8,7,6,5,...$, with the target date for the forecasts in each subsequence being Q4 of the following year. The dataset begins in 1996Q2, so the first available revisions to inflation forecasts are dated 1996Q3 in figure 2. The GDP growth questions were introduced in 1998Q1, so figure 3 begins with forecast revisions dated 1998Q2.

The inflation shocks in figure 2 have an outlier that is very easy to interpret, caused by a change in the question that matched the change in the Bank of England's inflation target variable announced by the Government in December 2003, from RPIX inflation, with a target of 2.5 per cent, to CPI inflation, with a target of 2 per cent. The fact that the average deviation between the two measures over the previous fifteen years was three-quarters of a percentage point was widely reported. The 2004Q1 survey now asked for forecasts of CPI inflation, with the grid for the probability questions shifted down by half a percentage point. On average, the point forecasts for the current year were duly revised down by the reported historical average deviation.

The next largest revisions to the inflation survey average forecasts are the adjacent observations, of opposite sign, in the second and third quarters of 2001. We relate these directly to the news in the most recent inflation data

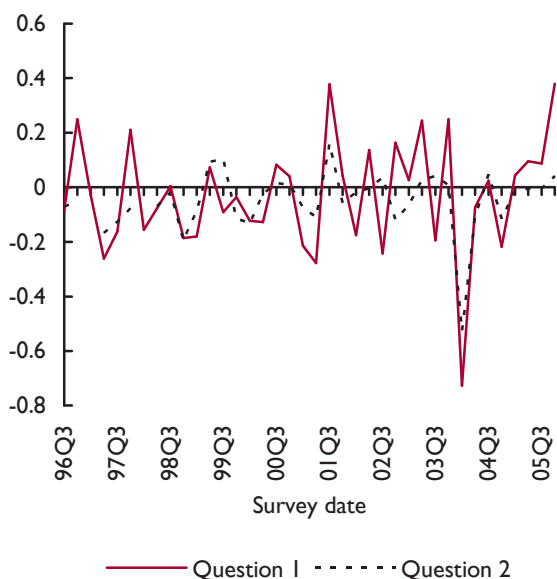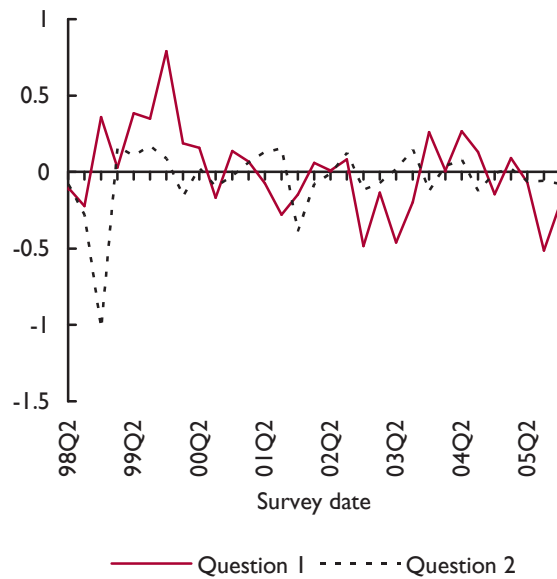Figure 2. Aggregate shocks to expected inflation

Figure 3. Aggregate shocks to expected GDP growth

available at the time, having in mind the high significance of this variable in the augmented regressions used to test strong efficiency. The first quarter of 2001 saw historically low RPIX inflation, at an annual rate of 1.9 per cent. This figure became available in mid-April, shortly before the second-quarter survey was carried out, and forecasts for 2001Q4 were duly revised down. Inflationary pressures were increasing rapidly, however, mainly through the food price components of the index, reflecting falls in the domestic supply of agricultural crops, following flooding, and of meat, following the outbreak of foot-and-mouth disease. The result was inflation of 2.4 per cent in May and June. Again the latter outcome was released, in mid-July, shortly before the next survey was carried out, which in turn saw large upward revisions to the forecasts for 2001Q4.

The extreme observations on GDP growth forecast revisions in figure 3 have a similar paired relationship, with a large downward revision being followed by upward revisions to the forecast for the same target date. 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 – was followed by well-publicised reductions in forecasts of world economic growth by such organisations as the International Monetary Fund and the World Trade Organisation. For the UK, Chancellor Gordon Brown reduced his growth forecast to 1 per cent in a speech at the World Bank in early October. The 1998Q4 SEF is seen to have revised forecasts of the annual rate of GDP growth for 1999Q4 downwards, on average by one percentage point. The November 1998 *Inflation Report* found this particularly noteworthy, observing (p.50) not only this reduction but that the new survey average forecast was for ¾ per cent growth. Perhaps influenced by this finding, the Monetary Policy Committee (MPC) reduced interest rates by 50 basis points in November, and repeated this action in December 1998; further cuts totalling 1.25 per cent by June 1999 took interest rates to 5 per cent, the lowest level for 22 years. Manufacturing confidence reached a two-year high in August, and growth rose due to strong performance by the service sector. Global forecasts were revised upwards, as were the domestic forecasts in the SEF, the upward revisions in the 1999Q2–Q4 surveys more than offsetting the extreme downward revision in 1998. In the event, the National Statistics preliminary estimate of the 1999Q4 outcome was 2.9 per cent, which at the time of writing has been revised to 3.5 per cent.

The impact of the news is greater the nearer is the

forecast target date, as reflected in calculations of the standard deviations of the annual series of plotted values at the same lead time. For inflation, deleting the RPIX/CPI outlier noted above, the standard deviations at $h = 1,...8$ are 0.199, 0.190, 0.179, 0.100, 0.077, 0.080, 0.080, 0.045, showing a declining tendency, as hypothesised, bearing in mind that the sample size for each of these statistics is between seven and ten. For GDP, the corresponding figures are 0.400, 0.273, 0.261, 0.098, 0.355, 0.168, 0.077, 0.107. In this case the maximum number of observations for each calculation is eight. The main departure from a monotonic decline is at $h = 5$, as a result of the overreaction to the global financial crisis evidenced in the 1998Q4 survey, discussed above.

4.4 Revisions to forecast uncertainty

Revisions to measures of the dispersion of the density forecasts are interpreted as revisions to forecast uncertainty, and we describe this revision process by calculating the standard deviations of the survey average density forecasts. The changes in standard deviation are then plotted for separate sequences of fixed-event forecasts in figures 4 (inflation) and 5 (GDP growth). The key identifies the target date for each sequence of up to eight forecast revisions. We limit attention to nine more-or-less complete sequences of inflation forecast revisions, excluding two revisions of forecasts of 1996Q4 outcomes and four of 2006Q4. The horizontal axes indicate the date of the survey that collected the 'new' forecast, as in figures 2 and 3 respectively. Figures 4 and 5 are derived from information presented in figure 2 of our previous article (2007a), with one change in the details. The set-up of the histogram bins in the survey questionnaire has changed during our sample period, on two occasions for the inflation question and one for GDP growth. Calculated standard deviations, more than means, are sensitive to changes in the width of the interior bins and the associated width of the open-ended bins assumed for computational purposes, and we find that there are fewer distortions on these occasions in the present figures if we assume that the open-ended bins are of the same width as the interior bins.

The revisions are generally expected to be negative, with uncertainty being reduced at an increasing rate as the forecast target date approaches, as discussed above. This negative downward trajectory emerges as a general shape, subject to short-term deviations, in figures 4 and 5. The overall amplitude has fallen over time, reflecting

Figure 4. Revisions to inflation uncertainty (aggregate standard deviation)
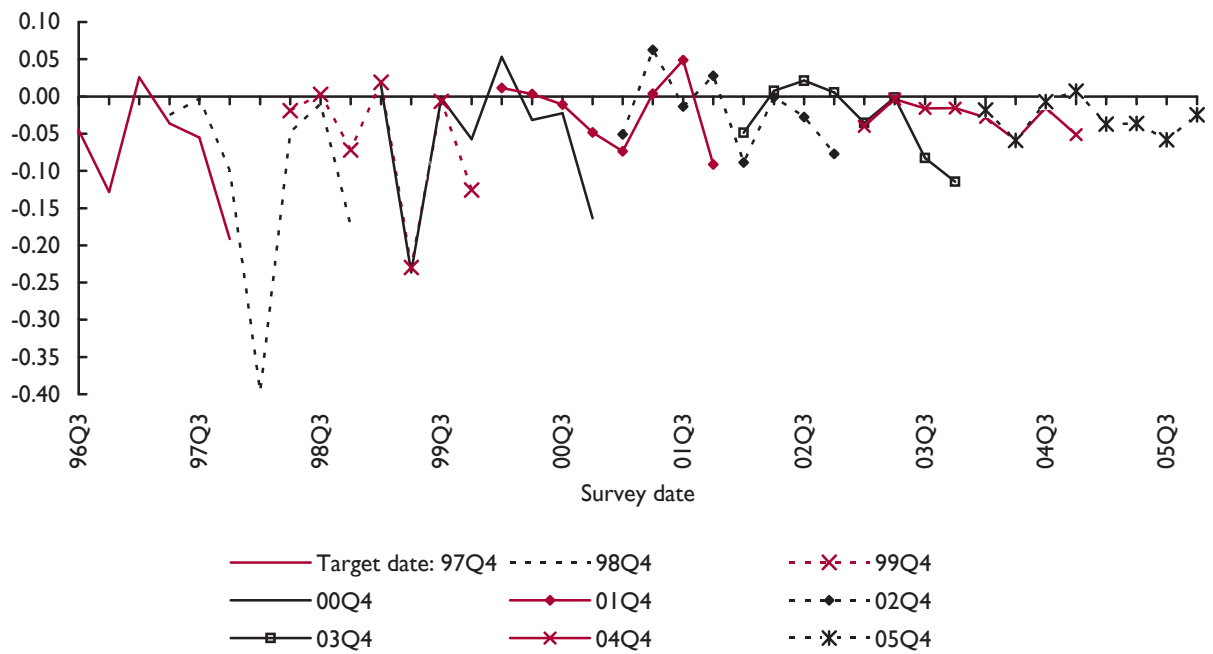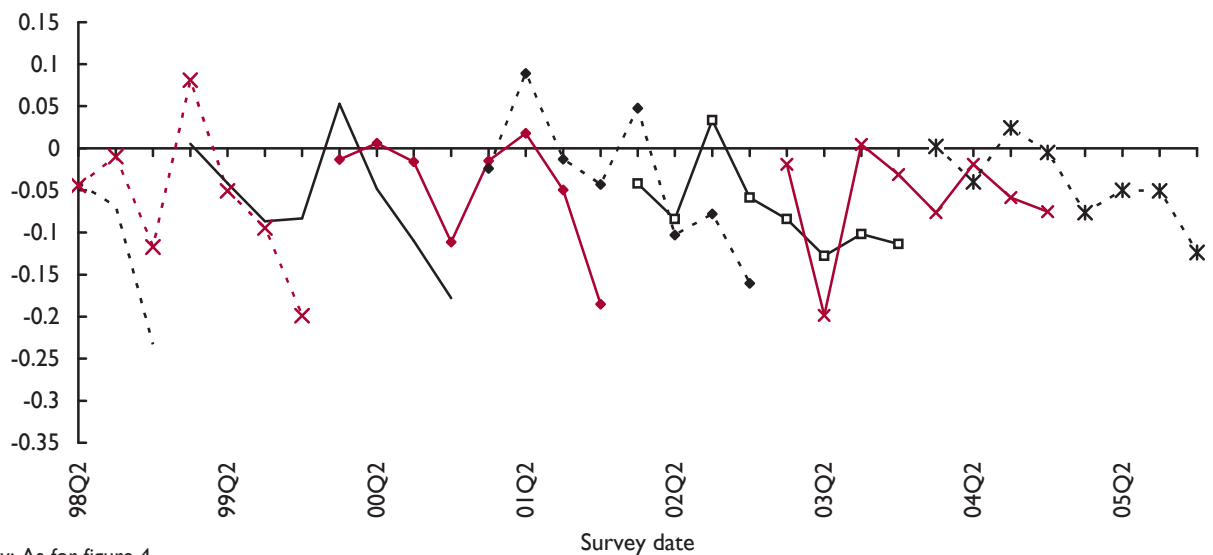

Figure 5. Revisions to GDP growth uncertainty (aggregate standard deviation)

Key: As for figure 4.

the general reduction and stabilisation of uncertainty over this period documented in our previous article (2007a). Positive revisions, implying increased uncertainty about the future outcome, or unusually large negative revisions within the sequence, represent departures from the general pattern, subject to sampling error, and some of these are worthy of comment.

First, the negative outliers within the sequence, in 2003Q2 for GDP growth, and 1998Q1 and to a lesser extent 1999Q2 for inflation, are possible distortions associated with changes in the configuration of the histogram bins noted above. Second, there are prominent positive outliers for GDP growth, three of the first four of which occur in the first quarter of the year,

likewise the positive inflation uncertainty revisions in 1997Q1 and 2000Q1. At these times the question about the target Q4 outcome shifts from being the second question on the questionnaire in the last survey of the previous year, to the first question, and this may have hindered clarity of response. On the other hand the increases in uncertainty in both variables in 1999Q1, relating to 1999Q4 outcomes, may have been influenced by the MPC's hyperactivity between the 1998Q4 and 1999Q1 surveys, in the circumstances discussed in the previous section.

Also discussed in the previous section are the rapid changes in inflation experienced in 2001. In November 2001 the MPC almost doubled the standard deviation of its inflation fan chart in the initial quarter of its forecast due to "volatile short-term movements in inflation from month to month", and the increases in inflation uncertainty in the SEF short-term forecasts in 2001 may have shared this cause. Finally the MPC also increased the standard deviation of the inflation fan chart in 2003Q1, to reflect the uncertainties associated with the war in Iraq, and although the Committee judged that the exceptional uncertainty had dispersed by 2003Q3, the survey forecasters' increases in inflation forecast uncertainty over this period may again have shared the same cause.

5. Conclusion

This article studies the revision process of the successive forecasts of inflation and GDP growth supplied by individual respondents to the Bank of England Survey of External Forecasters between May 1996 and November 2005. The analysis builds on our previous work on this recently released dataset and exploits two important features of the survey, namely its fixed-event forecast structure over this period and the availability of density forecasts as well as point forecasts. The former enables us to study the efficiency of the forecast revision process, and to extract estimates of period-by-period aggregate shocks. The latter allows a dual revision evaluation, of the reported point forecasts and the alternative point forecasts implied by a measure of location of the reported density forecasts. We also consider the revision process of forecast uncertainty, as represented by changes in the standard deviation of the density forecasts; this appears to have no counterpart in existing studies of similar forecast surveys.

Tests of weak and strong efficiency, based on pooled sequences of up to nine forecasts of the same fourth-quarter outcome, are conducted on the survey average, the pooled individual forecasts, and a subsample of regular respondents, examined individually. These indicate, overall, a good performance. Greater individual heterogeneity is observed in respect of forecasts of GDP growth rather than inflation. This is not surprising given the well-known problems of forecast construction and evaluation caused by the longer publication delay of national accounts data, and the subsequent extended revision process of the preliminary 'real-time' data. A general mistrust of the data by the forecaster is hard to incorporate into simple models of the forecast process, but it can be expected to mask any inference about that process that it might be possible to draw in its absence. This prompts several questions about the forecasters' objectives and methods, which their anonymity prevents us from exploring.

Estimates of the perceived impact of aggregate quarterly shocks extracted by averaging forecast revisions over the survey respondents show a clearly declining impact of the news the longer is the forecast horizon. This appears to be an important extension to the model used as a framework for analysis, which is otherwise well-supported by the data. Revisions to forecast uncertainty in general move in the opposite direction, with uncertainty being reduced at an increasing rate as the forecast target draws nigh. Some particular episodes can be well understood in relation to the economic history of the time. If there are departures from efficient incorporation of the latest news into the forecasts, they seem to be towards over-reaction to the news. This tendency also emerged to some extent in our evaluation (2007b) of these forecasts against the actual outcomes. Again, further exploration of the forecasters' objectives and methods would be worthwhile.

REFERENCES

Boero, G., Smith, J. and Wallis, K.F. (2007a), 'Uncertainty and disagreement in economic prediction: the Bank of England Survey of External Forecasters', *Economic Journal*, forthcoming.

—(2007b), 'Evaluating a three-dimensional panel of point forecasts: the Bank of England Survey of External Forecasters', presented at the Royal Economic Society Annual Conference, University of Warwick, April.

Clements, M.P. (1997), 'Evaluating the rationality of fixed-event forecasts', *Journal of Forecasting*, 16, pp. 225–239.

Davies, A. and Lahiri, K. (1995), 'A new framework for analyzing survey forecasts using three-dimensional panel data', *Journal of Econometrics*, 68, pp. 205–27.

Nordhaus, W.D. (1987), 'Forecasting efficiency: concepts and applications', *Review of Economics and Statistics*, 69, pp. 667–74.