Estimating the uncertainty around forecasts for GDP growth and inflation: a stress testing approach

Paul Ormerod, Bridget Rosewell* and Will Cook

Volterra Consulting

February 2005

* corresponding author: brosewell@volterra.co.uk

Summary

- The Bank of England has introduced fan charts around the published forecasts for real GDP growth and inflation. These represent the Monetary Policy Committee's view of the probability of observing outcomes which differ from those of the central forecasts. Specifically, they represent the range within which the Committee expect the outcome to take place 90 per cent of the time, and are based upon the experience of the relevant variables over the most recent ten year moving window of data.

- This is a welcome step in recognising the inherent uncertainty which surrounds forecasts of the macro economy. But it raises the question of how to measure this uncertainty and what shape the fan charts should take.

- The fan charts reflect an implied probability density function of the outcomes which is based on the judgement of Monetary Policy Committee members. In this paper we compare the empirical evidence with the judgmental views of the Monetary Policy Committee. We do this by analysing post-war data on growth and inflation in the UK economy, and use a variety of standard techniques to calibrate empirical probability density functions.

- The 90 per cent probability range of outcomes based on analysis of past data is much wider than the range which appears in the Inflation Report, for both GDP growth and inflation. For 2005, the approximate 90 per cent range for growth plotted in the fan chart for GDP growth in the February 2005 *Inflation Report* is between +1.5 and +3.5 per cent. The empirically estimated range is from –1.5 to + 5.0 per cent. For inflation for 2005, the MPC range is between 1 and 3 per cent, but the empirical range is between –1 and +3.5 per cent.

- The wide ranges obtained from empirical probability density functions, even for 2005, are entirely consistent with the actual macro-economic forecasting record. Large errors can be made even on a one-year ahead basis, particularly at turning points, where macro forecasts are known to be at their worst. Forecasters find it particularly difficult to anticipate changes in the most recent pattern of events.

- Of course, this does *not* mean that either growth or inflation necessarily will fall below zero. But the MPC judges that the probability of either is very close to zero, whereas the empirical evidence suggests that there is approximately an 11 per cent chance of the former and a 13 per cent chance of the latter.

- We consider several arguments to try to reconcile the MPC judgmental range with the empirically estimated ranges. The most plausible is that the MPC ranges are constructed on the basis of market interest rate expectations being fulfilled in the future. In other words, the MPC fan charts abstract from interest rate uncertainty. This has the effect of making the range of the fan charts as published less than they would otherwise be if interest rate uncertainty were incorporated, but even so it is difficult to believe that the differences in the judgmental and empirically estimated ranges can be accounted for entirely by this point.

1 Introduction

The May 2002 Bank of England *Inflation Report* stated that 'Nobody can predict the future evolution of the economy with absolute certainty. It is more realistic for forecasters to recognise that uncertainty when describing their projections' (p.48). This is a very welcome advance in accepting the reality of the economy, but it raises a further question in identifying measures of that uncertainty. This paper addresses this question

The actual forecasting record of macro-economic variables underlines the difficulty. Indeed, it seems a distinct understatement to say that the economy cannot be predicted 'with absolute certainty'. For example, a detailed study of the one-year ahead forecasts for real GDP growth by the UK Treasury[1] showed that over the 1971-1996 period the mean absolute annual forecast error was 1.45% of GDP, compared to an actual mean absolute change of 2.10%. A more general examination of one year ahead real GDP growth rates in 17 Western economies concludes that 'the poor forecasting record of GDP growth by economists appears to be due to inherent characteristics of the data, and cannot be improved substantially no matter what economic theory or statistical technique is used to generate them. Over what is thought of as the time period of the business cycle in economics, in other words the period over which any regularity of behaviour of the growth of GDP might be postulated to exist, the genuine information content of correlations over time in the data is low'[2].

The Bank of England has therefore taken an important step in recognising the uncertainty. Fan charts are now drawn around the forecasts for real GDP growth and inflation. These represent the Monetary Policy Committee's view of the probability of observing outcomes which differ from those of the central forecasts. A discussion of the fan charts is given in the May 2002 *Inflation Report*, and a more detailed and technical explanation is available in the *Bank of England Quarterly Bulletin*, February 1998.

The Monetary Policy Committee is made up of five leading members of the Bank of England, plus four external members chosen for fixed periods of time by the Chancellor of the Exchequer. The Committee's fan charts of inflation and GDP growth 'represent the best collective judgement of the Committee, conditional on a particular path for interest rates'[3], and are essentially base upon the experience over the most recent ten year moving window of each of the relevant variables. The full range of the charts spans 90 per cent of the probability density function. In other words, the Committee expects the outcome to lie somewhere within the entire fan chart on 90 out of 100 occasions.

To examine how uncertainty is being measured in these fan charts we compare the range of the fan charts in the February 2005 *Inflation Report* with the range in the variables implied by their past history. In other words, we compare the probability density function formed judgementally by the members of the Monetary Policy Committee with

[1] C.Mellis and R.Whittaker, (1998), 'The Treasury forecasting record: some new results', *National Institute Economic Review*, 164, pp.65-79
[2] P.Ormerod and C.Mounfield, (2000),'Random Matrix Theory and the Failure of Macro-economic Forecasting', *Physica A*, 280, 497-504
[3] May 2002 Bank of England *Inflation Report*, p.48

the probabilities computed from the past history of real GDP growth and inflation in the UK.

Since they are formed judgementally, the fan charts of the Monetary Policy Committee (MPC) are not necessarily symmetric around the central projection, although in practice the assumption of symmetry does not seem to be violated in an immediately obvious way[4].

2 The data

2.1 GDP

Data on real GDP in the UK is available online from the Office for National Statistics on an annual basis from 1948 to 2004[5]. The rate of growth of this variable is plotted over the 1949-2004 period.

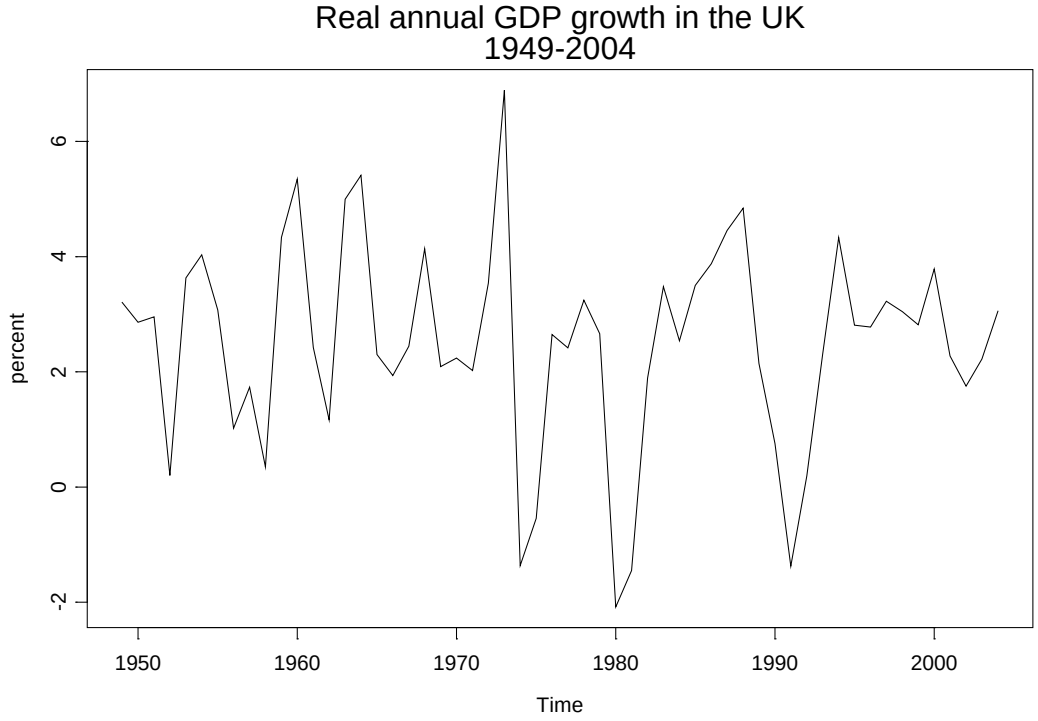

Figure 1 Real annual GDP growth in the UK, per cent

An obvious first step with the data in Figure 1 might seem to be to obtain estimates of the 5 and 95 percentile points in its probability distribution. These would give us the 90 percentage point spread. There are several options available as to how to compute these, and they are discussed in the Appendix.

[4] the fan charts are simply displayed graphically, rather than a table being presented with their ranges in it
[5] http://www.statistics.gov.uk/statbase/tsdintro.asp

However, there is, as always with economic time series data, a question of what is the relevant part of the sample of data 1949-2004 to use. The year 1973, for example, is often used to mark a watershed in post-war economic history, with the whole post-war period being divided into the period up to and including 1973, and the period afterwards[6].

We therefore split the data period into two, from 1949-1973 and 1974-2004, and carried out two statistical tests. First, a test of the null hypothesis that the mean of the growth rate was the same in the two periods and, second, a test of the null hypothesis that the variance of the growth rate was the same in the two periods[7]. The null hypothesis of equality of the means was rejected at a p value of 0.081 and of the variances at p = 0.54. In other words, the null hypotheses were not rejected at the conventional level of significance p = 0.05. The null hypothesis that the two distributions are the same more generally is rejected on a Kolmogorov-Smirnov test only at p = 0.68.

We then advanced the first period year-by-year, comparing 1949-1974 with 1975-2004, and so on up to 1949-1998 and 1999-2004. The null hypothesis of equality of means was never rejected at p = 0.05. The null hypothesis of equality of the variances was never rejected until the periods 1949-1991 and 1992-2004. During the latter short period of not much more than a decade, there have of course been no economic recessions in the UK, the last being in1991.

An additional point to consider is the autocorrelation function of real GDP growth. If there is strong autocorrelation, this would need to be taken into account in assessing the probability of any particular outcome in the future, given the recent past history of growth. The empirical autocorrelation function at lag 1 has a value of 0.281, and the null hypothesis that this is zero is rejected at p = 0.036. However, at no other lag of the function is the null hypothesis that the coefficient is zero rejected at p = 0.05, computing the function from 1 to 15 lags. There is therefore some weak, positive autocorrelation at lag 1, but the effect is decidedly weak. Further, the initial estimate for GDP growth in 2004 is just over 3.0 per cent, not too dissimilar from the mean over the entire period of 2.5 per cent.

The February 2005 *Inflation Outlook* states that the central projection for GDP growth is 'close to trend'. On the basis of the preceding paragraph, there is a (very) mild presumption that the potential range of outcomes might be biased very slightly above trend. But in the results reported in section 3 below, we ignore this potential complication, which is very much of second order of importance.

2.2 Inflation

The inflation variable which is targeted by the MPC has changed over the course of the MPC's existence. From being RPIX – prices excluding mortgage costs – it had become the Harmonised Index, now known as CPI. This excludes all forms of housing costs

[6] see, for example, A.Maddison, *Monitoring the World Economy 1820-1992*, 1995, OECD, Paris

[7] these were carried out using the commands 't.test' and 'var.test' in the statistical package S-Plus. In general, with t.test, the option that the variance in the two periods was not necessarily equal was used. Snedecor, G. W. and Cochran, W. G., *Statistical Methods*, 7th ed., 1980, Ames, Iowa: Iowa State University Press.

since the EU cannot agree how to include them. The Office for National Statistics has data on this particular variable only from 1989. For the period 1949-1988, we used the rate of change of the overall consumer expenditure deflator.[8]

Inflation presents us with more issues about the choice of sample period over which to calibrate a probability density function than does GDP growth. Figure 2 plots the rate of inflation.

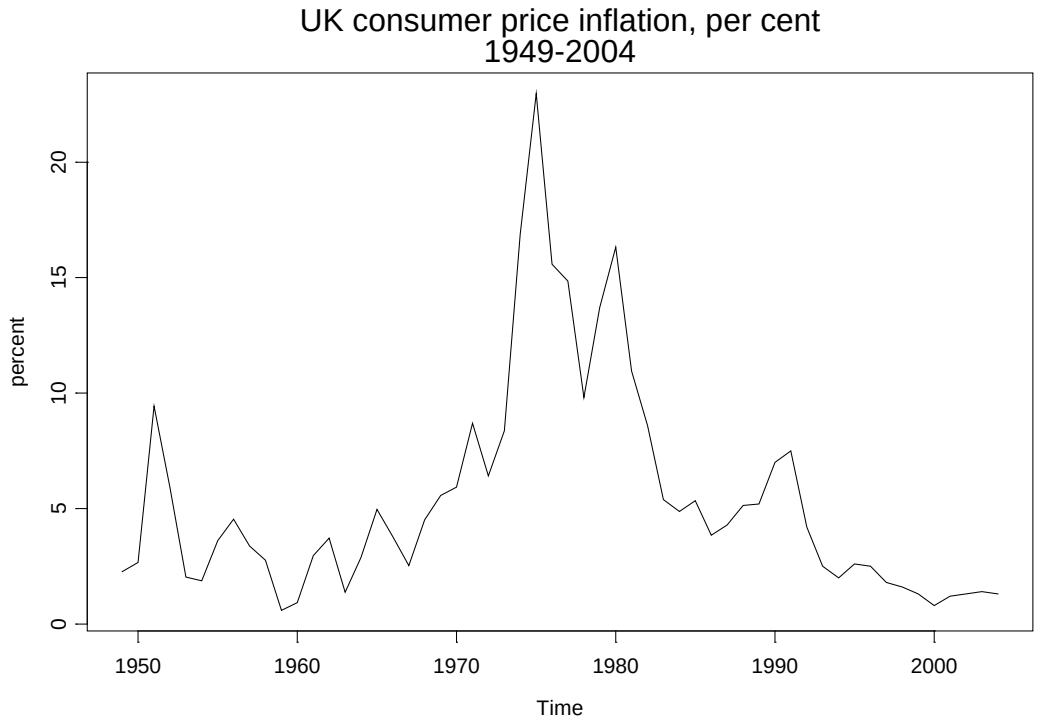

Figure 2 UK consumer price inflation, annual basis, per cent

An immediate point is that there is very strong positive autocorrelation of the inflation rate. This is entirely to be expected, because changes in costs today feed through into price and wage changes in the near future, introducing substantial inertia into the process of inflation. At lag 1, for example, the value of the autocorrelation function is 0.83.

However, this question can be dealt with in a straightforward way. The simple difference of the inflation rate gives an autocorrelation function which is effectively zero from lags 1 through 15, with the null hypothesis that the coefficient is zero only being rejected for lag 2, where the estimated coefficient is –0.29, and the null is rejected at $p = 0.032$.

Much more problematic is the choice of sample period. Figure 3 plots the difference in the annual inflation rate 1950-2004.

[8] data for 1963-1988 is available on the Office for National Statistics website, and the deflator for 1948-1962 was taken from C.H.Feinstein *National Income, Output and Expenditure in the United Kingdom 1855-1965*, CUP, 1972

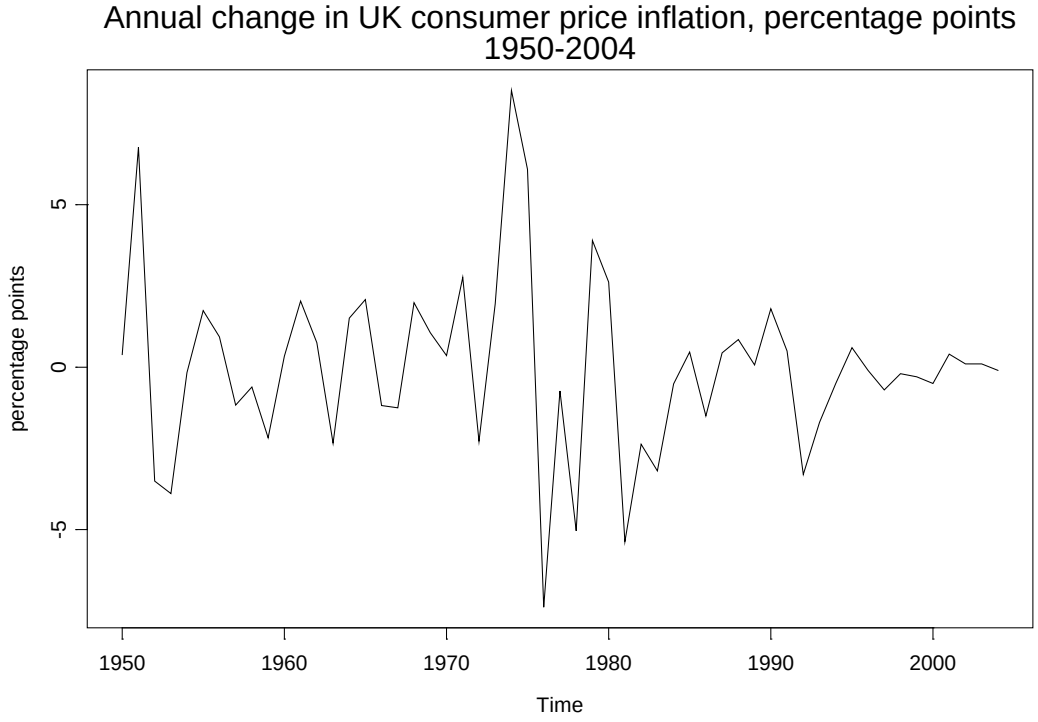

Figure 3 Annual change in UK consumer price inflation, percentage points

We can identify three years in which the UK suffered an external shock which had the effect of increasing inflation. This is not to say that the external shock was the only factor operating on inflation in these years, but on each occasion it was important. First, the Korean War in 1951; second, the first oil shock in 1974 and third the second oil shock in 1979/80. In each case, the increase in inflation triggered responses in the economy, in part by deliberate policy and in part by the endogenous movement of variables, which led to the sharp reductions shown in Figure 3 which followed the initial upward shocks.

The variance of the change in inflation is dramatically different depending upon whether the whole sample period is chosen, or whether the periods of inflationary shock and their aftermath (1951-53, 1974-83) are excluded.

We err on the conservative side and restrict our focus to the two more stable periods of historic inflation, namely 1954-1973 and 1984-2004. For the purpose of deriving confidence levels around inflation forecasts this restriction narrows the range of possible bands.

The null hypothesis that the means of the change in inflation are the same in the periods 1954-73 and 1984-2004 is only rejected at $p = 0.24$, and the null hypothesis that the variances are the same is rejected at $p = 0.06$, so it legitimate to combine these two periods into a single sample.

3 Results

The fan charts in Charts 1 and 2 of the *Inflation Report* show ranges of GDP growth and inflation over the forecast period. The outturns are expected to lie within these bands with a probability of 90 per cent. We do not have the exact numbers which correspond to the bands, but reading off the published charts, we can see that for 2005 the approximate ranges for the variables within the 90 per cent probability bands are:

GDP growth	1.5 to 3.5 per cent
Inflation	1.0 to 2.0 per cent

3.1 GDP

Consider first of all the growth rate of real GDP. In the Appendix, we note three potential ways of estimating a probability density function: linear interpolation, kernel density estimation, and using a hypothesised distribution. The latter can only be used if the relevant data can be shown to be compatible with a known statistical distribution.

Figure 4 plots a histogram of real UK GDP growth 1949-2004.

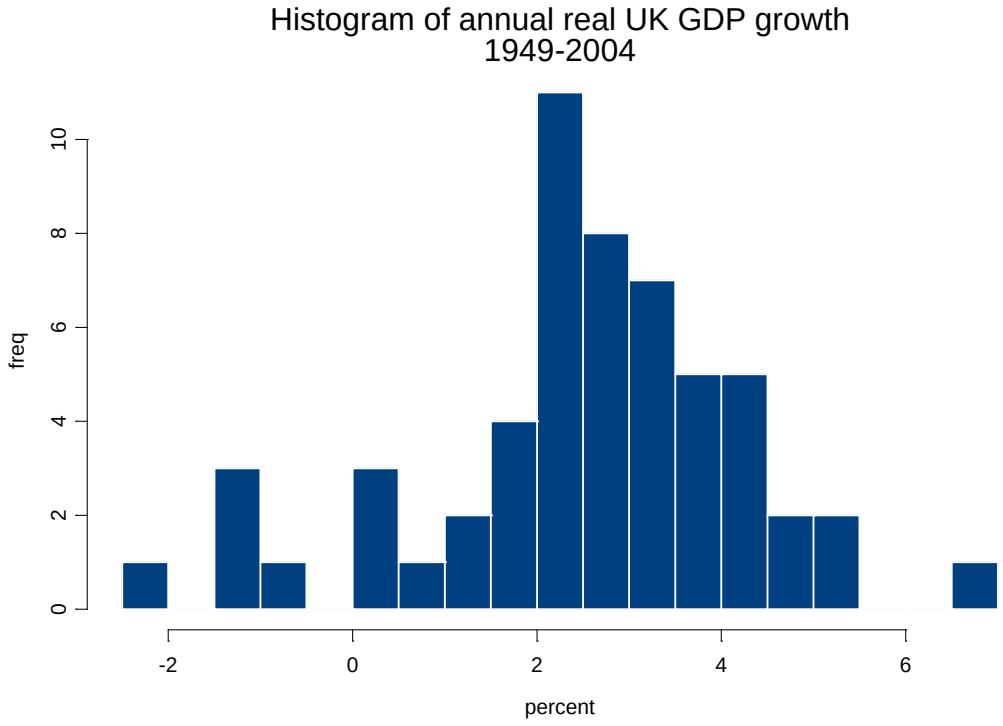

Figure 4

The data appears to follow as a broad approximation a normal distribution, but somewhat skewed to the left. In other words, there are more years with weaker growth than would be the case if the data were genuinely distributed normally. This is confirmed by a Kolmogorov-Smirnov test. The null hypothesis of normality is rejected at p = 0.013.

We therefore use linear interpolation and kernel density estimation. Of the two, the

latter should be regarded as being in general the more accurate.

Table 1 Range for real annual GDP growth rate (per cent) in the UK in 2005 : 90 per cent range

	Low	High	Probability of growth < 0
Monetary Policy Committee judgement	1.5	3.5	≅ 0
Linear interpolation of probability density function of data 1949-2004	-1.4	5.1	0.10
Kernel density estimation of pdf 1949-2004	-1.3	5.4	0.11

The statistically based estimates of the 90 per cent range of the data are much higher than the judgmentally based ones of the MPC. It may be thought implausible, for example, that a recession could take place in 2005. But anticipating turning points, and recessions in particular, is well known to be the weakest aspect of macroeconomic forecasting[9]. In recession years, forecasts made early in the year have often projected positive growth.

The fan chart drawn up by the MPC extends to the year 2008. The bands widen as the time horizon is extended, indicating increasing uncertainty about outcomes. It must be said that whilst this has a certain initial plausibility, it is not clear that the actual macroeconomic forecasting record bears this out. We are not aware of any systematic, published comparison of forecasts more than one year out but the impression of two of the present authors, with over 50 years practical experience of macro forecasting between them, is that the forecast errors do not in general increase much when the horizon is extended beyond one year. This is because the one year ahead forecasts themselves are often so inaccurate.

The MPC estimates that the 90 per cent band in 2008 is between 1.5 and 5.0 per cent. The upper end is consistent with the evidence of the post-war period. But the lower end remains very much higher. Indeed, a figure of +1.5 per cent growth corresponds to the 24th percentile, using the kernel density estimation technique. So the range for 2008 put forward by the MPC is of the order of 66 percentage points and not 90.

3.2 Inflation

In terms of inflation, the current rate is around 1.5 per cent, which the MPC sees as persisting during 2005, within a 90 per cent probability band of 1 to 2 per cent.

[9] see, for example, H.Stekler and R.Fildes, (1999), 'The state of macroeconomic forecasting', George Washington University, Center for Economic Research Discussion Paper 99-04; V.Zarnowitz and P.Braun, (1992), 'Twenty-two years of the NBER-ASA Quarterly Outlook Surveys: aspects and comparisons of forecasting performance', NBER Working Paper 3965

Table 2 sets out the various estimates for the 90 per cent probability band for 2005 using the data on the change in inflation. With the inflation data we cannot reject the hypothesis of the data being normally distributed, with a Kolmogorov –Smirnoff normality test returning a p-value of 0.5.

Table 2 Range for inflation (per cent) in the UK in 2005 : 90 per cent range

	Low	High	Probability of inflation < 0
Monetary Policy Committee judgement	1.0	3.0	≅ 0
Linear interpolation of probability density function of data	-0.8	3.5	0.14
Kernel density estimation of data	-1.0	3.8	0.13
Hypothesised distribution of data (normal distribution)	-0.7	3.8	0.12

It may at first sight seem surprising that negative inflation is possible given a history in which this has yet to occur. The current level of inflation is, however, very low, and thus a further downward shock could take it below zero. The chances are not high, being between 12 and 14 per cent depending upon which method of estimation is used. But the MPC is effectively ignoring this possibility completely.

Yet in fairly substantial sectors of the economy, negative inflation has been a feature for several years. Manufacturing producer prices, for example, fell in every year between 1996 and 2002. The annual rate of increase in average earnings in the construction sector, after having been around 5 per cent for several years, has fallen to zero because of an increased labour supply from the new EU countries.

3.3 Explaining the differences

How can we reconcile the much wider bands obtained from the analysis of economic history, both for GDP growth and inflation, with the judgmental bands drawn up by members of the MPC?

There are three possible arguments. First, the existence of the MPC might mean that the macro economy can be managed much more effectively than in the past and hence more extreme outcomes can be avoided. As far as we know, no member of the MPC itself has made this point, so it is not a criticism of them to say that it appears to be without substance. There is no evidence that the Bank of England has unique insights into macro economics which enables it both predict the future better than any outside forecasters, and to understand the impact of policy instruments better than any outsider. It might be argued that the MPC will take fewer risks than when decisions are made by politicians, but since neither group is in charge of all the risks, this is not convincing either.

The second is that there has been in general a more decisive shift of behaviour in the UK economy which means that most of the post-war data is irrelevant to the future. The creation of the MPC may or may not be one of the factors behind such a hypothesised shift of behaviour. It is simply not possible to provide decisive evidence on this, not least since this argument means that there are very few data points in the 'new' regime.

But the onus here is on those who believe there has been a decisive shift to attempt to demonstrate it, particularly given the fact that over much of the post-war period the statistical distribution of GDP growth has been stable with respect to the choice of most possible sample sub-periods, and that the same is true of the change in inflation except for the periods following the major external shocks of the Korean War and the two oil price shocks. For what it is worth, the opinion of the present authors is that the supply side reforms of the 1980s did raise the sustainable long-term growth rate of the economy. But the differences between the MPC probability bands and the bands estimated from economic history are so substantial that it is implausible that the performance of the economy has improved to such a dramatic extent.

The third point is the most effective. The 90 per cent bands drawn up by the MPC are prepared on the basis of market interest rate expectations being fulfilled. Relaxing this assumption would undoubtedly widen the band, though we cannot say how far, since this depends upon the judgement of the members of the MPC about what the bands around interest rates are. As they are presented, however, this point means that the bands are undoubtedly an underestimate of the true potential range.

4 Conclusion

The Monetary Policy Committee takes into account the uncertainty around any given macroeconomic forecast for GDP growth and inflation. The Committee does this by forming a judgement about the range within which they expect the outcome to occur 90 per cent of the time.

We compare the judgmental bands with the evidence from post-war data on growth and inflation in the UK. We estimate empirically the 90 per cent probability ranges for both growth and the change in inflation. There are several technical ways of doing this and the results are very robust with respect to the techniques used.

There are issues concerning the choice of sample period of the estimates, particularly for inflation. The estimates are very robust with respect to choice of sample period for GDP growth. In our estimates for inflation, we exclude the inflationary shocks of the Korean War and the two oil shocks, and their immediate aftermaths. The inclusion of these periods in the estimates would widen even further the 90 per cent probability range.

We find that the empirically estimated ranges are considerably wider than the judgmental ones of the MPC. For growth, the MPC's 90 per cent probability range for 2005 is between 1.5 and 3.5 per cent, whereas the empirically based estimates are between –1 and + 5 per cent. For inflation, the MPC's range is +1 to +3 per cent, compared to around –1 and +3.5 per cent.

This evidence is consistent with the actual macroeconomic forecasting record. The level of accuracy even of one-year ahead forecasts is low, but it is particularly bad at turning points. Forecasters seem to find it especially hard to anticipate changes in recent trends.

The results do not of course mean that either growth or inflation will be below zero, but that the 90 per cent probability range encompasses values below zero for both these variables. The MPC appear to be assigning far too low a probability to the chance of either an economic recession or deflation, when inflation becomes negative. We estimate that the probability of growth being below zero is approximately 11 per cent, and of inflation becoming negative of 13 per cent.

Appendix

Techniques for calculating quantiles of data

1 Linear interpolation

To estimate the p percentile point, we sort the observed data into ascending order, and work our up from the lowest observation until we find the observation for which we have p percent of the data below (inclusively). That observation is then the p percentile point. Clearly we will not always hit the p percent exactly, so in this case we take a weighted average of the observed value, and the next lowest value. Expressed mathematically:

$$X = x_k + r(x_{k+1} - x_k)$$

where X is our p percentile point and x_i are the sorted data and:

$$k = \left\lfloor \frac{(n-1)p}{100} \right\rfloor + 1 \quad \text{and} \quad r = \frac{(n-1)p}{100} - \left\lfloor \frac{(n-1)p}{100} \right\rfloor$$

n = # of data points

2 Kernel density estimate

This approach requires no assumptions about the distribution of the data. A probability density function is fitted to the data, and the relevant percentile point is estimated from this. The density function is fitted with what is known as a kernel estimate, with Gaussian window and standard histogram-bin bandwidth. This is described in more detail in the seminal work by Silverman, *Density Estimation for Statistics and Data Analysis*, Chapman and Hall,1996.

The density function results come in the form of a set of coordinates (x_i, d_i), where i = 1, ... m. These coordinates give the relative probabilities, d_i, of observing different values, x_i. Up to a certain limit, the larger the number of coordinates, the better the estimate of X, the p percentile point, will be.

The approach requires us to find k such that $\sum_{i=1}^{k} d_i \le \frac{q}{100} \le \sum_{i=1}^{k+1} d_i$ then the estimate of X is given by

$$X = x_k + \frac{\left(\frac{q}{100} - d_k\right)}{(d_{k+1} - d_k)}(x_{k+1} - x_k)$$

This is an attractive approach in many contexts, since it does not require us to make any assumptions about the distribution of the data - assumptions which carry non-zero probabilities of being wrong.

3 Hypothesised distribution

If the distribution of data can be shown to match that of a classic statistical distribution then it is possible to calculate the desired quantile directly from the hypothesised distribution, rather than relying on the specific observations within the data.

The Kolmogorov-Smirnov test is a powerful statistical test of the hypothesis that a set of data from a continuous distribution follows a particular distribution (normal, lognormal, exponential, or whatever). It is known to be more powerful than the chi-square goodness of fit test (S-Plus ref here).

We can use the K-S test to test the null hypothesis that our observed data is draw from a particular distribution. If the null hypothesis is not rejected (at a suitable p-value), we can hypothesize that the data is drawn from such a distribution.

Suppose, by way of example, that the null hypothesis of normality is not rejected. We consider, then, that the data are drawn from a normal distribution with mean equal to the sample mean and variance equal to the sample variance. The required quantile can then be estimated on the basis of this simply specified normal distribution.

It is important to note, however, that this option will not always be available. In practice, it might not be possible to find a known distribution such that the null hypothesis that the data is drawn from such a distribution is not rejected.