

**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, revised October 2006

Abstract This article 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 significant, 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 for providing extended data from one of his articles, and to Roy Batchelor, Jagjit Chadha, Michael Clements, Kajal Lahiri, John Whitley, and this journal's editor and referees for helpful comments. Readers wishing to gain access to the data should write to the Publications Editor, Inflation Report and Bulletin Division, Bank of England, Threadneedle Street, London EC2R 8AH, UK.

*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 subjective probability distributions – so-called density forecasts. Our purpose in this article is to introduce a new source of such 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 asked 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 responses to the US survey, 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) and Rich and Tracy (2006); 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 article. 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 subjective distributions of individuals' prospects, 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 (2007) 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 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

some possible direct measures of uncertainty and disagreement, and the results of implementing our preferred measures 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 each quarterly *Inflation Report*, which then presents a summary of the 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.

From its first appearance in February 1993 the *Inflation Report* included information on the point forecasts of inflation made by outside forecasters, variously described as “other inflation projections”, “private sector inflation forecasts” and so forth. In February 1996, for the first time, the Bank itself asked the forecasters not only about their central projections for inflation but also about the probabilities they attach to various possible inflation outcomes. Inflation was defined with respect to the Retail Prices Index excluding mortgage interest payments (RPIX), in terms of which the official inflation target was defined; the survey definition switched to the Consumer Prices Index (CPI) from 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 article.

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 , $1.5-2$, $2-2.5$, $2.5-3$, $3-3.5$, >3.5), each change reducing the width of the interior intervals. From February 2004 the whole grid was shifted downwards by 0.5, following the switch from the RPIX to the CPI inflation measure. 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).

The dataset of individual SEF responses made available by the Bank covers 39 surveys, beginning with the May 1996 survey and continuing 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 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 the dataset, 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%. 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 adjacent surveys. These investigations uncovered two further inconsistencies which have also been corrected. A note detailing these amendments to the dataset is available 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 forecast densities and the distribution of point forecasts (see Wallis, 2005, for fuller discussion and references to precursors). Next, some practical considerations in calculating these measures are discussed. 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 $f_i(y)$, $i = 1, \dots, n$. 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

$$f_A(y) = \frac{1}{n} \sum_{i=1}^n f_i(y). \quad (1)$$

From a statistical point of view, this is an example of a finite mixture distribution. 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 σ_i^2 . Then the mean of the aggregate density is

$$\mu'_1 = \frac{1}{n} \sum_{i=1}^n \hat{y}_i = \hat{y}_A, \quad (2)$$

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

$$\mu'_2 = \frac{1}{n} \sum_{i=1}^n (\hat{y}_i^2 + \sigma_i^2).$$

Hence the variance of f_A is

$$\sigma_A^2 = \mu'_2 - \mu_1'^2 = \frac{1}{n} \sum_{i=1}^n \sigma_i^2 + \frac{1}{n} \sum_{i=1}^n (\hat{y}_i - \hat{y}_A)^2. \quad (3)$$

This expression decomposes the variance of the aggregate density, σ_A^2 , a possible measure of uncertainty, 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 variance of 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 GDP 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. Denoting the average individual variance, the first term on the right-hand side of equation (3), as $\overline{\sigma^2}$:

$$\overline{\sigma^2} = \frac{1}{n} \sum_{i=1}^n \sigma_i^2, \quad (4)$$

the *RMSV* is its square root,

$$RMSV = \sqrt{\overline{\sigma^2}}. \quad (5)$$

An implication of equation (3) is that, given the variance of the aggregate density and the disagreement measure, an implied *RMSV* can be calculated as the square root of their difference, without calculating the variance of each individual forecast density.

3.2 Some practical considerations

Estimating moments from histograms A first question is how to estimate means and variances from histograms, which in the SEF data have rather few bins, between four and six, 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. 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.

For a histogram with k intervals, denote the upper boundaries of the first $k-1$ intervals as b_j , $j=1, \dots, k-1$, and the cumulative probabilities at these boundaries as θ_j , $j=1, \dots, k-1$. 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,

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

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. However in the SEF data the aggregate distributions typically look quite normal, and the time series of variances estimated by the crude method and the normal approximation are virtually indistinguishable from one another, hence in practice we choose to retain the simpler method. The individual density forecasts differ considerably in their dispersion, and present a problem for the normal approximation when forecasters assign zero probability to the outer intervals, as often occurs in the dataset. At first sight this might appear to be an estimation problem, since there may be inadequate degrees of freedom in the above non-linear regression, but more fundamentally the assumption of a normal distribution is inconsistent with the specification of a finite range for the forecast density. When some forecasters further assign non-zero probability to only two intervals, the minimum observed in the dataset, individual variance calculations by any method become less reliable, and we avoid these difficulties by using the implied *RMSV* as our measure of average individual uncertainty in what follows, with the exception of Section 3.6 below.

Our assumption about the open-ended intervals is that they have twice the width of the interior intervals, following Batchelor and Dua (1996). Whereas the crude calculation of variances is generally insensitive to assumptions about the width of the open-ended bins, because their reported probabilities are small, on two occasions in the sample period, associated with a reduction in the number of bins, this condition does not apply. The more prominent concerns the reduction in the number of histogram bins for GDP growth from six to four in May 2003. As noted above, this was achieved by combining the outer, open-ended bins with their nearest neighbours, leaving only two interior bins. In the aggregate density for question 3, the last two bins in February 2003 (3-4, >4%) have a combined probability of 26%, while in May 2003 the single, newly combined upper bin (>3%) has a probability of 22%. At the individual level, 17 of the 22 reported densities in February 2003 have non-zero probabilities that growth will exceed 4%, with four responses of 15% or greater, while in

May 2003 none of the 23 reported densities give zero probability to the possibility that growth will exceed 3%, indeed there are four responses of 40% or greater. If a unit width is assumed for the upper interval, and weights of this order of magnitude are placed on mid-points of 3.5 and 4.5, separately, in February, and on 3.5 in total in May, then a spurious reduction in uncertainty is observed. Our preferred double-width assumption implies corresponding mid-points of 3.5 and 5.0 in February, and 4.0 in May, which results in series without an obvious break at this point, and in which we have greater confidence throughout.

Measures of disagreement The measure defined by the square root of the last term in equation (3) is a measure of disagreement among individual density forecast means, whereas the variable that is relevant to the question of the usefulness or otherwise of disagreement as a proxy for an unavailable direct measure of uncertainty is the reported point forecast. The notation used above assumes that these coincide, although we observe differences between point forecasts and density forecast means in the SEF dataset. The optimal point forecast under a quadratic loss function is the mean of the density forecast, but differs from the mean under other symmetric loss functions – linear, or bounded, for example – if the density is asymmetric. And if the loss function is asymmetric, again the optimal point forecast differs from the density forecast mean. For the US SPF data, 1992-2004, Engelberg, Manski and Williams (2008) document the extent to which reported point forecasts deviate from calculated measures of central tendency (mean/median/mode). They observe that in such cases forecasters are inclined to present favourable scenarios, in the sense that, more often than not, their point forecasts anticipate lower inflation and higher output growth than indicated by the central tendencies of their density forecasts. For the SEF data, some summary evidence is given in Table 1, which reports the percentage of cases in which the point forecast deviates from the density forecast mean by more than 0.2 percentage points. It is seen that there is a greater tendency towards “favourable” scenarios in respect of GDP growth than inflation, and that this increases as the forecast horizon increases. However in the two-year-ahead inflation forecasts the deviations are almost evenly balanced, suggesting that, in an inflation targeting regime, a favourable scenario is one in which the official target is achieved in the medium term, and this may lead to a positive or negative adjustment to an initial forecast. A fuller analysis of the point forecasts is given by Boero, Smith and Wallis (2006).

For our statistical measure of disagreement, using either variable, we follow Giordani and Soderlind (2003) and choose a measure that is robust to outliers, specifically the quasi-standard deviation (qsd), equal to half the difference between the 16th and 84th percentiles of the sample of point forecasts or density forecast means, appropriately interpolated. With normally distributed data this calculation delivers the standard deviation; in the SEF data it prevents the measure being overly sensitive to extreme observations, which may be mistakes, on two or three occasions. As an example we use the responses to question 1 of the survey, and plot the $qsds$ of the point forecasts and the density forecast means for inflation and GDP growth in the two panels of Figure 1. There appear to be no statistical reasons to prefer one series to the other, and in the remainder of this article we use the qsd of the point forecasts as our measure of disagreement, since these are the forecasts that are relevant to the proxy variable question, as discussed above.

3.3 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. The six panels of Figure 2 present, for each of the three questions and two variables in turn, the aggregate standard deviation, σ_A , the average individual uncertainty, implied *RMSV*, and the qsd measure of disagreement among point forecasts.

It is seen that the uncertainty measures obtained from the responses to the first question show strong annual periodicity, but this 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 relative amplitude of the periodic fluctuations reflects the behaviour of the mean square forecast error for a relevant time series model, such as a first-order autoregression, which falls much less between 8- and 5-quarter-ahead forecasts than between 4- and 1-quarter ahead forecasts. This pattern of decline is emphasised when we consider fixed-event forecasting schemes below. The disagreement series for question 1 show evidence of similar periodicity but with a smaller amplitude than the uncertainty series, and for inflation this appears only in the first part of the sample period.

The major change in the general level of the inflation uncertainty series that catches the eye 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. We test for a structural break in the series using Andrews' (1993) sup-Wald statistic, treating the break point as unknown, with approximate asymptotic p -values obtained as described by Hansen (1997). For the first two questions this shows that a significant break, at the 5% level, can be dated at any time between February 1998 and November 1999 for the aggregate standard deviation and average individual uncertainty (implied *RMSV*) series. The third question did not appear until February 1998, and there are insufficient observations to apply a formal test in the early part of the series, nevertheless the aggregate and average individual uncertainty series both show a prominent fall in May 1999. There may have been some initial learning about how to answer the probability questions on the part of the forecasters, but we attribute these reductions in uncertainty about future inflation to 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 committee is chaired by the Governor of the Bank of England and has eight other members, of whom four are senior Bank staff and four are externally appointed.) 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. Our measures of uncertainty provide further evidence of the rapid gains in monetary policy credibility over this period.

Application of the formal tests to the GDP growth uncertainty series shows no equivalent structural breaks, although the series have a tendency to decline towards the end of the sample period. In the first seven surveys the two inflation disagreement series are at a significantly higher level than subsequently, with a local peak in February 1997, but from February 1998, when the third inflation question and the GDP growth questions first appeared, none of the disagreement series shows any systematic variation.

3.4 Fixed-event forecasts

We use the fixed-event scheme to study the effect of forecast horizon on uncertainty and disagreement. Except for the first two years of the 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 2, and present it as separate trajectories for selected Q4 target outcomes in Figure 3. Of the nine available Q4 inflation outcomes, 1997-2005, and the seven available Q4 GDP growth outcomes, 1999-2005, in each case we omit three (2001-03) to avoid overloading the figure. The implied *RMSV* measure of uncertainty is shown in the upper panels, and the *qsd* of the reported point forecasts is shown in the lower panels.

Looking first at the inflation panels of Figure 3, we note that the earliest trajectory, for the 1997Q4 outcome, starts, with the 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 average individual uncertainty measures (upper left panel), 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 two recent trajectories (and the three not shown) 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 average individual uncertainty measure is reduced from approximately 0.7 percentage points at $h = 9$ to 0.4 at $h = 1$. In time-series forecasting, the reduction in forecast mean square error between long-horizon and short-horizon forecasts is greater the more persistent is the series, and the reduction observed here is greater than that implied by the persistence of inflation over this period. The degree of persistence of UK inflation has been falling in recent years, perhaps to a greater extent than forecasters have recognised.

Disagreement measures scarcely feature in forecasting textbooks, however, and there are some similarities and some differences between the two inflation panels. The first two disagreement trajectories, like their uncertainty counterparts, start high and then decline, although the later ones are rather flatter. Nevertheless in all nine cases there is more

disagreement at the beginning ($h = 9$ or slightly less) than at the end ($h = 1$), but 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.

3.5 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, $h = 1, \dots, 8$, 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. Over the available period, 1998Q1-2005Q4, the correlation between the *RMSV* uncertainty series and the *qsd* disagreement series shown in the bottom panels of Figure 2 is 0.03 for inflation and -0.11 for GDP growth. 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 shown in the top and middle panels of Figure 2. For inflation, over the full period 1996Q2-2005Q4, we then obtain much higher correlations: 0.63 for question 1 and 0.86 for question 2. It is clear from the figure that these results are driven by the initial observations in each series, covering the “pre-credibility” and “learning” periods, but also when the very early observations seem less reliable: from 1999Q2 the correlations become 0.02 and -0.08 respectively. For GDP growth, over the full period used above, we again obtain negative correlations, namely -0.28 for the adjusted

question 1 series and -0.18 likewise for question 2. These results show that, at least over this period, disagreement is not a good proxy for uncertainty.

In relation to the US literature, we note that earlier results more favourable to the proxy variable view were based on much more turbulent experience, going back as far as 1968. Our sample period is shorter and much more stable: RPIX inflation has varied between 1.9% and 3.2%, and GDP growth between 0.7% and 3.95% (quarterly observations on annual rates) over the ten years of the SEF dataset. At the same time, however, we note that recent research on the SPF data that brings the sample period up to date and studies the robustness of previous findings to the choice of measures finds less support than hitherto for the proposition that disagreement is a useful proxy for uncertainty (Lahiri and Liu, 2006; Rich and Tracy, 2006).

3.6 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 4, 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: for this purpose, though not for our computations, as noted above, these are drawn with 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. (Corresponding variation in the dispersion of the SPF individual forecast histograms is represented by Lahiri and Liu (2006) as a box-and-whisker plot of their variances.)

To summarise the impact of this sample variation on our collective uncertainty measure, $RMSV$, as defined in equation (5), we calculate its standard error under the conventional random sampling assumption. 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 obtained as

$$se(RMSV) = \sqrt{\frac{\sum(\sigma_i^2 - \overline{\sigma^2})^2}{4n(n-1)\overline{\sigma^2}}}$$

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 4, which refers to the quarterly constant-horizon series at $h = 9$. (This does not exactly correspond to the series shown in the lower left panel of Figure 2, since for the present purpose we need to calculate the variances of the individual forecast histograms, and hence on this occasion we report the direct *RMSV* as given explicitly by equation (5)). 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 GDP 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, implied *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 public attention. Results for inflation and GDP growth are shown in Figure 5.

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 5. 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 5 in August 2005, increasing the standard deviation of the GDP growth forecast at $h = 1$ to almost the same value as that at $h = 5$ (centre right panel), 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 5 suggest that the same conclusion will emerge more strongly when these forecasts face the same tests. Finally we note that the relative positions of the time series plots at $h = 1$ and $h = 9$ suggest that the MPC views inflation as being more persistent than viewed by the external forecasters.

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 SPF data reported by Giordani and Soderlind (2003, Fig.5), kindly updated by the second-named

author, and calculate the implied *RMSV* measure of collective uncertainty. The SPF questions have an end-year focus, corresponding to questions 1 and 2 of the SEF, although the SPF asks about year-on-year growth rates, whereas the SEF asks about the Q4 annual growth rates. Like many other SPF researchers, Giordani and Soderlind concentrate on the first-quarter surveys, hence results are available at horizons $h = 4$ and $h = 8$. Corresponding SEF data on uncertainty and disagreement are extracted from Figure 2 above, and 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 6.

The results for inflation at both horizons show that, after the initial credibility-gaining period, UK uncertainty and disagreement are 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 relatively stable, possibly reflecting the credibility of the Greenspan policy regime over the whole of this period. 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 time of the next survey, at the turn of the year.

The two countries’ GDP growth uncertainty measures are broadly similar, irrespective of differences in their monetary policy regimes and somewhat different macroeconomic experience, the US recession of March-November 2001 having had no counterpart in the UK. The most prominent feature in the US GDP forecast data is the peak in disagreement in early 2002. This may reflect a wide dispersion of views about the short-term effects on the real economy of the terrorist attacks of 11 September 2001, or the prospects for recovery from recession, the November 2001 turning point not having been announced by the NBER Business Cycle Dating Committee until July 2003.

The disagreement measures for the UK survey in general lie below their US counterparts, suggesting that the SEF forecasters are more homogeneous than the SPF forecasters. It is certainly the case that the SPF sample is much more widely dispersed, across a much larger country, than the SEF sample. Although modern communications technology can be expected to offset any simple location effects, there may also be some City of London “herd” effects in the SEF data.

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), who use the standard deviation of the implied probability density function of the short-term interest rate, six months ahead, 1997-99, for illustrative purposes (Chart 9). A fan chart describing financial market participants’ uncertainty about future interest rates, based on these data, appears in the Bank’s *Inflation Report*.

For comparative purposes the corresponding daily series for the short-term interest rate, three, six and twelve months ahead, are plotted in Figure 7, together with the “question 1” *RMSV* measure of collective inflation uncertainty taken from the top left panel of Figure 2. The options-based series is noisy, day-by-day, and the inflation uncertainty measure exhibits the periodic 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 the survey includes 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 article 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, while many questions remain to be explored.

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, and this is implemented in respect of forecasts of inflation and GDP growth. In the absence of such a direct measure of uncertainty it is sometimes suggested that a measure of disagreement be used instead, but our evidence indicates that this is not an adequate substitute, at least during what has been a relatively stable period for both inflation and GDP growth.

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 measure 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 options-implied interest rate uncertainty have no counterparts in inflation forecast uncertainty, whose lack of movement emphasises the secure anchoring of inflation expectations by a credible monetary policy.

REFERENCES

- Andrews, D.W.K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica*, 61, 821-856.
- Batchelor, R.A. (2007). Bias in macroeconomic forecasts. *International Journal of Forecasting*, 23, 189-203.
- Batchelor, R.A. and Dua, P. (1996). Empirical measures of inflation uncertainty: a cautionary note. *Applied Economics*, 28, 333-341.
- Boero, G., Smith, J. and Wallis, K.F. (2006). Evaluating a three-dimensional panel of point forecasts: the Bank of England Survey of External Forecasters. Unpublished paper, University of Warwick.
- Carruth, A., Dickerson, A. and Henley, A. (2000). What do we know about investment under uncertainty? *Journal of Economic Surveys*, 14, 119-153.
- Clements, M.P. (2004). Evaluating the Bank of England density forecasts of inflation. *Economic Journal*, 114, 844-866.
- Clews, R., Panigirtzoglou, N. and Proudman, J. (2000). Recent developments in extracting information from options markets. *Bank of England Quarterly Bulletin*, 40(1), 50-60.
- Das, M. and Donkers, B. (1999). How certain are Dutch households about future income? An empirical analysis. *Review of Income and Wealth*, 45, 325-338.
- Dominitz, J. and Manski, C.F. (1997). Using expectations data to study subjective income expectations. *Journal of the American Statistical Association*, 92, 855-867.
- Engelberg, J., Manski, C.F. and Williams, J. (2008). Comparing the point predictions and subjective probability distributions of professional forecasters. *Journal of Business and Economic Statistics*, forthcoming.

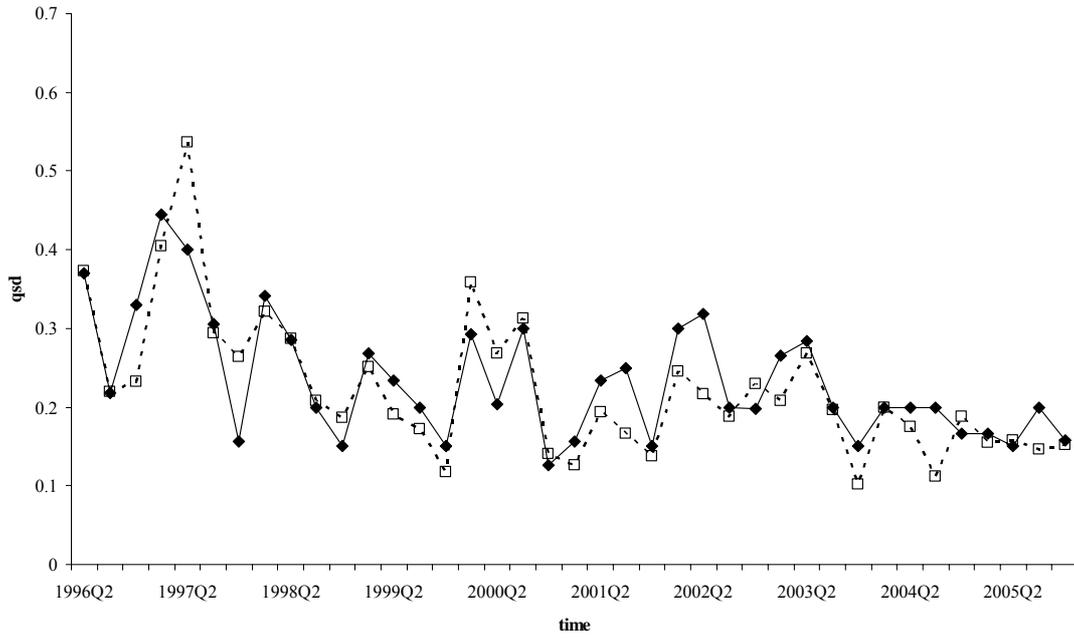
- Giordani, P. and Soderlind, P. (2003). Inflation forecast uncertainty. *European Economic Review*, 47, 1037-1059.
- Giordani, P. and Soderlind, P. (2006). Is there evidence of pessimism and doubt in subjective distributions? Implications for the equity premium puzzle. *Journal of Economic Dynamics and Control*, 30, 1027-1043.
- Guiso, L. and Parigi, G. (1999). Investment and demand uncertainty. *Quarterly Journal of Economics*, 114, 185-227.
- Guiso, L., Jappelli, T. and Terlizzese, D. (1992). Earnings uncertainty and precautionary saving. *Journal of Monetary Economics*, 30, 307-337.
- Hansen, B.E. (1997). Approximate asymptotic P values for structural-change tests. *Journal of Business and Economic Statistics*, 15, 60-67.
- King, M.A. (2005). Monetary policy: practice ahead of theory. The 2005 Mais lecture delivered at City University, London on 17 May 2005.
(<http://www.bankofengland.co.uk/publications/speeches/2005/speech245.pdf>)
- Lahiri, K. and Liu, F. (2006). Modelling multi-period inflation uncertainty using a panel of density forecasts. *Journal of Applied Econometrics*, 21, 1199-1219.
- Lahiri, K., Teigland, C. and Zaporowski, M. (1988). Interest rates and the subjective probability distribution of inflation forecasts. *Journal of Money, Credit, and Banking*, 20, 233-248.
- Manski, C.F. (2004). Measuring expectations. *Econometrica*, 72, 1329-1376.
- Rich, R. and Tracy, J. (2006). The relationship between expected inflation, disagreement and uncertainty: evidence from matched point and density forecasts. Staff Report No.253, Federal Reserve Bank of New York.
- Wallis, K.F. (2004). An assessment of Bank of England and National Institute inflation forecast uncertainties. *National Institute Economic Review*, No.189, 64-71.
- Wallis, K.F. (2005). Combining density and interval forecasts: a modest proposal. *Oxford Bulletin of Economics and Statistics*, 67, 983-994.
- Zarnowitz, V. (1969). The new ASA-NBER survey of forecasts by economic statisticians. *American Statistician*, 23(1), 12-16.
- Zarnowitz, V. and Lambros, L.A. (1987). Consensus and uncertainty in economic prediction. *Journal of Political Economy*, 95, 591-621.

Table 1. Divergences between point forecasts and density forecast means

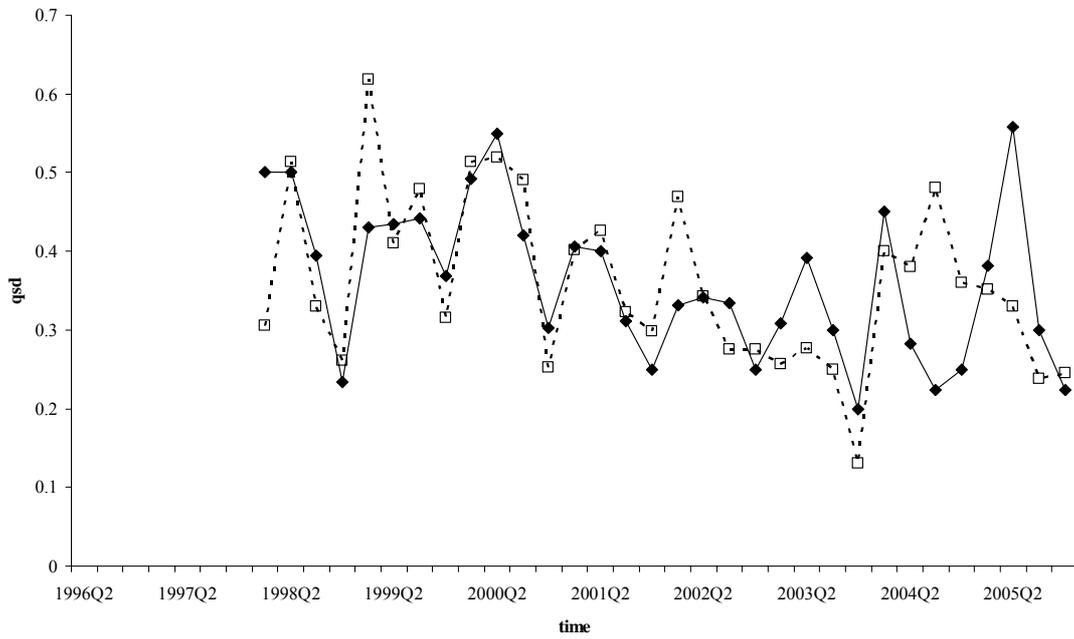
Percentage of cases (across all individuals and time periods) in which the point forecast lies above (below) the density forecast mean by more than 0.2 percentage points

	Inflation		GDP growth	
	above	below	above	below
Question 1	6.2	11.5	27.4	8.8
Question 2	10.2	14.1	44.2	6.6
Question 3	9.4	9.8	46.6	4.6

Inflation



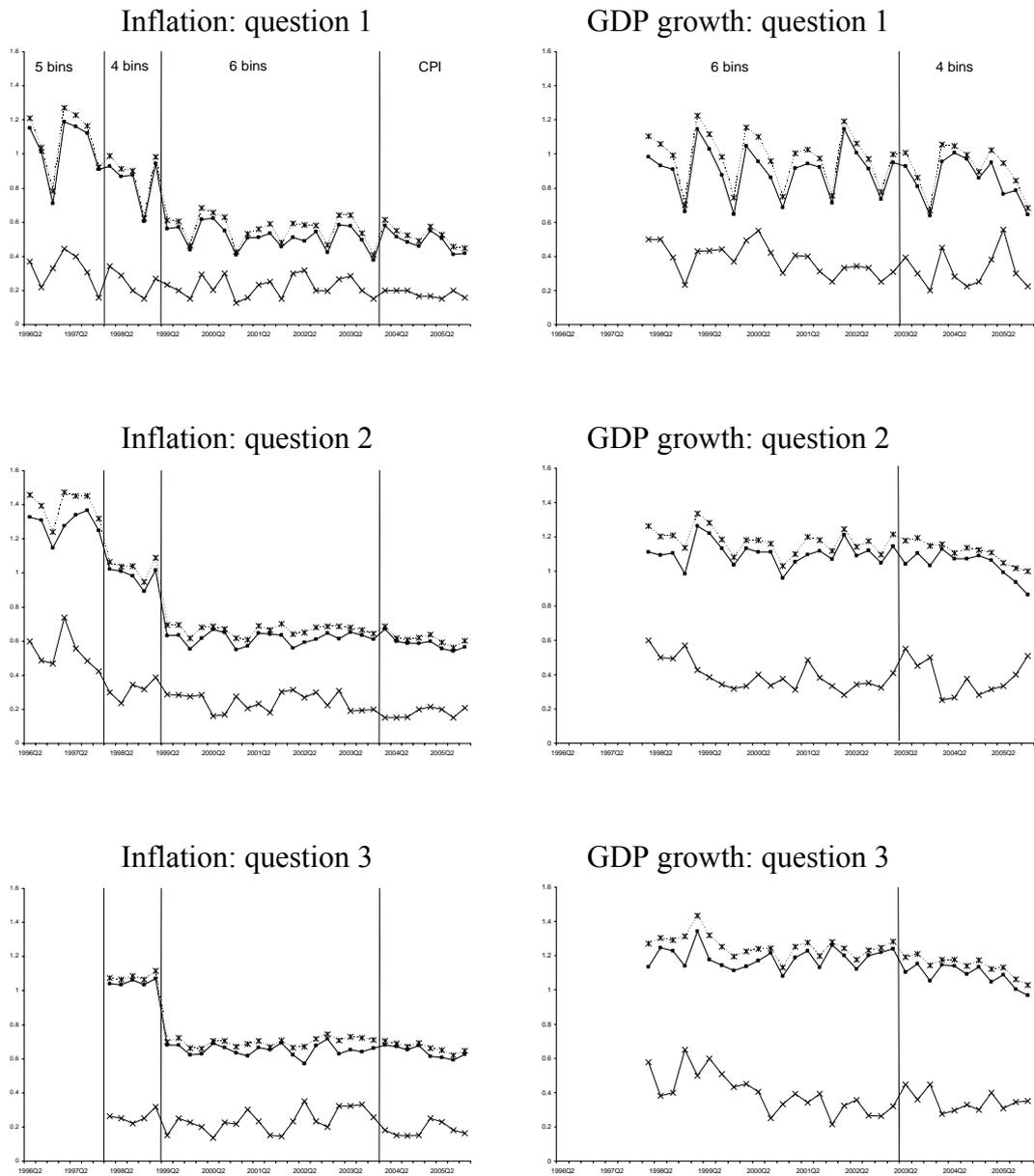
GDP growth



- □ - density forecast means

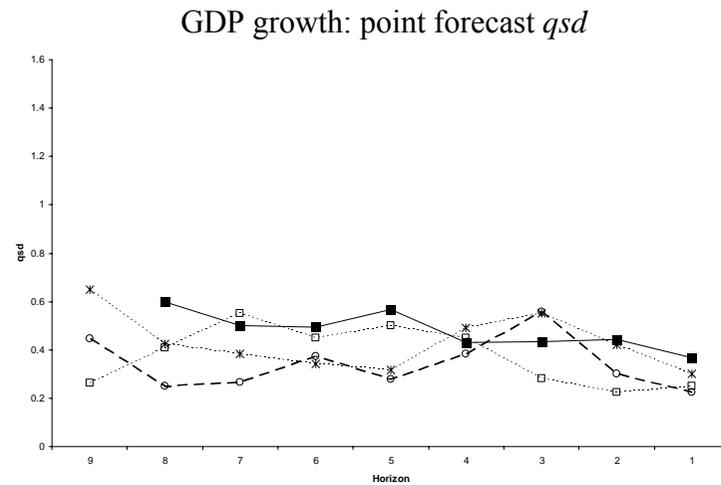
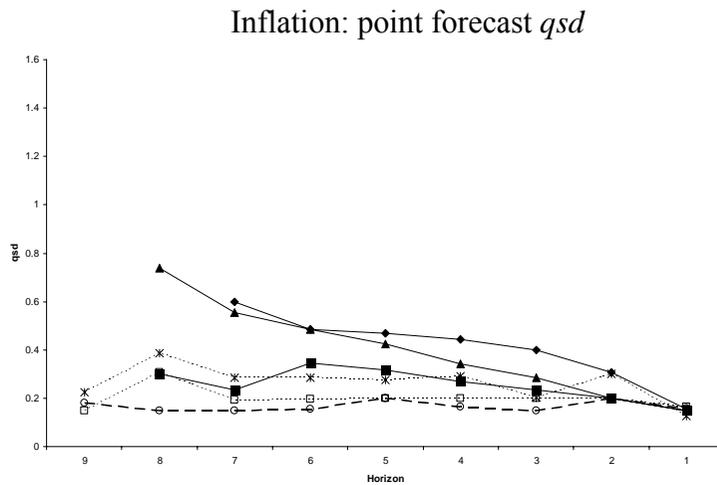
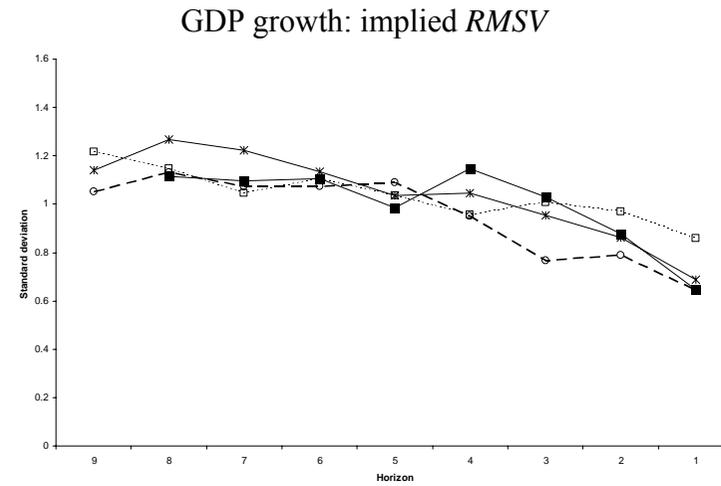
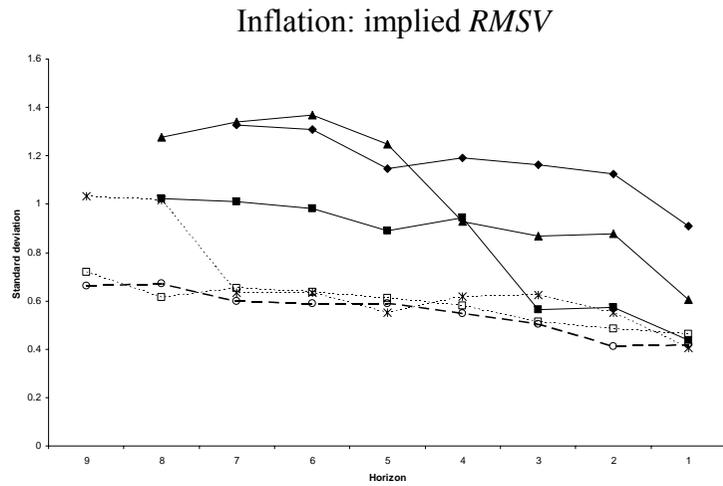
—◆— point forecasts

Figure 1. Alternative measures of disagreement: *qsds* of density forecast means and point forecasts, question 1



\dots aggregate sd, σ_A \perp implied *RMSV* \times point forecast *qsd*

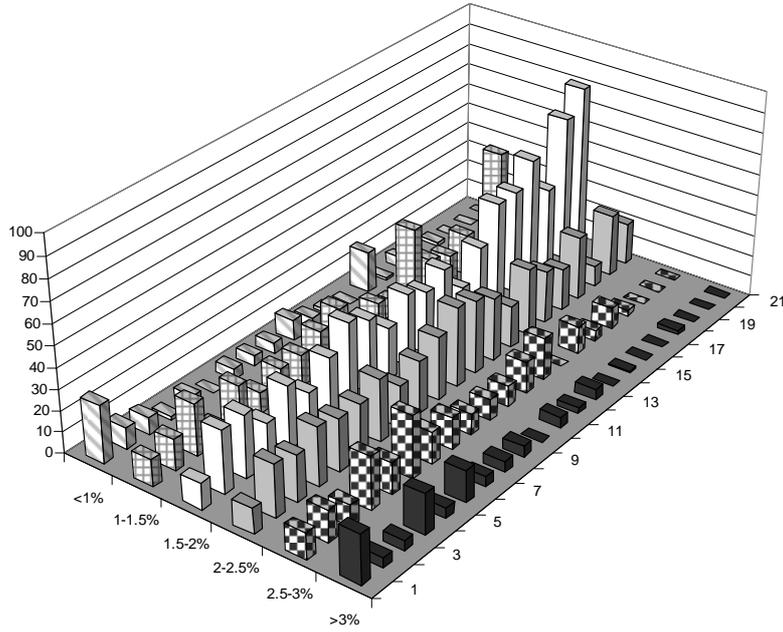
Figure 2. Measures of uncertainty and disagreement



◆ 1997Q4 ▲ 1998Q4 ■ 1999Q4 ⋈ 2000Q4 □ 2004Q4 ○ 2005Q4

Figure 3. Fixed-event forecasts: variation of uncertainty and disagreement with forecast horizon

Individual histograms two years ahead, November 2005



$RMSV \pm 2$ standard errors, $h = 9$

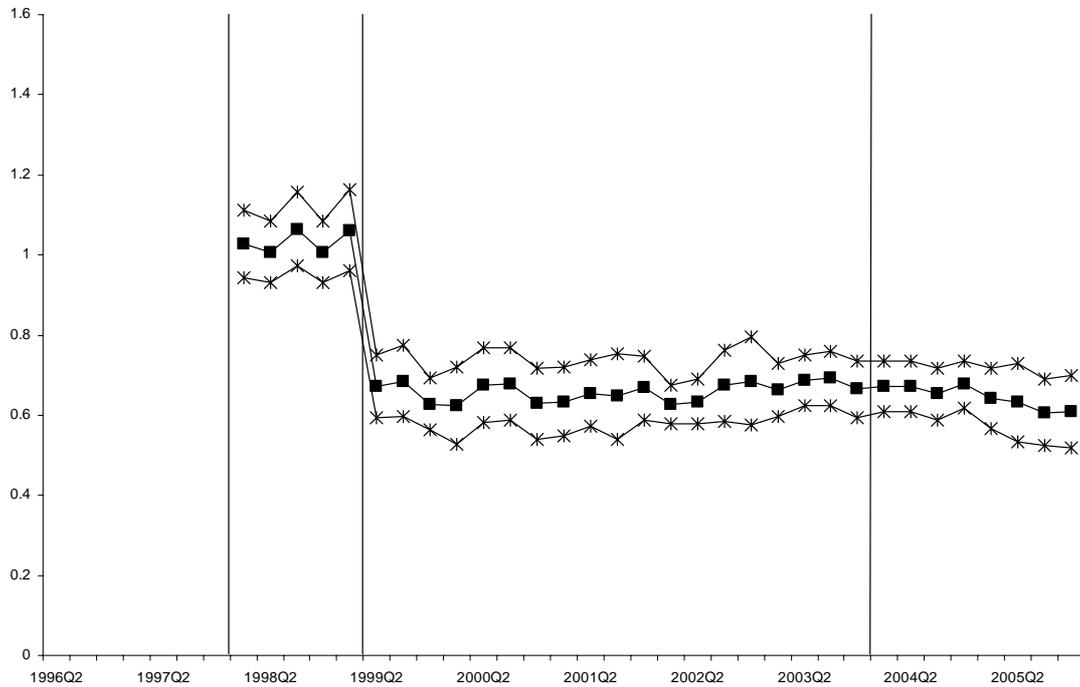


Figure 4. Disagreement about inflation uncertainty

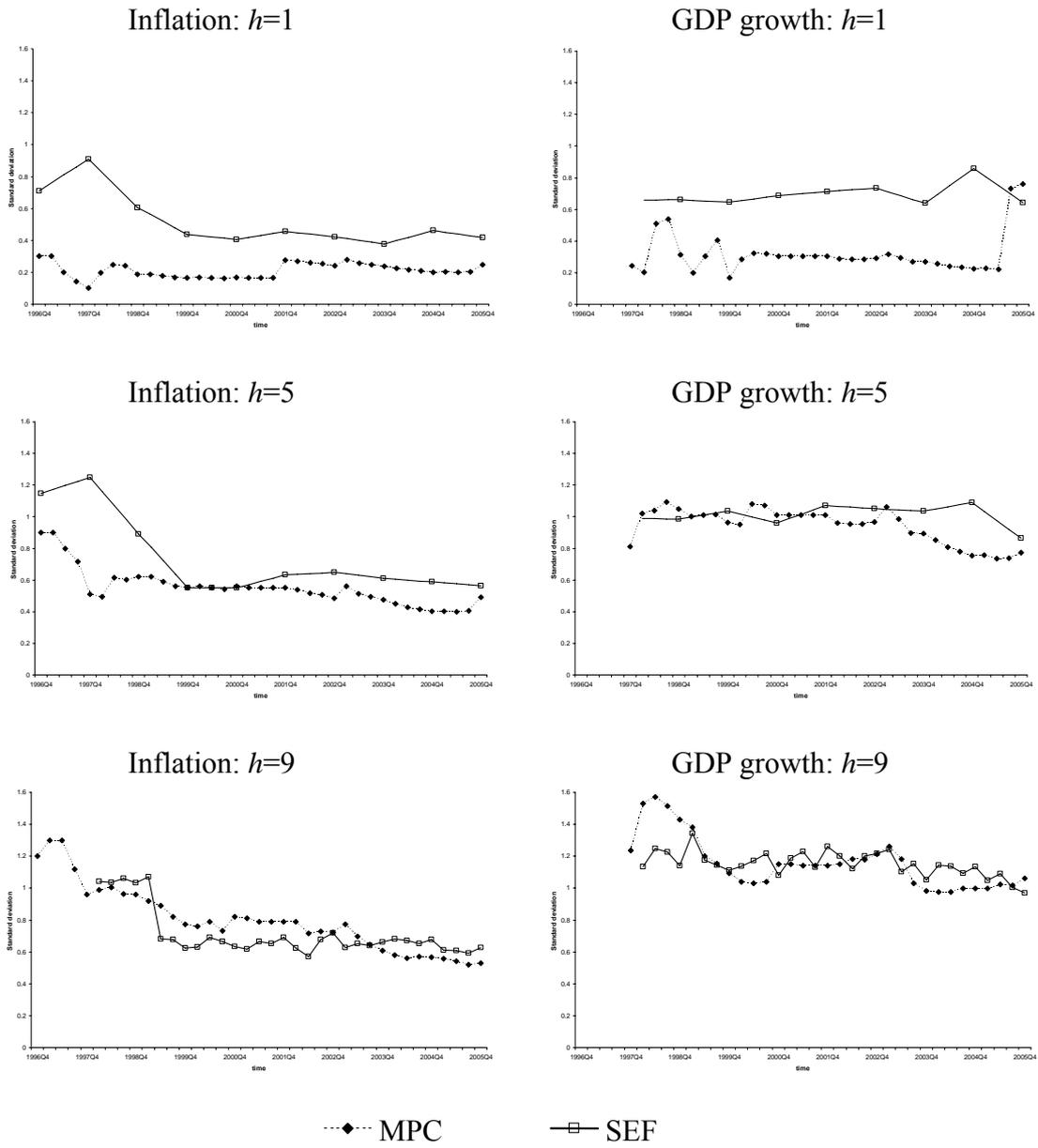
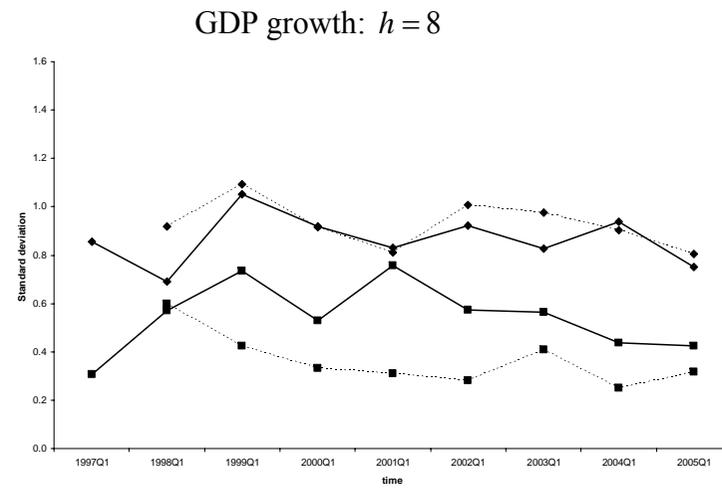
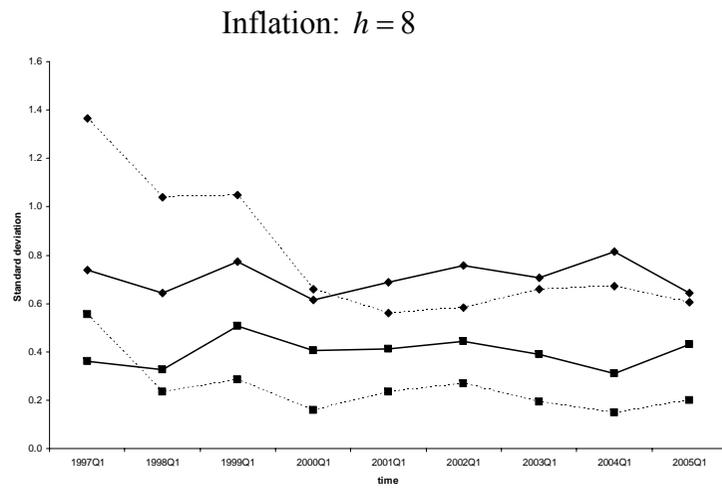
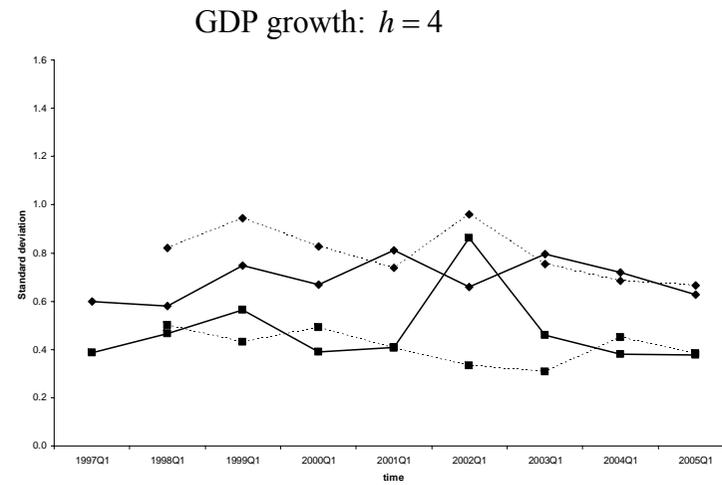
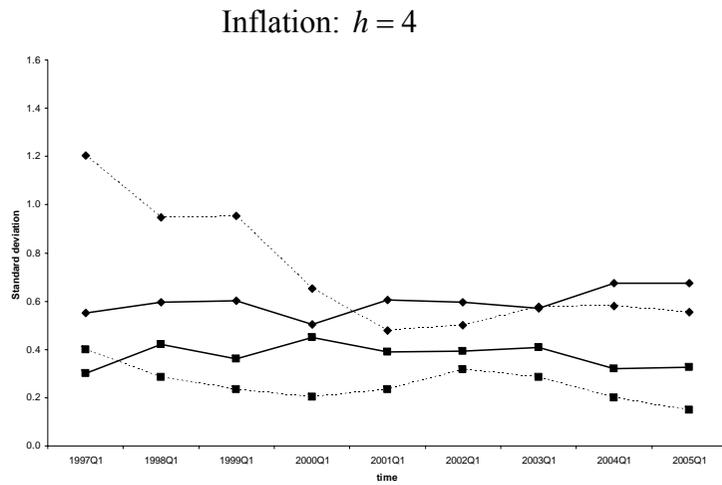


Figure 5. Uncertainty, MPC and SEF, 1996Q4-2005Q4



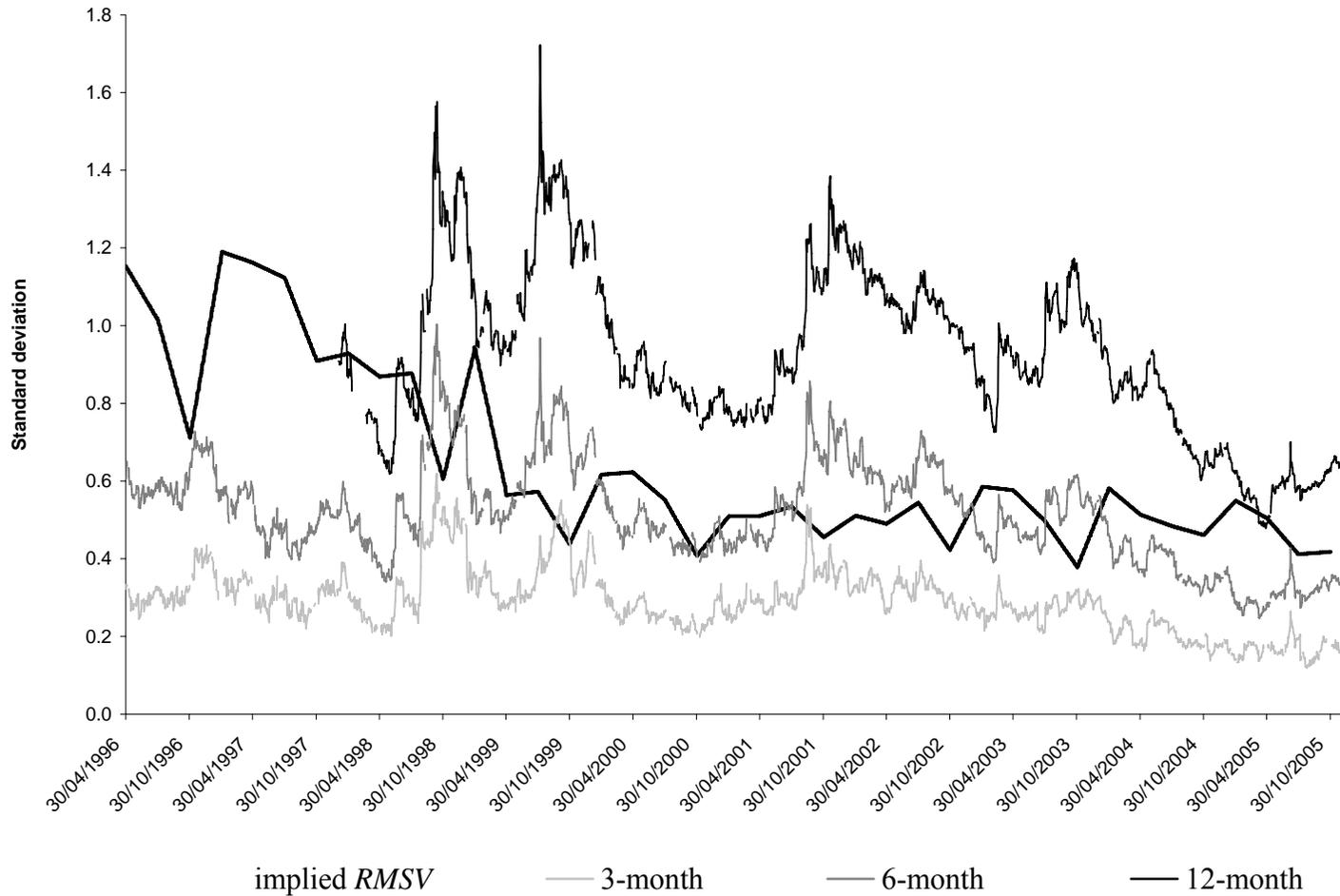


Figure 7. SEF inflation uncertainty (quarterly) and option-implied interest rate uncertainty (daily)