Evaluating forecast performance

Independent Evaluation Office | November 2015

BANK OF ENGLAND

BANK OF ENGLAND

Evaluating forecast performance

Independent Evaluation Office, Bank of England

Foreword from the Chairman of Court

The Bank of England's Court of Directors, like all company Boards, is collectively responsible for the sustainable success of the institution. Court's statutory duties, laid down by Parliament, include keeping the performance and financial management of the Bank under review.

All well-functioning Boards need to engage in constructive challenge and robust debate. Since its formation in September 2014, the Bank's Independent Evaluation Office (IEO) has played an important role in ensuring that this challenge and debate takes place, alongside providing support to Court in the effective discharge of its statutory duties.

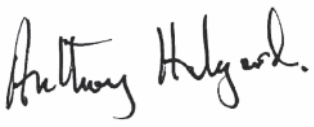
Operating at arm's length from the business areas, and reporting directly to me in my capacity as Chair of Court, the IEO has strengthened Court oversight of the Bank's performance across the full range of its policy functions. It has enhanced the Bank's regular reporting mechanisms, supported the external review of Monetary Policy Committee (MPC) transparency by Governor Kevin Warsh and put in train in-depth work programmes, such as the one set out in this paper.

In this paper, the IEO details the work programme initiated towards the end of 2014 to provide Court with a better basis for evaluating the Bank's forecast performance. Forecasting is an essential input to the work of the MPC, and, increasingly since the advent of the 'One Bank' strategy, into the work of the Financial Policy Committee and the Prudential Regulation Authority Board. The publication of forecasts, which occurs quarterly in each *Inflation Report*, is also an important output of the MPC, and forms a key part of the Committee's communication strategy.

As the IEO's work programme makes clear, many of the Bank's core macroeconomic forecasts perform well when assessed both against standard statistical criteria and against forecasts produced by other institutions. Nevertheless, as Court has discussed, there remains scope to do more to improve performance. The work presented in this paper has already served as an important input to a range of initiatives agreed by the MPC and set out in the November 2015 *Inflation Report*. Implementation of these initiatives will be monitored by Court. As part of its regular review of MPC processes, Court will also be monitoring forecast performance.

In putting this paper into the public domain, the IEO hopes to facilitate debate about its empirical methods and analytical approach, as well as to build wider understanding of its work.

We welcome feedback.



Anthony Habgood
November 2015

The Bank of England's Independent Evaluation Office can be contacted at independentevaluation@bankofengland.co.uk

Contents

Summary

At the meeting of the Court of the Bank of England (hereafter 'the Bank') on 29 September 2014, the Independent Evaluation Office (IEO) was requested to put in place a work programme that would give Court Directors 'a better basis for evaluating the Bank's forecasting performance'. That work programme has been an input into a number of proposed changes to the Monetary Policy Committee's (MPC's) forecasting processes,(1) and, more generally, is intended to facilitate Court oversight of this important aspect of the Bank's work.(2)

This paper provides details of the IEO's work programme, including methodology, some important issues of context, statistics on forecast performance for eleven macroeconomic variables over the period 1997–2014 and some recommendations for change. The work programme has benefited from access to a wide range of unpublished data held by the Bank, and consequently represents the most comprehensive statistical study of Bank forecast performance that has been undertaken to date. A relatively novel aspect of our analysis is the 'benchmarking' of Bank forecast accuracy to that of individual private sector forecasters and other central banks.

We would note at the outset that, as with all statistical exercises, interpretation of empirical results needs to take account of background and context. An important aspect of this is sample size and statistical significance. For some aspects of our evaluation exercise — most notably, our benchmarking work and the assessment of post-crisis forecast performance — data constraints meant that our sample was necessarily smaller than ideal. Small samples usually imply greater uncertainty about the estimated parameters in statistical tests, and make it more difficult to determine the degree to which empirical results are driven by pure chance rather than forecasting ability.

More generally, the scope of this paper is, by construction, focused primarily on only one aspect of evaluating forecast performance — statistical diagnostics. At the Bank, and at central banks more broadly, forecasting encompasses much more than providing point estimates of macroeconomic variables, and our empirical exercise should be regarded as complementary to the broader assessment of Bank forecast process and culture provided by the independent 2012 Court Review of the MPC's Forecasting Capability (Stockton (2012)).

A founding principle of the Bank's Independent Evaluation Office is that it operates at arm's length from local business areas and reports directly to the Chairman of the Bank's Court. To help safeguard the quality of the IEO's work, as well as its independence, our empirical approach has been peer reviewed by James Mitchell, Professor of Economic Modelling and Forecasting at Warwick Business School.

Annex 2 provides a summary of Professor Mitchell's views, concluding that this paper provides: 'a fair and balanced statistical evaluation of the Bank's forecasts using a battery of appropriate forecast evaluation tests; sensible conclusions are drawn with worthy recommendations made to improve forecast performance in some areas'. The IEO would like to thank Professor Mitchell for his advice and support.(3)

In setting out our conclusions and recommendations, we would note that this is necessarily a review of past performance. For two year ahead forecasts, for example, our sample included forecasts made in *Inflation Reports* up to and including 2013. We would also note that the Bank's forecast processes have changed in various ways since the 2012 Stockton Review; it is too soon to assess the impact of these changes.

(1) See the box in Section 5 of the November 2015 *Inflation Report*.
(2) This paper was approved for publication by the Court of the Bank of England on 26 October 2015.
(3) We would note that the analysis and views expressed in this report, as well as any errors contained herein, are the responsibility of the IEO, and not of Professor Mitchell, nor of the Bank of England more widely.

Our main empirical results, by variable, can be summarised as follows:

- For **UK GDP growth**, there was no statistically significant evidence of bias; in other words, projections did not systematically under or overpredict outturns. Neither was there, for one year ahead forecasts, any statistically significant evidence that forecasts could have made more systematic use of information known at the time, such as past forecast errors or data outturns (in other words, these forecasts were strongly efficient). For two year ahead forecasts, however, errors were positively and significantly related to both past forecast errors and data outturns. And we observed a systematic tendency for two year ahead GDP forecasts to cluster at around the pre-recession (1997–2007) average growth rate of around 3%. The accuracy of the Bank's UK GDP growth forecasts compared favourably with the accuracy of UK private sector forecasts, and with domestic GDP forecasts made by other central banks, particularly so at the one-year horizon.

- For **UK inflation**, as with UK GDP growth, we found no statistically significant evidence of bias at any forecast horizon. Neither was there, for one quarter and one year ahead forecasts, any statistically significant evidence of strong inefficiency. For two year ahead inflation forecasts, however, errors were systematically related to data outturns known at the time the projections were made; we additionally observed a tendency for two year ahead forecasts to cluster around, or a little below, the inflation target. In our benchmarking exercise, the Bank's one year ahead inflation projections performed well — for example, they were typically as accurate, if not more so, than those of the majority of UK private sector forecasters in our sample. The accuracy of the two year ahead projections compared less favourably, however, particularly in the post-crisis period, although our sample period is too short to draw definitive conclusions from this observation.

- For the **UK unemployment rate** and **wage growth**, we found statistically significant evidence of bias (Bank forecasts typically too high) and strong inefficiency. Despite this, at the one year ahead horizon, the accuracy of the Bank's wage growth projections tended to compare favourably with that of private sector forecasters. Meanwhile, the accuracy of the unemployment rate forecasts did not compare favourably with either UK private sector forecasters or other central banks. There was also evidence that the Bank's unemployment rate forecasts were less accurate over the sample period as a whole than a forecast generated by a simple statistical rule, although the difference in accuracy was no longer statistically significant if the 1998–2000 period was excluded (we observed relatively large forecast errors for the unemployment rate at the beginning of the sample period). Similarly, when we excluded the pre-2001 period, the observation of bias in the unemployment rate forecasts was no longer statistically significant. We did note some improvement in the relative accuracy of the Bank's unemployment forecast at the end of our sample, but that improvement was too recent to be confident that it represented a systematic improvement.

- Other forecasts considered in this evaluation were growth in **UK consumer spending, investment, house prices, household lending, corporate lending, US GDP** and **euro-area GDP**. There was no statistically significant evidence of bias in UK consumer spending or investment growth projections, although there was evidence of strong inefficiency at some forecast horizons. There was statistically significant evidence of bias in the Bank's UK house price inflation forecasts over the full sample (forecasts systematically too low). Despite this observation, the accuracy of the Bank's house price projections in recent years has typically compared favourably with those of private sector forecasts, although the short sample of available data means we cannot be confident in the robustness of this comparison. In terms of UK nominal lending growth forecasts, data constraints meant that the sample period under consideration was materially shorter than those for other variables; consequently, our observation of statistically significant bias in the Bank's one year ahead corporate lending projections may not prove robust. Finally, Bank forecasts for US and euro-area GDP growth have tended to be as accurate as those made by the US Federal Reserve and the European Central Bank.

The empirical results summarised above are based on analysis of point estimates of central tendency (specifically, the mode) of Bank projections. This reflects data constraints — for most of the variables under consideration, only forecasts of the mode were available for our full sample period. The exceptions were UK GDP growth and inflation, where we had data on the full probability distributions, as summarised by the Bank's fan charts. Formal statistical tests of these probability distributions suggested that the fan charts have not tended to provide an accurate guide to the eventual distribution of UK GDP growth and inflation outturns. That may reflect a

tendency, particularly since 2007, to understate the probability of especially low outturns for GDP growth, and of especially high outturns for inflation.

In general, our empirical results suggest that there are a number of areas where the Bank's core forecasts have performed well from a statistical standpoint. Our work did, however, suggest two aspects of performance that merit further investigation.

First, we observed a tendency for some two year ahead projections to be less accurate than one year ahead projections, particularly so in the 2010–14 period. Moreover, for some variables, the Bank's two year ahead projections have tended to be less accurate than those of private sector forecasters/other central banks. The post-crisis period and the comparisons with external forecasts necessarily involve small samples; consequently, these findings should be interpreted cautiously. They are nevertheless consistent with a tendency to assume too little persistence of economic shocks.

The second area of note relates to the Bank's unemployment rate forecasting record: of the eleven variables evaluated, the unemployment rate forecasts performed the least well against our statistical tests, including tests of unbiasedness and efficiency. To some extent, this reflected relatively large unemployment rate forecast errors in the early part of our sample, however, a period when the UK labour market was thought to be undergoing structural change.

These general observations suggest that there are areas where the Bank's forecast performance may be improved by more systematic use of alternative statistical models. This is consistent with the findings of a number of other published studies of Bank forecast performance (eg Fawcett *et al* (2015) which notes that, for UK GDP growth and inflation, statistical models have had a tendency to outperform Bank forecasts at longer horizons). More broadly, our empirical work, together with the numerous interactions we had with producers and (internal) users of the Bank's macroeconomic forecasts as part of this study, suggest that the following recommendations could assist the Bank in improving its forecasting capability:

- **Learn more from other forecasters and models**. While there is a balance to be struck between models that are optimised for forecast performance, and models that aid policymakers in constructing a coherent economic narrative, our work suggests that there could be gains from investigating (possibly in conjunction with external academics/forecasters) more systematic use of a range of statistical models;

- **Learn more systematically from the past**, including more systematic monitoring of forecast performance, which we would recommend is reported to the Bank's Court on a regular basis;

- **Challenge conventions more**, for example on the speed with which economic shocks are assumed to unwind. We recognise, however, that considerable changes have been made, following the 2012 Stockton Review, to generate more internal challenge to the Bank's forecasts, and it is too soon to assess properly the impact of these changes;

- **Provide more support for non-MPC internal users of the Bank's forecasts**. The Bank's macroeconomic forecasts are increasingly being used to inform analysis presented to the Financial Policy Committee and the Prudential Regulation Authority Board. Some members of these Committees, and some of the staff supporting them, have not previously worked closely with the Bank's macroeconomic forecasts. To help them use these forecasts more effectively, Bank staff should consider giving key forecasts a score based on past performance and the degree of oversight by the MPC.

1 Forecasting at the Bank of England

1.1 The purpose of forecasting at central banks

Forecasting at the Bank, and at central banks more widely, makes a core contribution to policymaking. That contribution is considerably broader than the production of numerical forecasts for GDP and inflation. As Stockton (2012) states: 'Much external attention to the forecasting enterprise is focussed on the numbers. (…) But most policymakers at central banks would likely agree that point forecasts alone provide insufficient information on which to take decisions.'

Specifically, most practitioners within the central banking community (see, for example, Stockton (2012)) would identify three broad, and mutually reinforcing, functions served by an effective forecast process:

- First, that the forecasting round should help to foster a shared understanding of the evolving economic narrative, the likely outlook for the economy and the risks around that outlook. Greenspan (2005) summarises the 'risk management' approach that typically characterises policymaking at central banks as follows: 'the paradigm upon which we have settled has come to involve… crucial elements of risk management. In this approach, a central bank needs to consider not only the most likely future path for the economy, but also the distribution of possible outcomes about that path'.

- Second, that the forecasting round should allow consideration of the various merits and demerits of alternative monetary policy strategies. The Bank, for example, has for many years published forecast projections under alternative assumptions about monetary policy (specifically, an assumption that policy rates evolve in line with the market yield curve, and an alternative assumption that policy rates remain constant). This is supplemented by regular discussions during the forecast round, and at MPC meetings themselves, of monetary policy strategy.

- Third, that the forecasting round should serve as an effective communication vehicle for the central bank's current thinking on the economic outlook, the risks, and the appropriate monetary stance. It is widely accepted within the macroeconomics community that good communication can play a core role in shaping external expectations of policy, thereby helping central banks fulfil their mandates more effectively. As Warsh (2014) states: 'Effective communication is essential to the successful implementation of monetary policy. (…) The goal is to transmit effectively, so that the signal from communications takes precedence over the noise.'

At the Bank, this multi-faceted nature of the forecast process has helped to shape the institution's long-standing approach to formulating its economic projections. And although the Bank has used a number of differing forecast models and techniques over the past two decades, its approach has been rooted in a small number of core insights. Specifically:

- the importance of evaluating the risks around a central view, as illustrated by the Bank's 'fan charts' (the visual representation of the probabilistic distribution of outcomes for GDP, inflation, and, since August 2013, unemployment);

- that there is no one 'best' model to use, and that a suite of models, using a variety of differing approaches, provides a better way to gauge the likely economic outlook and the risks; and

- that, while statistical models help to inform the policy process, policymakers are ultimately required to use their judgement when making economic assessments. As George (1999) sets out: 'The Bank's use of economic models is pragmatic and pluralist. In an ever-changing economy, no single model can possibly assimilate in a comprehensible way all the factors that matter for policy. Forming judgements about those factors, and their

implications for policy, is the job of the [Monetary Policy] Committee, not something that can be abdicated to models or even to modellers. But economic models are indispensable tools in that process.'

This multi-faceted nature of central bank forecasting, and its importance to the policymaking process, provides context for the 2012 decision by the Bank's Court of Directors to commission the Stockton Review into the MPC's forecasting capability.(1) The Bank subsequently made a series of changes to its forecasting process and practices,(2) including a variety of changes designed to facilitate greater consideration of alternative points of view. Broader changes to the transparency and accountability of monetary policy followed the Warsh Review (2014).(3)

1.2 The forecasting approach at the Bank

The Bank publishes economic projections in its *Inflation Report* four times a year (February, May, August and November). In the six weeks or so leading up to the publication of the *Report*, a structured set of discussions between Bank staff and the Monetary Policy Committee — the 'forecast round' — helps the MPC to settle on its projections for core macroeconomic variables. For more on the structure of a typical forecast round, see the box on page 10.

At the end of the forecast round, the MPC settles on its projections for GDP growth, CPI inflation (the MPC's target variable) and the LFS unemployment rate. These projections are published in the *Inflation Report* in the form of a fan chart; the parameters of the underlying probability distribution (the central tendency and the skew) are also made publicly available. The fans represent the 'best collective judgement' of the MPC, although not every member will necessarily agree on every assumption upon which the projections are based.

The projections for GDP growth, CPI inflation and unemployment are the only ones that are owned and explicitly signed off by the MPC; for unemployment, this has only been the case since August 2013. However, for many years, Bank staff have produced indicative projections for a wide range of other macroeconomic variables that are consistent with the MPC's modal projections for GDP growth, inflation and, since August 2013, the unemployment rate. The degree of scrutiny and attention that these indicative projections receive — both from the MPC, and from Bank staff — varies over time, and not all MPC members will agree with the staff's indicative projections. Section 2 returns to this issue in the context of data quality.

A notable feature of the Bank's forecasting process is the use of 'conditioning assumptions' for Bank Rate. Typically, two different conditioning assumptions for Bank Rate are used in the *Inflation Report* projections. One assumes that Bank Rate rises in line with the market yield curve; a second assumes that Bank Rate remains constant. The MPC publishes projections for GDP growth, CPI inflation and the unemployment rate under both these 'market rate' and 'constant rate' conditioning assumptions. A range of other conditioning assumptions for the exchange rate, energy prices and fiscal policy are also used when forming Bank forecast projections. Section 2 returns to the issue of conditionality in the context of interpretation of our empirical results.

The 2012 Financial Services Act materially expanded the Bank's statutory duties, with the institution gaining significant new responsibilities. These included the microprudential regulation of insurers, deposit-takers and major investment firms; macroprudential regulation of the financial system as a whole; and supervision of some critical post-trade financial market infrastructure providers.(4)

This expansion of the Bank's duties, together with its emphasis — as set out in its 2014 Strategic Plan(5) — of working as a single institution across the full range of its policy responsibilities, has led to its macroeconomic forecasts being used for a wider range of purposes than pure monetary policy making. For example, the insights

(1) This Review was conducted by David Stockton, former Director of Research and Statistics at the Board of Governors of the US Federal Reserve System.
(2) See www.bankofengland.co.uk/publications/Documents/news/2013/nr051_courtreviews.pdf for the Bank of England's response to the 2012 Court Reviews. An overview of subsequent changes to the Bank's *Inflation Report* is provided in McKeown and Paterson (2014).
(3) Warsh (2014) made a series of recommendations to improve transparency and accountability of monetary policy, including: providing more, and more timely, information about the monetary policy decision; releasing transcripts of that part of the MPC meeting where policy is decided with a deferral period of between five and ten years; releasing key staff briefing materials alongside the transcripts. The Bank's response to the Warsh Review can be found at www.bankofengland.co.uk/publications/Documents/news/2014/warshresponse.pdf.
(4) See Murphy and Senior (2013).
(5) See Carney (2014).

Box 1
A typical quarterly forecast round at the Bank

The typical quarterly forecast round at the Bank lasts around six weeks, and largely takes the form of structured discussions between the Monetary Policy Committee (MPC) and staff from the Monetary Analysis and International areas of the Bank. It is primarily focused on a number of set piece forecast meetings, where Bank staff typically make a relatively short presentation to MPC members that highlights key issues for debate, and then the Governor will chair the subsequent discussion. The MPC formally signs off the probability distributions for GDP growth, CPI inflation and the LFS unemployment rate shortly before publication; these represent the 'best collective judgement' of the Committee.

The main milestones in a typical forecast round are as follows:

- **Publication of Quarterly National Accounts (QNA) data** by the Office for National Statistics (typically around five weeks before forecast publication in the *Inflation Report*). Staff preparation for the quarterly round typically begins in earnest following publication of QNA data, the most detailed quarterly release of GDP expenditure components and income. Following QNA release, staff process the data that has been released since the Committee's previous forecast, revisit past Committee judgements, and discuss potential new judgements for the forecast.

- **The Benchmark meeting** (typically around three weeks before forecast publication). Staff will present their first formal cut of the new forecast to the MPC, highlighting notable data released in the preceding weeks and any proposed changes in judgement. The Benchmark meeting typically sees a detailed presentation of the latest international forecast and preliminary fan charts for UK GDP growth and inflation. Although there will have normally been a series of bilateral meetings, and informal discussions, between MPC members and Bank staff in the run-up to the meeting, the Benchmark provides the first formal opportunity for in-depth forecast discussions with the Committee as a whole.

- **The Key Issues meetings** (typically around two weeks before forecast publication). There are typically two Key Issue meetings in any forecast round, and these provide an opportunity for a detailed assessment of core aspects of the forecast. The content of the Key Issues meetings will vary from round to round, but will typically focus on some of the principal risks to the outlook — for example, one meeting may focus on aspects of the GDP projection (eg the extent to which it is dependent on a revival in credit growth), and another may focus on aspects of the inflation projection (eg the nature of the assumptions about pass-through of changes to relative prices such as energy). For some forecast rounds, some or all of one of the Key Issues meetings will be devoted to questions of monetary policy strategy.

- **The Draft meetings** (typically in the week or so leading up to the policy decision). Following the Key Issues meetings, staff will revise the forecast projections, taking on board Committee discussions and requests, as well as any new data received since the Benchmark. These revised projections are then presented to the MPC for discussion. The number of Draft meetings varies from round to round, depending in part on the time it takes for the MPC to settle on key judgements underpinning the forecast. Two Draft meetings would not be uncommon.

- **Sign-off of the final forecast**. The MPC formally signs off the fan charts, and the associated probability distributions, for GDP growth, CPI inflation and the LFS unemployment rate shortly before their publication in the *Inflation Report*. These projections represent the 'best collective judgement' of the Committee, although not all members will agree with every assumption on which the projections are based. As part of the forecast process, the staff will produce indicative projections for a wide range of other macroeconomic variables that are consistent with the MPC's modal projections for GDP growth, CPI inflation, and the LFS unemployment rate. These are seen by the MPC, but not formally signed off, and the attention they receive from the Committee will vary from round to round.

The forecast process is set out in more detail in Bean and Jenkinson (2001) and Stockton (2012).

provided by the Bank's forecasts are used to inform the design of the institution's stress tests for the banking sector; the annual stress-testing exercise is, in turn, an important input to the work of the Bank's Financial Policy Committee (FPC) and the Prudential Regulation Authority Board (PRAB). One purpose of the IEO's work programme has therefore been to provide Court with a framework to assure itself of the quality and governance of the macroeconomic projections used by policy committees in the Bank other than the MPC.

The Bank's current forecasting set-up uses a central organising model (known as COMPASS — the Central Organising Model for Projection Analysis and Scenario Simulation), supplemented by a 'suite' of alternative models and tools used to cross-check and adjust the forecast, particularly in areas where the central model is thought to be lacking.

COMPASS began to support the MPC's forecasting process in November 2011, and replaced the Bank's previous central organising model, known as BEQM (the Bank of England Quarterly Model), which had been in use since 2003. COMPASS is a 'New Keynesian' general equilibrium model, and shares many features with similar models in use at other central banks and policy institutions. Box 2 summarises key features of COMPASS and the model suite; a fuller description is provided in Burgess *et al* (2013).

1.3 Published studies of Bank forecast performance

The Bank has for many years published an annual account of its forecasting record in its *Inflation Report*. These accounts have typically looked back over the period of a year or so, compared data outturns with MPC expectations, and provided some economic narrative to explain the degree to which the central expectations of the MPC had been borne out. These shorter-run accounts of MPC forecasting have been supplemented by a number of longer-run, more in-depth studies — both by Bank staff, and by external commentators.

Box 3 on pages 13–14 summarises the findings of selected longer-run studies of Bank forecast performance. These papers focus on GDP and inflation projections published externally by the Bank on its website; this IEO paper has additionally benefited from the use of internally held data to analyse forecast performance for a wider range of macroeconomic variables.

In common with our findings, previous studies find few signs of systematic bias in Bank GDP and inflation projections (although Hackworth *et al* (2013) observe evidence of bias in one year ahead projections at the 10% significance level), and mixed evidence of inefficiency. A number of studies (Fawcett *et al* (2015), Groen *et al* (2009), Clements (2004)) note that, for at least some variables, pure statistical models have a tendency to outperform Bank forecasts at longer horizons. The most recent studies (Stockton (2012), Hackworth *et al* (2013)) note the large forecast errors made by the Bank and external forecasters since the financial crisis; Stockton observes that the Bank's performance was 'marginally worse' than those of outside forecasters. In terms of formal fan chart evaluation, both Clements (2004) and Wallis (2004) find evidence that, historically, Bank inflation fan charts tended to be too wide; both these studies relate to the pre-crisis era, however.

Box 2
The Bank's forecasting platform — COMPASS and the suite of models

The forecast process at the Bank is informed by a range of economic models. Since November 2011, the Bank's forecasting platform has consisted of a central organising model (the Central Organising Model for Projection Analysis and Scenario Simulation, or COMPASS), a suite of around 50 separate models to supplement and cross-check the projections provided by COMPASS, and a range of IT tools to assist the forecasting process. Details of the Bank's forecast platform are provided in Burgess *et al* (2013).

COMPASS was designed to fulfil three purposes: it is the main organising framework for the construction of the forecast; it is used to analyse and explain the forecast projections; and it is used to assess the sensitivity of the forecast to alternative assumptions. COMPASS provides the basic set of relationships that articulate core macroeconomic mechanisms, and provides a disciplining framework by ensuring that forecasts are internally consistent. It is a relatively small model, generating forecasts for around 20 variables, including GDP, inflation, interest rates, trade, wages and consumer spending.

COMPASS is an open economy, New Keynesian, Dynamic Stochastic General Equilibrium (DSGE) model, estimated on UK data using Bayesian methods. It shares many features with similar models at other central banks. Wages and prices are assumed to be sticky, and so monetary policy can influence real variables such as output and employment over short to medium horizons, but not the long run. Expectations of monetary policy are an important determinant of current output and inflation.

While COMPASS acts as the central organising model for the MPC's forecast projections, Bank staff also have access to a range of alternative models, known as the 'statistical suite'. Models in the suite encompass a wide range of different frameworks and techniques. They typically fall into one of three broad categories:

- Models which articulate economic shocks and channels that are omitted from COMPASS. These include models that explicitly account for energy as an input to production and consumption and models that have a formal role for a financial sector intermediating funds between different sectors of the economy.

- Models which produce forecasts for additional variables not included in COMPASS. An example of this is the Bank's Balance Sheet Model (BSM), a backward-looking, recursive, non-linear model which generates forecasts for around 140 variables. The BSM focuses on money and banking-related variables such as components of household and corporate lending growth.

- Models that generate alternative forecasts for variables which are in COMPASS; these play a role in validating and adjusting the output from the central organising model. One set of models in this category are statistical forecasting models that typically employ simple econometric relationships such as Error Correction Models.

Some models in the suite (the 'inner' suite) would tend to always play a role in MPC forecast rounds; other models (the 'outer' suite) are typically only used on an occasional basis, and tend to have less resources devoted to their maintenance. As set out in Burgess *et al* (2013), formal economic models such as COMPASS and the wider model suite are an important part in the forecast process, but policymakers' judgement still continues to play a core role. 'Key elements of the forecasting process at the Bank of England include the ownership of the final forecast by the MPC, their intensive involvement in the production of the forecasts and the fact that the published forecasts in the *Inflation Report* are judgemental.'

Box 3
Selected studies of Bank forecast performance

Studies by Bank of England staff

Fawcett, Körber, Masolo and Waldron (2015) investigate the real-time performance of the Bank's main DSGE model, COMPASS, before, during and after the financial crisis with reference to statistical and judgemental benchmarks. Specifically, the forecast performance of a relatively 'judgement-free' version of COMPASS is evaluated against the performance of the judgemental forecasts made by the MPC and published in the *Inflation Report*, as well as against the performance of a statistical benchmark forecast from the Bank's Suite of Statistical Models.

The authors find that, at shorter horizons, the MPC's *Inflation Report* projections (both point forecasts and complete probability density forecasts) are more accurate for both GDP growth and inflation than either the COMPASS forecasts or the forecasts produced by the Statistical Suite. At longer horizons (horizons of more than one year), COMPASS has the more accurate inflation point forecasts, and the Statistical Suite the more accurate GDP point forecasts. For probability density forecasts, the MPC's *Inflation Report* projections are more accurate at shorter horizons than those produced by other methods, but not at longer ones. The authors note that not all these differences are statistically significant, and forecast accuracy in itself is not the only metric by which models should be assessed.

An in-depth evaluation of Bank forecasts for GDP growth and CPI inflation is provided by **Hackworth, Radia and Roberts (2013)**. The authors provide a detailed comparison of economic outturns over the 2010–13 period with MPC central expectations made in August 2010. The unexpected weakness in GDP growth over that period was largely attributed to weaker growth in UK trading partners, tighter domestic credit conditions, and the slower dissipation of uncertainty. The unexpected strength of CPI inflation was attributed to unanticipated rises in energy and other imported costs; weak effective supply was additionally thought likely to have counteracted the impact on inflation of weak demand.

Hackworth *et al* provide statistical diagnostics similar to those detailed in this paper on the performance of MPC mean projections for GDP growth and inflation over 1997–2013; results are presented for one quarter ahead and one year ahead projections. For GDP growth, mean projections were found to be unbiased at the 5% level at both horizons, although there was some evidence of bias at the 10% significance level for one year ahead GDP growth (MPC mean projections too high), as well as some evidence of strong inefficiency in one quarter ahead projections. For inflation, mean projections were found to be unbiased at the 5% level at both horizons, although there was some evidence of bias at the 10% significance level for one year ahead inflation (MPC mean projections too low), as well as some statistically significant evidence of weak inefficiency at both horizons.

Groen, Kapetanios and Price (2009)(1) evaluate the performance of central estimates for inflation and growth contained in the Bank's *Inflation Report* for the 1997 Q3 to 2006 Q2 period. Specifically, Groen *et al* compares RMSEs of Bank forecasts with those of pure statistical models, using the Diebold-Mariano-West test of significance. Groen *et al* conclude that while GDP forecasts produced by statistical models perform as well or better than Bank forecasts at all horizons, the reverse was typically true of Bank inflation forecasts. 'For inflation, the *Inflation Report* forecasts are clearly dominant, often significantly so.'

A formal evaluation of the Bank's fan charts — rather than point estimates of GDP and inflation — is provided by **Elder, Kapetanios, Taylor and Yates (2005)**. Elder *et al* conducted a variety of formal and informal tests of the Bank's fan charts for GDP growth and RPIX inflation.(2) They found that, at most horizons, outturns for GDP growth and RPIX inflation had been distributed broadly in line with the MPC's fan chart bands. There were some signs that, over the sample period under consideration (forecasts published between February 1998 and May 2003), GDP fan charts had been too narrow at shorter horizons. At the two-year point, outturns for both GDP growth and inflation had typically been more narrowly dispersed than implied by the fan charts, although

(1) Simon Price is a senior member of Bank of England staff. Jan Groen is an official at the Federal Reserve Bank of New York, and George Kapetanios is Professor at Queen Mary University of London.
(2) RPIX inflation was the target variable for the Bank's MPC until December 2003, at which point the target was redefined in terms of CPI inflation.

this was not seen as evidence that the fan charts had been drawn too widely, in part because the authors only had a small sample — six years' worth of data — with which to work.

Elder *et al* also evaluated mean point projections for growth and inflation against outturns. The authors noted that their sample was too small to draw firm conclusions about bias, but observed a tendency for RPIX inflation to be lower than predicted in the earlier part of the forecast period (which the authors associated with the unexpected strength of the exchange rate at that time), and higher than predicted in the later part of the forecast period (associated with unexpected strength in the housing market). Compared with external forecasters, MPC average absolute errors had tended to be a little smaller for GDP growth, and broadly the same for inflation, although the authors cautioned against reading too much into this observation given the small sample size.

Forecast evaluations by external commentators

The remit for **Stockton (2012)** included an analysis of MPC forecast performance since the financial crisis. Stockton provided: an informal comparison of the distribution of outturns for GDP growth and inflation to that implied by the fan charts; a comparison of errors (RMSEs, average absolute errors, and average errors) in MPC central estimates to errors made by external forecasters; and, a comparison of outturns for euro-area/US GDP growth with European Central Bank/Federal Open Market Committee forecasts. Stockton noted that, since the financial crisis, MPC forecast errors had been characterised by 'persistent over-prediction of output growth and persistent under-prediction of CPI inflation'. He concluded that the performance of MPC forecasts had been 'noticeably worse than prior to the crisis and marginally worse than that of the average of outside forecasters'.

Stockton found the narrative provided by the MPC for these errors — for growth, a larger and more persistent impact of the financial crisis, and, for inflation, unexpectedly high upward pressure from external price shocks — to be 'in broad terms, … persuasive'. But he additionally observed that this narrative 'may not fully explain' the serial persistence of post-crisis errors. 'That persistence could simply reflect a string of bad luck, but it also could reflect some inertia imparted by the forecast process, or point to problems with the paradigm underlying the Bank's forecasts.'

Clements (2004) conducted a formal evaluation of the Bank's inflation fan charts from the 1997 to 2003 period. The paper also includes evaluation of the central tendencies of the inflation fans. The MPC's point forecasts were found to be largely unbiased; they additionally outperformed naïve statistical benchmarks at shorter horizons. At longer horizons (one year ahead), naïve benchmarks tended to outperform MPC point estimates, and the fan charts appeared to overestimate the upside risks to inflation. Clements concludes: 'The year ahead MPC forecasts fare poorly however evaluated.'

Formal evaluations of the Bank's inflation fan charts are also provided by **Wallis (2004)**. Wallis evaluates the Bank's fan charts for inflation in the 1997 to 2003 period against the distribution of outturns, and conducts an equivalent exercise for inflation projections provided by the National Institute of Economic and Social Research (NIESR). Wallis looks at one quarter ahead and one year ahead projections. For both the Bank and NIESR, central point forecasts were found to be unbiased, but their density forecasts were found to substantially overstate uncertainty.

2 Data, methodology and context

This section first describes the data and methods that we used to assess Bank forecasts and to compare them to forecasts from UK private sector forecasters and other central banks. Further detail is available in Annex 3. The section then discusses some key pieces of context that should be borne in mind when interpreting the empirical results in Section 3, with sample size being particularly important.

2.1 Data

Past published evaluations of Bank forecasting performance have focused on GDP and inflation, the two macroeconomic variables for which the Bank has published projections since the inception of the MPC (see Box 3 on pages 13–14 for a summary of previous evaluations). This evaluation has benefited from being able additionally to draw upon internally held, unpublished, Bank projections, expanding the set of forecasts assessed to eleven macroeconomic variables: real GDP, inflation, the unemployment rate, real private consumption, real total investment, nominal wages, nominal house prices, nominal household lending, nominal corporate lending, US GDP (real) and euro-area GDP (real). This subsection describes key features of the Bank's forecast data, before turning to the data on external forecasts that were used to benchmark the Bank's forecast performance.

Bank forecast data

Table 2.A provides more information on Bank forecast variables used in our evaluation exercise. As set out in Section 1.2, the MPC only formally signs off on projections for GDP growth, CPI inflation and the unemployment rate. These are additionally the only variables for which full probability distributions (expressed visually as 'fan charts' in the *Inflation Report*) are produced. The other variables considered in this report are indicative projections made by Bank staff to be consistent with the modal path agreed by the MPC for GDP growth, inflation and (since August 2013) the unemployment rate.

Table 2.A Summary of Bank forecast variables evaluated in this report

Variable[a]	Forecasts available from:	Forecasts published?	Ownership of forecasts[b]	Forecast distribution produced?
Real GDP growth	August 1997	Yes	MPC	Yes, including modal and mean projections
Inflation[c]				
Unemployment rate[d]	May 1998	Yes (since August 2013)	MPC (since August 2013)	Yes, including modal and mean projections (since August 2013)
Real private consumption growth	August 1997	Yes (calendar-year basis since February 2014)	Bank staff	No — only modal projections available
Real total investment growth				
Nominal wage growth				
Nominal house price growth	August 1997	No		
Nominal household lending growth	February 2007			
Nominal corporate lending growth	February 2009			
Real US GDP growth	February 2001	Yes (calendar-year basis since February 2014)		
Real euro-area GDP growth				

(a) Unless stated otherwise, variables are domestic (eg real GDP growth is UK real GDP growth).
(b) Some forecast variables are explicitly and collectively agreed by the MPC every quarter. Other forecasts are produced by Bank staff to be consistent with these MPC-owned forecasts and to be consistent with MPC judgements, which can be made on any forecast variable.
(c) The Bank's target variable was changed from RPIX inflation to CPI inflation in December 2003.
(d) For most of the period since the MPC's inception, the Bank has forecast the unemployment rate as measured by the Labour Force Survey. A small number of forecasts made prior to May 1998 used the claimant count measure. These have been excluded from our sample.

As noted in Section 1.2, the degree of scrutiny afforded to these modal staff projections — both by Bank staff, and by the MPC — will vary over time. Generally speaking, staff projections made in the early part of the sample period will typically have benefited from less scrutiny (both by staff themselves, and by the MPC), than those made in the later part of the sample period. And even latterly, the degree of attention given to any individual staff profile will vary from round to round, depending on the economic issues under consideration.(1)

As discussed further in Section 2.3, forecasts are inherently uncertain, and, ideally therefore, our evaluation exercise would focus on the full probability distributions for forecast variables, rather than on point estimates of central tendency. However, as **Table 2.A** sets out, we only have data on full probability distributions for a sufficiently long sample period for GDP growth and inflation. While we evaluated these probability distributions using standard techniques (see Box 4 on pages 50–51), most of our analysis was, by necessity, focused on modal forecasts. For GDP and inflation, we were able to repeat our main diagnostic tests using mean forecasts (which take into account the balance of risks to the modal projection),(2) as well as modal forecasts. Our key findings were robust to using modal or mean projections.

Throughout this paper, we use the following conventions when assessing and describing forecasts:

- We assessed Bank forecasts at three horizons: one quarter ahead, one year ahead and two years ahead.(3) Forecast horizons are defined with respect to data availability not calendar time. For example, a one quarter ahead forecast for the UK GDP growth outturn in 2014 Q2 would have been published in the May 2014 *Inflation Report*, because at that time published UK GDP data were only available up until 2014 Q1.

- Similarly, when using calendar-year forecasts (in comparisons with external forecasters, described below), we defined one year ahead forecasts as those made when the data for Q4 of the preceding year were available. For example, a one year ahead forecast for the inflation outturn in 2014 would have been made in February 2014 (when inflation data for 2013 Q4 were available).

- Unless otherwise stated, dates are defined with respect to data outturns, not forecasts. So 2008 refers to forecasts made *for* 2008 and not necessarily to forecasts made *in* 2008.

- Finally, throughout this paper, we define forecast 'error' as outturn less forecast. So a positive forecast error, for example, means that the forecast was below the eventual data outturn.

For our standard statistical tests (for bias, inefficiency, and so on), our main results exclude the worst years of the recent global financial crisis (2008 and 2009). This is because we used these tests to evaluate evidence of any systematic errors that were made in more 'normal' times, and thereby to identify scope for process improvement. Including the large forecast errors made by the Bank at the height of the financial crisis had the potential to skew our diagnostic results, and hence our recommendations. Results covering the full sample period are included in Annex 4 for reference; generally speaking, however, the inclusion of the 2008–09 period made relatively little difference to our main results. We do include 2008 and 2009 when comparing the forecast performance of the Bank to that of other institutions in our benchmarking exercise, however, as it is relevant to ask how Bank forecasts compared to those of others at both the peak of the crisis, and in more normal times.

Forecasts are assessed against published data.(4) For all of the variables that we assess in this report, except for inflation, data outturns are subject to revision over time. This is particularly true for GDP, consumption and investment. For these variables, the majority of the analysis was conducted using the 'latest' vintage of data, defined as the published history of the variable that was available at the time of our analysis. That is consistent

(1) Archiving and data storage processes for Bank forecast information were also not as robust in the early part of the sample period, particularly pre-2001. As part of our wider robustness checking, therefore, we reran our main diagnostic tests excluding data from the pre-2001 period. Our findings were broadly unchanged, the main exception being the unemployment rate, where excluding the relatively large forecast errors in the early part of the sample did affect some of our results (see Sections 3 and 4 for further discussion).
(2) For a discussion of the concepts of mean and mode in relation to the Bank's fan charts, see Elder *et al* (2005). For a description of the fan charts, and what they represent, see the box on pages 48–49 of the May 2002 *Inflation Report*.
(3) The exceptions are the household and corporate lending growth forecasts, where we did not assess two year ahead forecasts because the sample of available forecasts was very short, only starting in the late 2000s.
(4) Much of the analysis in this report was done shortly after the publication of the Bank's November 2014 *Inflation Report*, meaning that data were available to 2014 Q2 or Q3, depending on the variable. We subsequently performed robustness checks, which confirmed that our key results were largely robust to extending the data sample to 2014 Q4. For calendar-year forecasts, we extended our main analysis to include 2014 once data became available.

with the stated aim of the MPC to forecast 'mature' vintages of UK GDP, which have been reviewed and revised by the UK Office for National Statistics over a number of years.(1) But it also means that we assessed forecasts using a data history that may look different to the history that was available to the forecasters at the time they made their forecasts.

Given these data revisions, we performed robustness checks using real-time data (defined as the third release for a particular quarter of data) when assessing the performance of the Bank's forecasts of growth in UK GDP, US GDP and euro-area GDP. The key impact on our results of using real-time data was that forecasts of UK GDP growth appeared to have been significantly too high; a result that we did not find when using the latest data vintage. We discuss in Section 3.2 how this likely reflects the Bank's emphasis on anticipating data revisions to early estimates when forecasting.

Data on external forecasts

As part of our benchmarking exercise, we compared Bank forecasts of domestic macroeconomic variables directly to those from UK private sector forecasters. Our main source of private sector forecasts were those reported to the international surveys run by Consensus Economics. These surveys were chosen because they are consistent across economies and available over our whole sample period. The private sector forecasts for the variables evaluated were only available on a calendar-year basis.

We compared these private sector forecasts to Bank forecasts published in the same month.(2) We used individual forecasts reported to these surveys, rather than the 'consensus' or average forecast. We restricted our sample to include only those forecasters that: (i) reported forecasts in at least 10 of the 17 years from 1998 to 2014, and (ii) reported forecasts for both 2008 and 2009. We imposed these restrictions to ensure that our sample only included regular forecasters, and because 2008 and 2009 were objectively difficult years to predict, so the Bank's forecast accuracy could appear unfairly low if we compared to forecasters who did not submit forecasts for those years. Our restricted sample included between 17 and 22 individual forecasters, depending on the variable.

Consensus surveys do not include projections for house prices. We therefore evaluated Bank projections for house price inflation using forecasts provided by the HM Treasury survey of external forecasters for the United Kingdom. Given the short sample period of house price forecasts available (2009 to 2014), we set a threshold of four forecasts from six years for inclusion in our sample, and we only assessed house price forecasts that were based on either the Halifax or Nationwide house price indices. Six forecasters met these criteria.(3)

We also compared Bank forecasting performance to published forecasts from four other central banks. Our sample included the European Central Bank (ECB) Staff Macroeconomic Projections and the US Federal Reserve (Fed) Monetary Policy Report forecasts(4) as these are the world's most prominent central banks. We also included the Reserve Bank of New Zealand (RBNZ) Monetary Policy Statement forecasts and the Swedish Riksbank Monetary Policy Report forecasts, as these central banks were identified in the 2014 Warsh Review as being more transparent than the Bank in terms of the information that they publish.(5) This allowed us to benchmark the performance of a wider range of Bank forecasts than would otherwise have been the case, since we only included published forecasts from other central banks in our comparisons.

Table 2.B summarises the five forecast variables included in the international comparison. For the Fed, RBNZ and Riksbank, forecasts are generally available since 1998, while ECB forecasts have been published since December 2000. For the most part, these central banks publish calendar-year forecasts. The exception is the Fed, which publishes forecasts for Q4 only. We were able to evaluate Fed forecasts against Bank forecasts on a comparable basis using unpublished Bank forecasts. But the comparison of Fed forecasts with private sector forecasts, which are all on a calendar-year basis, was more problematic. Since we would expect forecasting calendar-year averages

(1) See, for example, the box on page 39 of the November 2007 *Inflation Report*.
(2) Although we know the month in which the surveys of private sector forecasters were published, we do not know how up-to-date the forecasts submitted to those surveys were. It is therefore possible that the private sector forecasts may have been less up-to-date than the corresponding Bank forecasts, which would bias our results towards being more favourable to the Bank relative to the private sector forecasters.
(3) Also because of the short sample of private sector forecasts available, we did not assess two year ahead house price growth forecasts in this part of the evaluation.
(4) For each Monetary Policy Report, there is a collection of forecasts made by Federal Open Market Committee participants. As a simplifying assumption, we used mid-points of central tendency ranges in our analysis; further details are provided in Annex 3.
(5) See the comparison of economic transparency at central banks provided in Warsh (2014), page 31.

Table 2.B Summary of forecast variables included in comparison across central banks

Variable	Bank of England	European Central Bank	Reserve Bank of New Zealand	Swedish Riksbank	US Federal Reserve
Real GDP	✓	✓	✓	✓	✓
Inflation	✓	✓	✓	✓	✓
Unemployment	✓	Only published by ECB since December 2013.	✓	Published by Riksbank, but not included in Consensus survey.	✓
Private consumption	✓	✓	✓	✓	Not published by Fed.
Total investment	✓	✓	✓	✓	Not published by Fed.

to be easier than forecasting a particular quarter, our comparison of the Fed with private sector forecasters probably biases the evaluation against the Fed. Nevertheless, we believe that there is useful information in including the Fed in our analysis; see our empirical results for further details.

As well as comparing forecasting performance directly across central banks we also compared central bank forecasts to corresponding forecasts made by private sector institutions, again using individual forecasts submitted to the Consensus Economics surveys.(1) Comparing central bank forecasting performance to that of other forecasters in the same country provides a way to control for structural differences in national economies or differences in historical economic shocks; see Section 2.2 for more details. **Table 2.C** provides further detail of the private sector forecasts used in our evaluation exercise.

Table 2.C Summary of private sector forecast variables included in international comparisons of forecast accuracy

Variable	Sample period used in evaluation(a)	Number of forecasters included in sample(b)				
		UK	US	EA	NZ	SW
Forecasts provided by Consensus Economics						
Real GDP growth	1998–2014 (one year ahead forecasts) 1999–2014 (two year ahead forecasts)(c)	21–22	17–19	19–21	11–12	7–9
Inflation		20–21	17–19	18–21	11	11
Unemployment rate(d)		17	17–19	n.a.	10–11	n.a.
Real private consumption growth		19–20	n.a.	19–21	11–12	8–9
Real total investment growth		17–19	n.a.	19–21	11–12	8–9
Forecasts provided by HM Treasury						
Nominal house price growth	2009–14 (one year ahead forecasts only)	6	n.a.	n.a.	n.a.	n.a.

(a) The sample period is shorter for the EA forecast variables (2003–14 for one year ahead forecasts and 2004–14 for two year ahead forecasts), and for NZ real private consumption and total investment growth (2000–14 for one year ahead forecasts and 2001–14 for two year ahead forecasts).
(b) Variables which are not included in evaluation are marked with 'n.a.'; for example private consumption growth forecasts are not published by the Fed.
(c) Sample size can vary depending on the forecast horizon (one year or two year ahead), as some forecasters did not submit forecasts in every year.
(d) For most of the period since the MPC's inception, the Bank has forecast the unemployment rate as measured by the Labour Force Survey. A small number of forecasts made prior to May 1998 used the claimant count measure. These have been excluded from our sample.

2.2 Methodology

This section describes how we assessed Bank forecasting performance in four main ways:

- How closely forecasts have matched what actually happened (accuracy).
- Whether forecasts have been consistently too optimistic or too gloomy (unbiasedness).
- Whether forecasts have reflected all the information available at the time they were made (efficiency).
- Whether forecasts have been more or less accurate than forecasts from a simple model, from UK private sector institutions, and from other central banks (benchmarking exercise).

(1) We applied the same threshold restrictions as we applied to our sample of UK private sector forecasts, except for the euro area where the threshold was eight forecasts from fourteen years.

We recognise that accuracy, unbiasedness and efficiency are not independent of each other. If forecasts are consistently too optimistic or gloomy then they will typically be less accurate than if they were unbiased. Similarly, forecasts could be more accurate if they made more efficient use of available information. But it is helpful to look at forecast performance in different ways to try to identify scope for process improvement.

2.2.1 Accuracy

We measured the accuracy of all forecasts using root mean squared forecast errors (RMSE), defined as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} e_i^2} \quad (1)$$

where e is a forecast error, defined as outturn less forecast.

If forecast errors are larger, then RMSEs are also larger. The relationship is not linear, however. RMSEs will be disproportionately large (small) if errors are very large (small). RMSEs are a standard loss function used in the forecast evaluation literature. More importantly, a quadratic loss function is appropriate for our purposes because policymakers will care more about big forecast errors that could lead to big policy mistakes and have a damaging impact on the economy, than small errors, which may not impact much on policy. We recognise, however, that RMSEs represent one particular form of loss function — a quadratic loss function — and that alternative loss functions could lead to different results. We also recognise that different loss functions may be optimal for different measures of central tendency. For example, Gneiting (2011) recommends that RMSEs are used to evaluate mean forecasts rather than the modal forecast. But as noted in Section 2.1, data constraints mean that most of our analysis is necessarily focused on modal forecasts; forecasts of mean outturns are only available over a sufficiently long period for UK GDP growth and inflation. For these variables, our key findings were robust to using modal or mean projections.

We also note that RMSE is not independent of other metrics used in this evaluation. In particular,

$$RMSE = \sqrt{bias^2 + forecast\ error\ variance} \quad (2)$$

To allow comparability of RMSEs across variables, we scaled RMSEs by the standard deviation of the data outturns (over the same period as the RMSE was calculated) (unscaled RMSEs are reported in Annex 4). This is a crude way to account for data volatility, since greater volatility makes a given variable inherently more difficult to forecast. It is important to note that the value of the scaled RMSE is not itself informative (ie whether it is above or below one), and is only useful to compare with scaled RMSEs of other variables or of the same variable in other periods of time. By scaling by the standard deviation of realised outturns we are essentially comparing forecast accuracy to that of the (*ex post*) unconditional mean of the outturns. But this is not a viable benchmark to judge the Bank's forecasts against as the *ex-post* mean would not have been known at the time the forecast was made. An alternative would be to scale by the recursively computed (*ex ante*) unconditional mean — in which case a value of RMSE higher (or lower) than one would indicate that forecast performance is worse (or better) than such a simple benchmark.

2.2.2 Unbiasedness

Unbiasedness of Bank forecasts was assessed using ordinary least squares (OLS) regression.(1) Forecast errors were regressed on a constant with a null hypothesis that the constant was zero, which would be the case if the forecasts were unbiased. Otherwise, the forecasts could have been made more accurate by adding a constant amount to them. As in equation **(1)**, we defined forecast errors as:

$$e_t^{t-h} = y_t - y_t^{t-h} \quad (3)$$

where y_t is the outturn of variable y in period t and y_t^{t-h} is the forecast for variable y in period t made in period t-h.

(1) Consistent with the use of RMSE to measure accuracy, all of our statistical tests use a quadratic loss function as they are estimated using least squares, minimising the sum of squared residuals.

To test for unbiasedness, we estimated the following regression:

$$e_t^{t-h} = \beta_0 + u_t \quad (4)$$

where u_t is a zero-mean error term. Under the null hypothesis of unbiasedness $\beta_0 = 0$. If $\beta_0 > 0$, forecasts have been systematically too low. If $\beta_0 < 0$, forecasts have been too high. We estimated the regression using OLS with Heteroscedasticity and Autocorrelation Consistent (HAC) standard errors.(1)

We used HAC standard errors for all of the statistical tests in this report. That is particularly important for forecasts of growth rates at horizons longer than one quarter. By construction, such forecasts made in consecutive quarters will cover mostly the same forecast period and the associated forecast errors are therefore likely to be autocorrelated. Using HAC standard errors should account for this and other potential autocorrelation and heteroscedasticity issues.

2.2.3 Efficiency

To assess the efficiency of Bank forecasts we conducted a number of tests. 'Strong' efficiency was assessed by regressing forecast errors on information that was known when the forecasts were made. Ideally, forecast errors will be uncorrelated with any such information. Otherwise, those errors could have been reduced by incorporating that information when the forecasts were made.

To perform tests of strong efficiency, we estimated the following equation using OLS with HAC standard errors:

$$e_t^{t-h} = \beta_0 + \beta_1 z_{t-h} + u_t \quad (5)$$

where e is a forecast error, as defined in equation **(3)**, z_{t-h} is a variable that was known to the forecaster at time t-h, and u_t is a zero-mean error term. Under the null hypothesis of strong efficiency we test whether $\beta_1 = 0$ since unbiasedness, defined as $\beta_0 = 0$, is tested in a separate regression (see equation **(4)**).

For each of the variables in our sample, we performed strong efficiency tests using the following for z: the previous forecast error (for the same variable) known to the forecasters at t-h; and the previous data outturn (for the same variable), also known at t-h. We tested each z variable separately. For these tests, we used real-time data, to be consistent with the information set available to forecasters at t-h. For forecasts of UK GDP growth, we additionally performed the test using UK-weighted world GDP growth as z to assess whether international influences were sufficiently captured in the Bank's UK GDP forecasts.(2) We conducted the test separately using both the latest known outturn of world GDP growth and using the contemporaneous Bank staff forecast of world GDP growth (ie the forecast for time t).

Other tests for efficiency that we conducted include a joint test of unbiasedness and 'weak' efficiency (also known as Mincer-Zarnowitz regressions). This tests whether forecasts could have been made more accurate if they were scaled by a constant and if a constant was added/subtracted from them. Again we estimated OLS regressions with HAC standard errors:

$$y_t = \beta_0 + \beta_1 y_t^{t-h} + u_t \quad (6)$$

where u_t is a zero-mean error term, y_t is the outturn of variable y in period t and y_t^{t-h} is the forecast for variable y in period t made in period t-h. Under the null hypothesis, we jointly tested whether $\beta_0 = 0$ and $\beta_1 = 1$, reporting the p-value on the F-statistic from the Wald Test for coefficient restrictions. In the simplest case, when $\beta_0 = 0$ but $\beta_1 \neq 1$, the interpretation is straightforward: forecast errors could have been reduced by scaling the forecasts up

(1) The results are based on a HAC adjustment using an Andrew's Automatic bandwidth method. But since results can be sensitive to the choice of the bandwidth in the HAC adjustment, we also performed the tests using alternative bandwidth selection methods, specifically the Newey West fixed method (which chooses the bandwidth based on the sample size) and a fixed bandwidth of h+2 where h is the forecast horizon. The results were largely robust to the use of these alternative methods.

(2) UK-weighted world GDP is constructed by weighting together the GDP growth of UK trading partners by their share in UK exports. For more on the Bank's approach to 'nowcasting' world activity, see Stratford (2013).

(or down). In this case, there is a *conditional* bias; forecasts have generally tended to be too low (high), but the forecast error increases as the forecast value increases (decreases). The interpretation is more complicated when both $\beta_0 \neq 0$ and $\beta_1 \neq 1$. In this case, the forecasts could have been improved by scaling them (up or down) and adding a (positive or negative) constant.

We also used a similar test to assess whether past revisions to forecasts predict the final forecast revision (ie the last forecast revision before the data outturn is released). The results of such tests indicate whether forecasters have tended to under or overreact to data outturns; ideally forecast revisions should be uncorrelated over time.

We conducted two versions of the test. The first test assesses whether past revisions to forecasts for a given quarter predict the final forecast revision for the same quarter. So each of the past forecast revisions will be for forecasts of differing forecast horizons. For example, in the November 2014 *Inflation Report*, there was a (final) one quarter ahead UK GDP forecast for 2014 Q4. That forecast will have been revised from the two quarter ahead forecast made in August 2014, also for UK GDP in 2014 Q4. Similarly, the forecast made in August 2014 will have been revised from a three quarter ahead one made in May 2014, and so on. We tested whether those previous revisions predict the final revision made in the November 2014 *Inflation Report*. To do that, we estimated the following equation using OLS with HAC standard errors:

$$R_t^{t-1} = \alpha + \beta_1 R_t^{t-2} + \beta_2 R_t^{t-3} + \beta_3 R_t^{t-4} + \beta_4 R_t^{t-5} + \beta_5 R_t^{t-6} + u_t \qquad \textbf{(7)}$$

where R_t^z is the revision to the forecast for period t made in period z, ie $R_t^z = y_t^z - y_t^{z-1}$ where y_t^z is the forecast for variable y in period t made in period z. We included lags up to five quarters. The null hypothesis was that forecast revisions were uncorrelated over time and implied that $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$. We tested these parameters jointly using a F-test. We also used t-tests to determine which lags were statistically significant in explaining the final revision.

The second test of forecast revisions held the forecast horizon fixed, such that we tested if past revisions to one quarter ahead forecasts predict future revisions to one quarter ahead forecasts. We did the same for one year and two year ahead forecasts. We conducted the tests by estimating equations **(8a)** (for one quarter ahead forecasts), **(8b)** (one year ahead forecasts) and **(8c)** (two year ahead forecasts) using OLS with HAC standard errors:

$$R_t^{t-1} = \alpha + \beta_1 R_{t-1}^{t-2} + \beta_2 R_{t-2}^{t-3} + \beta_3 R_{t-3}^{t-4} + \beta_4 R_{t-4}^{t-5} + \beta_5 R_{t-5}^{t-6} + u_t \qquad \textbf{(8a)}$$

$$R_t^{t-4} = \alpha + \beta_1 R_{t-1}^{t-5} + \beta_2 R_{t-2}^{t-6} + \beta_3 R_{t-3}^{t-7} + \beta_4 R_{t-4}^{t-8} + \beta_5 R_{t-5}^{t-9} + u_t \qquad \textbf{(8b)}$$

$$R_t^{t-8} = \alpha + \beta_1 R_{t-1}^{t-9} + \beta_2 R_{t-2}^{t-10} + \beta_3 R_{t-3}^{t-11} + \beta_4 R_{t-4}^{t-12} + \beta_5 R_{t-5}^{t-13} + u_t \qquad \textbf{(8c)}$$

Again, revisions should ideally be uncorrelated over time, such that the null hypothesis was $\beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$. We tested these parameters jointly using a F-test. And again we also used t-tests to determine which lags were statistically significant in explaining the final forecast revision.

2.2.4 Benchmarking of Bank forecasts

As described earlier in this subsection, accuracy, unbiasedness and efficiency are the three metrics that we used to assess the performance of Bank forecasts. For unbiasedness and efficiency, these metrics alone are sufficient for us to judge if the forecasts are 'good enough', since forecast errors should not be systematically positive or negative, and they should not be systematically related to information available when the forecasts were made. Judging how accurate forecasts should be is less straightforward: forecasts may be inaccurate because something happened that was unforeseeable, not because forecasters made systematic mistakes.

To set benchmarks for accuracy, we compare Bank forecast accuracy to that of a simple forecasting rule, and to the forecasting accuracy of private sector forecasters and other central banks. While the statistical tests described in Sections 2.2.1 to 2.2.3 are standard forecast diagnostics, this benchmarking aspect of our evaluation is relatively novel, especially the comparisons with private sector forecasts.

As described later in this subsection, we compare central bank forecasting accuracy to the accuracy of individual private sector forecasters, in contrast to previous forecast evaluations which have typically used an average forecast calculated across private sector forecasts.(1) We are aware of only one paper, Blix *et al* (2001), which conducted forecast comparisons using individual private sector forecasts. That paper compared individual forecasts from the Consensus surveys to forecasts made by the IMF and OECD across six countries (not including the United Kingdom), but it did not compare the forecasting performance of these institutions to central banks' forecasting performance.

When benchmarking Bank forecasting performance to that of a simple rule, we used random walk processes without drift — in other words, each variable was assumed to remain constant at its latest known value. Such a benchmark is commonly used in the forecast evaluation literature and has been found to be quite robust to structural changes which often negatively affect forecasts based on econometric models.(2)

A series x follows a random walk process (without drift) if

$$x_t = x_{t-1} + e_t \tag{9}$$

where $E(e_t) = 0$. The random-walk forecast for variable x in period t, made at time t-h is defined as:

$$x_t^{t-h} = x_{t-h} \tag{10}$$

We constructed these random-walk forecasts using real-time data, such that we used the latest known data outturn at the time the forecasts were made. We then used a Diebold-Mariano test to assess whether differences in accuracy (measured by squared forecast errors) between the Bank forecasts and random-walk generated forecasts were statistically significant. To conduct this test, a difference in squared forecast errors for forecaster A and forecaster B at time t is defined as:

$$d_t = e_t^{2A,t-h} - e_t^{2B,t-h} \tag{11}$$

where $e_t^{2A,t-h}$ is the h-quarters ahead squared forecast error made by forecaster A at time t.

The following equation is then estimated using Ordinary Least Squares (OLS) with HAC standard errors:

$$d_t = \beta_0 + u_t \tag{12}$$

where u_t is a zero-mean error term. The null hypothesis is that there is no difference in accuracy between the two forecasts, ie $\beta_0 = 0$. If $\beta_0 > 0$, then forecaster B has tended to be more accurate than forecaster A, and *vice versa* if $\beta_0 < 0$.

To compare Bank forecasting accuracy with that of private sector forecasters we again used RMSEs. We first compared (unscaled) RMSEs over the full sample period and used Diebold-Mariano tests to assess whether differences in accuracy between Bank and private sector forecasts were statistically significant. We also compared (unscaled) RMSEs calculated over a four-year rolling window in order to highlight changes in relative accuracy over time.(3)

We compared Bank forecasts with forecasts from other central banks in two key ways. First we directly compared the RMSEs for each variable, scaled by the standard deviation of data outturns over the same period as we calculated the RMSEs. Scaling by data volatility helps to account for the fact that different economies are affected by different shocks; scaled RMSEs therefore provide a fairer comparison of forecast accuracy across countries than unscaled ones. Scaling RMSEs by the standard deviations of outturns does not, however, fully

(1) See for example, Andersson *et al* (2007) or IMF (2014).
(2) For a discussion, see for example Eitrheim *et al* (2004) and Clements and Hendry (1999).
(3) We chose a four-year window on the assumption that it would be long enough to smooth through year-to-year volatility in forecast errors, while being short enough to usefully illustrate how forecast accuracy was evolving over time.

account for the possibility that economies with particular characteristics may be more or less difficult to forecast than others. For example, the United States is a large, relatively closed economy, in comparison with the United Kingdom. That could mean that forecasting US GDP growth is easier than forecasting UK GDP growth because US GDP growth is less affected by shocks emanating from other countries.

Second, to address the issue that different economies have different structural characteristics, we compared each central bank's forecasting accuracy (again defined by (unscaled) RMSEs) with the accuracy of private sector forecasters in the same country. We then assessed the relative performance of forecasts made by the Bank against the relative performance of forecasts made by other central banks. And again, we used Diebold-Mariano tests to assess whether differences in accuracy between central bank forecasts and private sector forecasts were statistically significant.

2.3 Context

As with all statistical exercises, interpretation of results needs to take account of background and context. For example, as set out in Section 1, forecasts play an important role in policymaker debate, decision-making and communication, and so the 'value-add' of central bank forecasting processes cannot be judged solely by assessing the performance of point forecasts. Interpretation of results also needs to bear in mind the nature of the data. For example, as discussed in Section 2.1, the quality of unpublished Bank staff projections was likely to be higher towards the end of the forecast period than in the early days of the MPC.

Before presenting our results in Section 3, this subsection highlights some other key pieces of context that should be borne in mind when interpreting results, sample size being the most important.

Sample size

When assessing the performance of Bank forecasts, we need to bear in mind certain caveats. Among the most important is sample size. Depending on the variable, the sample sizes used to assess the performance of quarterly Bank forecasts varied between 48 and 69 observations, with the exception of household and corporate lending growth where the sample sizes were much smaller (**Table 2.D**). For the comparisons with other central banks and private sector forecasters, our maximum sample size was 17 observations; as noted earlier, we were obliged to use

Table 2.D Sample sizes used to assess Bank forecast variables in this report

Variable	Forecasts available from:	Number of observations available for:		
		One quarter ahead forecasts	One year ahead forecasts	Two year ahead forecasts
Real GDP growth	August 1997	68	65	61
Inflation	August 1997	69	66	62
Unemployment rate[a]	May 1998	65	62	58
Real private consumption growth	August 1997	69	66	62
Real total investment growth	August 1997	69	66	62
Nominal wage growth[b]	August 1997	67	61	53
Nominal house price growth	August 1997	69	66	62
Nominal household lending growth	February 2007	31	28	24
Nominal corporate lending growth	February 2009	23	20	16
Real US GDP growth	February 2001	55	52	48
Real euro-area GDP growth	February 2001	55	52	48
Calendar-year forecasts[c]	*1998*	*n.a.*	*17*	*16*

(a) For most of the period since the MPC's inception, the Bank has forecast the unemployment rate as measured by the Labour Force Survey. A small number of forecasts made prior to May 1998 used the claimant count measure. These have been excluded from our sample.
(b) The Office for National Statistics discontinued the Average Earnings Index in September 2010, after it had been replaced with the alternative Average Weekly Earnings series in January 2010. This causes a break in the nominal wage growth forecasts, which lowers the number of observations, particularly at the one and two year ahead forecast horizons.
(c) The sample size for Bank calendar-year forecasts was the same for all variables assessed, except for unemployment (which starts in 1999 and therefore has one less observation for the one and two year ahead forecasts) and the US and EA GDP growth forecasts (which start in 2001 and therefore have three less observations at the one and two year ahead horizon). Calendar-year forecasts were not assessed for household and corporate lending growth.

calendar-year forecasts for this part of the evaluation. As such, our benchmarking exercise and our assessment of Bank lending growth forecasts are particularly subject to the following caveats about sample size.(1)

Small samples usually imply greater uncertainty about the estimated parameters in statistical tests, which reduces the power of those tests to reject the null hypotheses (eg of unbiasedness, efficiency or equal forecasting performance of two forecasters). For example, it is possible that where the tests detect a systematic bias, the result is solely driven by a rare event that happened to occur during the sample period. To mitigate this problem to some extent, we perform our analysis on the full sample as well as on the sample excluding 2008 and 2009, as these years were associated with an arguably difficult to predict financial crisis (and its frequency in our sample is unlikely to reflect the true probability of it occurring).

The small sample problem also implies that it is more difficult to distinguish between accurate and less accurate forecasters, and in particular that a less accurate forecaster may appear to be accurate because of luck rather than superior forecasting ability. The probability of that occurring is related to the chosen confidence interval (ie the probability of a 'Type I' error). By construction, for 10% confidence intervals the probability of rejecting the null hypothesis when it is in fact true is 10%. Given the large number of tests performed for this evaluation, it is inevitable that in some cases when the test leads to a rejection of the null hypothesis, this will reflect a 'Type I' error.(2)

Moreover, as discussed in Broadbent (2013), the less predictable the series (ie the less signal about it can be extracted from available indicators) the longer the sample required to judge forecast performance.

Forecast uncertainty

Forecasting is inherently uncertain. As is widely discussed in the forecast literature (eg Clements and Hendry (1998)), forecast uncertainty can take numerous forms. There is uncertainty about future economic events (including the likely structural evolution of the economy and unexpected economic events (or 'shocks')). And there is uncertainty related to the economic model itself (including uncertainty about its estimation, its parameterisation, and uncertainty about the data in the base period in which the forecast begins). As set out by, among others, King (2010), the likelihood of any economic variable evolving precisely in line with a point forecast is close to zero.

The presence of uncertainty means that forecast 'errors' (defined throughout this paper as outturn less forecast), are inevitable, and the existence of forecast errors in and of themselves should not be regarded a shortcoming of a forecast process or a modelling approach. A more pertinent question is whether there are systematic patterns to the errors (for example, bias or inefficiency) or whether the pattern of forecast errors made by the Bank appears materially different to those made by other forecasting institutions (our benchmarking exercise).

The Bank's 'fan charts' — a visual depiction of the underlying probability distributions for its key economic forecasts — are a way of portraying the uncertainty that is inherent in economic forecasting. The fan charts show the probability distributions around modal forecasts.(3) Ideally, our evaluation would focus on probability distributions, rather than point estimates, for the full range of variables in our sample. However, as noted in Section 2.1, fan charts are only available over a sufficiently long period for UK GDP growth and inflation; for most of the variables under consideration, we only have modal forecasts. We provide an assessment of the performance of the Bank's inflation and GDP fan charts in Box 4 on pages 50–51.

Forecast independence

Throughout our evaluation we assess forecasts for different variables separately. We do recognise, however, that forecasts for individual variables are not made independently of one other. In particular, forecasts for variables such as consumption and investment are produced by Bank staff to be consistent with the MPC's GDP and inflation forecasts. And lending projections are produced using a model that is conditional on the forecasts for

(1) For a study of the small sample behaviour of the Diebold-Mariano test see Clark and McCracken (2013).
(2) A statistical technique called the Bonferroni adjustment could be used to mitigate this issue. But it requires that forecasts are independent of each other, which is unlikely in our sample, as forecast errors for various variables at various horizons all come from the same central forecasting process or are conditional on its outputs.
(3) For a fuller description of the fan charts, and their interpretation, see the box on page 39 of the November 2007 *Inflation Report*.

GDP and other variables from the central forecasting process. As a result, errors made in forecasting one variable, could be 'caused' by errors made in forecasting another variable. We do not explore this formally, but note that, in many cases, there was not a strong correlation between forecast errors for the variables in our sample (see **Tables A.4.A–A.4.C** in Annex 4).

Conditionality of forecasting

Forecasts of macroeconomic variables are described as being 'conditional' when they are produced on the assumption that a key underlying variable will follow a path that is not necessarily the forecaster's own expectation. 'Unconditional' forecasts are produced using only the forecaster's own information and expectations. The MPC (modal) forecasts for inflation, GDP and the unemployment rate are produced on the assumption that Bank Rate will follow the path implied by the market yield curve.(1) They are, therefore, said to be conditioned on the market-implied path for Bank Rate. The MPC's forecasts are also conditioned on a range of assumptions for other variables, including for the sterling effective exchange rate and energy prices.(2)

Conditionality complicates the assessment of forecast performance, because forecast errors may be caused by the conditioning path assumption, rather than an error in the forecaster's own expectation. If, for example, the market-implied path for interest rates is not aligned with the MPC's own view of the appropriate path for rates to return inflation to target over the medium term, then the forecast performance of the conditional forecasts may not be a good guide to the performance of the MPC's own (private) forecasts that they actually use to set policy.(3)

At least in terms of Bank Rate, conditionality is likely to be a more important issue for longer-horizon forecasts than shorter-term ones because of the lags that are generally assumed to operate between monetary policy being changed and it having an impact on the real economy. And it will only be an important issue if there is often a marked discrepancy between the market-implied path and the MPC's (unconditional) expectation for Bank Rate. Since MPC members do not divulge their expectations for Bank Rate we cannot test this formally. Nevertheless, our analysis does not suggest that discrepancies exist between the market-implied path for Bank Rate and the MPC's unconditional expectations. For example, as shown in Section 3, MPC inflation forecasts have tended to cluster around, or a little below, the 2% target at the two-year forecast horizon. That suggests that the MPC's (unconditional) expectation for interest rates has generally been close to the market-implied path. Moreover, if it were the case that the forecasts were conditioned on a path for Bank Rate that turned out to be much lower (higher) than the MPC's own expectation, and the MPC set Bank Rate according to its own expectation, we would expect the MPC's GDP and inflation forecasts that were conditioned on that market-implied path to both be too high (low). But as shown in Section 3, at the two-year horizon, GDP forecasts have tended to be too high, while inflation forecasts have been too low, which points to errors in forecasting the supply side of the economy, rather than errors stemming from the conditioning path.

A further potential complication associated with the conditionality of forecasts is that we may not be comparing 'like-for-like' when comparing the MPC's forecast performance with other central banks and private sector forecasters. The other central banks in our peer comparison have varying conventions in producing their forecasts, for example, the Federal Open Market Committee (FOMC) in the United States produces forecasts conditioned on the assumption of 'appropriate monetary policy', which may differ across FOMC participants.(4) And we cannot be sure about the nature of any conditioning paths used by private sector forecasters as such information is not provided by the organisations that collate these forecasts. A detailed assessment of the extent to which differences in conditioning assumptions may be causing differences in performance is beyond the scope of this project.

(1) Through most of its existence, the MPC has produced forecasts conditioned on both a constant Bank Rate and on forward market interest rates. Until the May 2004 *Inflation Report*, the MPC emphasised those forecasts conditioned on a constant path for Bank Rate as its main forecasts. Since August 2004, the MPC has typically emphasised those forecasts conditioned on forward market interest rates as its main ones. For more information on the conditioning path for market interest rates see www.bankofengland.co.uk/publications/Pages/inflationreport/conditioning_path.aspx.

(2) Specifically, in addition to assumptions about Bank Rate and the sterling effective exchange rate, the MPC's forecasts are also conditioned on assumed paths for oil prices, gas prices and nominal government expenditure. Since May 2009, they have also been conditioned on an assumed path for the stock of asset purchases financed by the creation of central bank reserves. The conditioning paths used in each forecast round are set out in the corresponding *Inflation Report*. See for example Table 1 in 'Conditioning assumptions, MPC key judgements, and indicative projections' August 2015; www.bankofengland.co.uk/publications/Documents/inflationreport/2015/augca.pdf.

(3) For further discussion of evaluating conditional forecasts, see Faust and Wright (2008).

(4) ECB macroeconomic projections are conditioned on: (i) a constant path for short-term interest rates before June 2006; (ii) market expectations for short-term interest rates since June 2006. Riksbank macroeconomic projections are conditioned on: (i) a constant path for the policy rate between 1999 and 2005 MPR no. 2; (ii) market expectations for the policy rate between 2005 MPR no. 3 and 2006 MPR no. 3; (iii) Riksbank projections for the policy rate since 2007. RBNZ macroeconomic projections incorporate an endogenous response of interest rates to the economic outlook since 1997.

Evolution of forecast processes over time

Forecasts assessed in this evaluation span a 17-year period, and were produced by different members of the MPC and of Bank staff over that time. Moreover, although the structure of the forecast process as described in the box on page 10 and in Bean and Jenkinson (2001) has remained broadly unchanged over that period, the tools used by the staff to implement that process have evolved. Bean and Jenkinson (2001) note that 'A central tool in the production of these forecasts is a relatively standard macroeconometric model (MM)'.(1) The MM was replaced in 2003 by the Bank of England Quarterly Model (BEQM) (see Harrison *et al* (2005)). And from the November 2011 *Inflation Report*, the forecast process has been supported by the forecasting platform including COMPASS and the suite of models, as described on page 11 and in Burgess *et al* (2013).

For most of the tests reported in this paper, and to mitigate the difficulties frequently encountered with small samples set out above, we included as many observations as possible in our analysis. But by including as long a time period as possible, it may be that some of the issues detected by our tests are historical in nature, and are not pertinent to the current forecasting approach — for example, difficulties associated with previous models that no longer form part of the Bank forecast process. Moreover, as discussed further in Section 2.1, the degree of internal scrutiny and oversight for some variables — in particular the unemployment rate — has changed over time. And, in the wake of the Court-commissioned Stockton Review (Stockton (2012)), a wide range of changes were made to the Bank's forecast processes; it is too soon to form a judgement about whether these changes have been associated with improved forecast performance. Where possible, this evaluation includes an assessment of how forecast accuracy has evolved over time. But in doing that, we are not assessing the performance of any particular individuals or models.

(1) The MM is described in Bank of England (1999) and Bank of England (2000).

3 Empirical results

In this section we set out our main empirical findings. Throughout this section, it is important to keep in mind the key pieces of context set out in Section 2.3, in particular the issue of small samples that is most relevant when comparing Bank forecasts to private sector forecasts and those from other central banks.

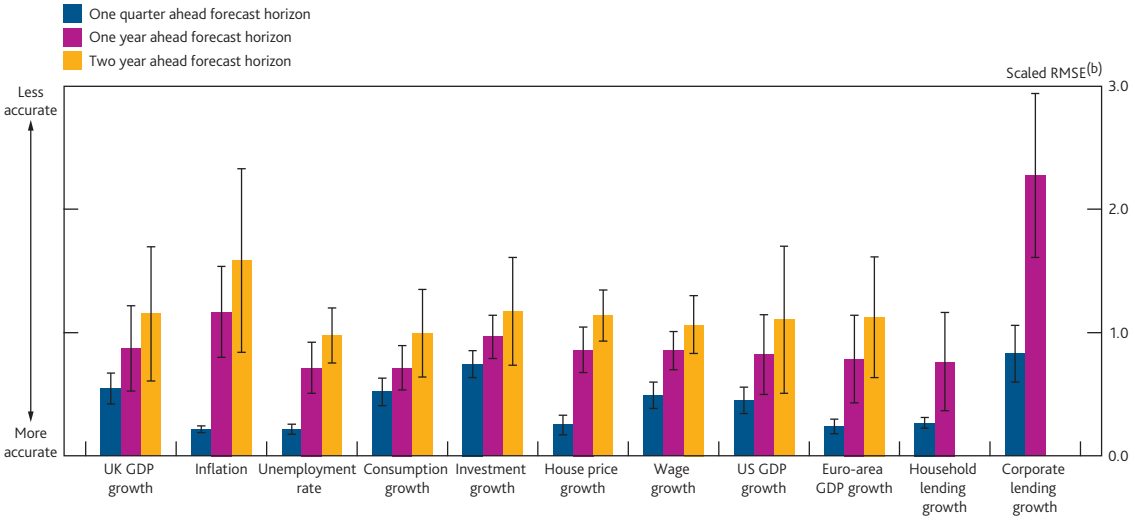
This section starts with the accuracy, unbiasedness and efficiency of Bank forecasts across the eleven macroeconomic variables assessed. We then compare the accuracy of Bank forecasts to forecasts from a simple model, and for a subset of those eleven variables, Bank forecast accuracy is compared with UK private sector forecasters and other central banks. As set out in Section 2.1, data constraints mean that the majority of our analysis is restricted to testing point estimates of central tendency. But for completeness, tests of the full forecast distributions for UK GDP growth and inflation are set out in Box 4 on pages 50–51. Further results, in particular the results of our robustness checking, are provided in Annex 4.

3.1 Accuracy of Bank forecasts

As described in Section 2.2, we measure forecast accuracy over a given period using RMSEs scaled by the standard deviation of data outturns over that same period. Scaled RMSEs for (quarterly) Bank forecasts at all three horizons across our full sample period are shown in **Chart 3.1**. Unscaled RMSEs are provided in Annex 4. As noted in Section 2.2, the value of a scaled RMSE in **Chart 3.1** is not in itself informative. But we can compare the degree of accuracy across variables and across time periods, with a higher scaled RMSE indicating that forecasts of a particular variable have tended to be less accurate, relative to the volatility of the data outturns.

Chart 3.1 shows that, for all variables, forecast accuracy has tended to lessen as the forecast horizon increases. That is intuitive; as the forecast horizon increases, available data will provide a weaker signal about the likely path of a given variable. Moreover, there is greater scope for unforeseen shocks to occur as the forecast horizon lengthens. For most of the assessed variables, the 90% confidence intervals for scaled RMSEs at longer horizons do not overlap with the confidence intervals for scaled RMSEs at short horizons. This means we can be reasonably confident that the apparent differences in accuracy between short and longer-term forecasts are statistically

Chart 3.1 Accuracy of Bank forecasts (full sample 1997–2014)(a)



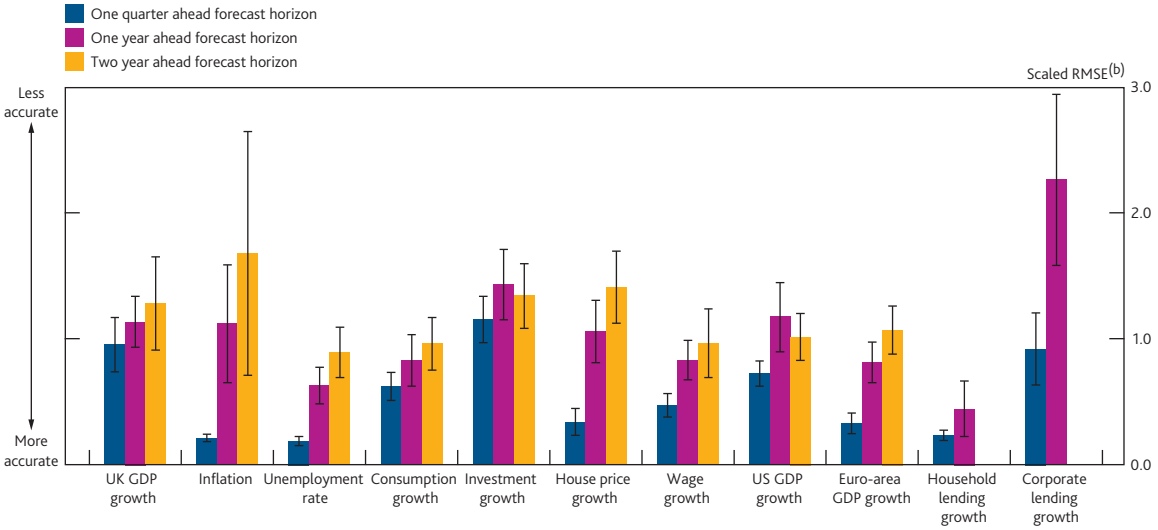
Note: Full details of data sources used for charts and tables in this report are provided in Annex 3.

(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3. Two year ahead forecasts for household and corporate lending are not assessed because the sample sizes are very short.
(b) Root mean squared errors scaled by the standard deviation of the data outturns. Error bars indicate 90% confidence intervals for each variable in each period of time, constructed using the delta method (see Annex 3 for more details).

significant at the 10% level.(1) These confidence intervals were constructed using the delta method (see Annex 3 for more details).

Comparing across variables, inflation and unemployment rate forecasts have tended to be the most accurate at the one quarter ahead horizon, while investment and corporate lending growth forecasts have tended to be the least accurate. At the one year ahead horizon, corporate lending growth forecasts have been materially less accurate than the other variables. In contrast to their relative accuracy at the one quarter ahead horizon, inflation forecasts have tended to be the least accurate at the two year ahead horizon. That remains the case when we exclude 2008 and 2009 from the sample, as illustrated in **Chart 3.2**. Indeed, all of the key results from **Chart 3.1** hold when we exclude 2008 and 2009.

Chart 3.2 Accuracy of Bank modal forecasts (full sample excluding 2008 and 2009)(a)



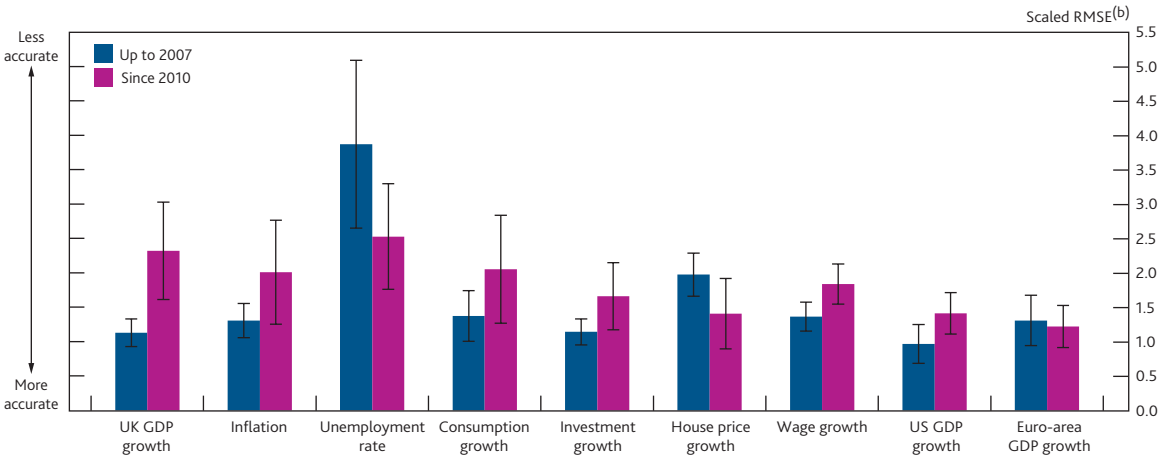
(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3. Two year ahead forecasts for household and corporate lending are not assessed because the sample sizes are very short.
(b) Root mean squared errors scaled by the standard deviation of the data outturns. Error bars indicate 90% confidence intervals for each variable in each period of time, constructed using the delta method (see Annex 3 for more details).

Comparing across time, we found that forecast accuracy at the two year ahead horizon appears to have generally decreased in the period from 2010 to 2014, compared with the period prior to 2008 (**Chart 3.3**). The exceptions were the forecasts of the unemployment rate and house price growth, where accuracy appears to have been greater in the 2010–14 period compared with the pre-2008 period; and euro-area GDP growth, where accuracy is similar across the two periods. We should note that, for all variables except for UK GDP growth, the 90% confidence intervals overlap, meaning that we cannot be confident that the differences in scaled RMSEs in the pre and post-crisis periods are statistically significant. To some extent that reflects the relatively small post-crisis sample size. Nevertheless, the broad-based nature of the apparent decline in forecasting accuracy is notable.

Although forecast accuracy appears to have decreased in general for two year ahead forecasts in the post-2010 period relative to the pre-2008 period, that is not the case for shorter-horizon forecasts, where forecast accuracy was generally little changed between the two periods (see **Charts A.4.1** and **A.4.2** in Annex 4). These contrasting results are consistent with the emergence of shocks in recent years that had greater persistence than those experienced pre-2008. The results could also be consistent with forecasters systematically underestimating how persistent observed shocks were likely to be, to a greater extent than they did pre-2008. Both of these hypotheses are plausible, given the unprecedented nature of the 2007–08 financial crisis and its aftermath.(2) We do not test these hypotheses formally, but our comparisons of Bank forecast errors with those of private sector

(1) Comparison of our RMSE confidence intervals is complicated by the possible interdependencies between forecasts of different variables, and forecasts of the same variable over time. In general, if confidence intervals do not overlap, we can be reasonably confident that the differences between RMSEs are statistically significant at the 10% level.
(2) For discussion of the nature of the recovery post the financial crisis see, for example, Bean (2012).

Chart 3.3 Accuracy of Bank modal two year ahead forecasts, comparison across time[a]

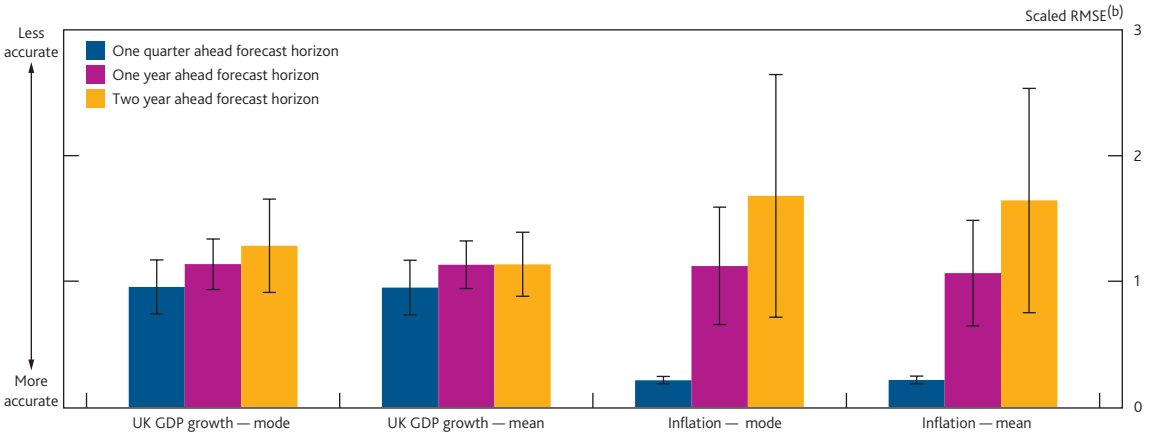


(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3. Forecasts for household and corporate lending are not assessed because the sample sizes are very short.
(b) Root mean squared errors scaled by the standard deviation of the data outturns. Error bars indicate 90% confidence intervals for each variable in each period of time, constructed using the delta method (see Annex 3 for more details).

forecasters later in this section indicate that the UK macroeconomy was more difficult to forecast in the 2010–14 period, relative to the pre-2008 period, particularly at the two-year forecast horizon.

Our key results on forecast accuracy remained intact when we calculated forecast errors using mean forecasts rather than modal ones (**Chart 3.4**); note that, as set out in Section 2.1, mean forecasts are only available for UK GDP and inflation. Our key findings were also unaffected by using real-time, rather than the latest available, data (see **Chart 3.5** and **Charts A.4.5, A.4.7 and A.4.9** in Annex 4).

Chart 3.4 Accuracy of Bank modal and mean forecasts (full sample excluding 2008 and 2009)[a]

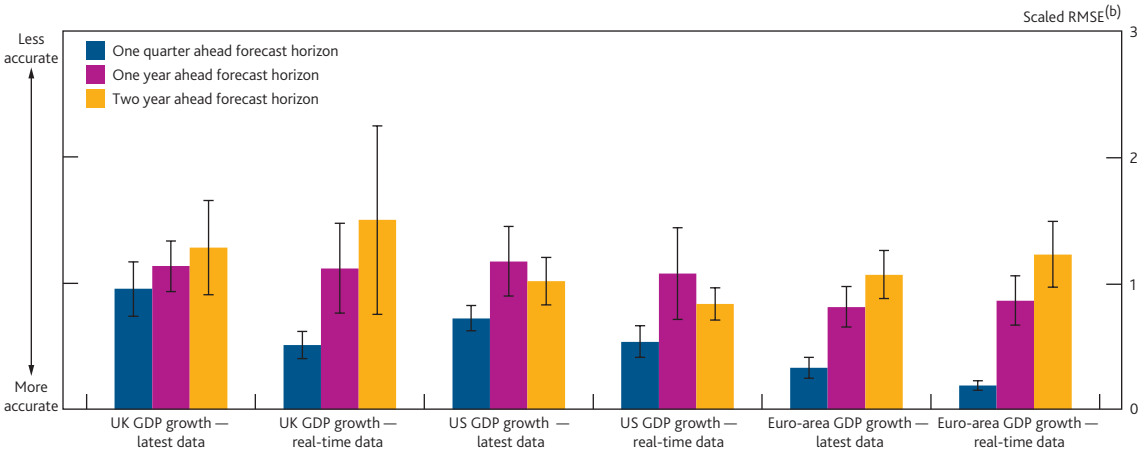


(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3. Forecasts for household and corporate lending are not assessed because the sample sizes are very short.
(b) Root mean squared errors scaled by the standard deviation of the data outturns. Error bars indicate 90% confidence intervals for each variable in each period of time, constructed using the delta method (see Annex 3 for more details).

3.2 Unbiasedness of Bank forecasts

Table 3.A shows the estimated bias coefficients, and associated p-values, for Bank forecasts over the full sample, excluding 2008 and 2009. Consistent with our measure of forecast accuracy, and again to allow comparability across variables and time periods, we present the estimated degree of bias in forecast errors (given by β_0 in equation **(4)**) scaled by the standard deviation of the data outturns (over the same period as the bias was calculated). Unscaled bias coefficients are provided in Annex 4.

Chart 3.5 Accuracy of Bank modal forecasts, against latest and real-time data (full sample excluding 2008 and 2009)[a]



(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3. Forecasts for household and corporate lending are not assessed because the sample sizes are very short.
(b) Root mean squared errors scaled by the standard deviation of the data outturns. Error bars indicate 90% confidence intervals for each variable in each period of time, constructed using the delta method (see Annex 3 for more details).

Table 3.A Unbiasedness of Bank modal forecasts (full sample excluding 2008 and 2009)[a]

	Bias (scaled)[b] — positive (negative) coefficient means that forecasts have tended to be too low (high)			Statistical significance of unscaled bias coefficient, p-value[c]		
	One quarter ahead	One year ahead	Two years ahead	One quarter ahead	One year ahead	Two years ahead
UK GDP growth	0.17	0.02	-0.39	0.31	0.95	0.29
Inflation	0.01	0.37	0.75	0.67	0.18	0.22
Unemployment rate	-0.07	-0.33	-0.41	0.04**	0.04**	0.05**
Consumption growth	0.14	0.24	0.18	0.20	0.20	0.56
Investment growth	-0.08	-0.12	-0.18	0.69	0.65	0.49
House price growth	0.13	0.69	1.14	0.04**	0.00**	0.00**
Wage growth	-0.02	-0.29	-0.65	0.74	0.07*	0.00**
US GDP growth	-0.22	-0.24	-0.45	0.07*	0.33	0.02**
Euro-area GDP growth	0.21	0.10	-0.30	0.00**	0.69	0.33
Household lending growth	0.03	-0.17	n.a.	0.66	0.25	n.a.
Corporate lending growth	0.04	-1.10	n.a.	0.85	0.07*	n.a.

(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3. Two year ahead forecasts for household and corporate lending are not assessed because the sample sizes are very short.
(b) Unbiasedness is assessed using the following regression, as described in Section 2.2.
$e_t^{t-h} = \beta_0 + u_t$ where e_t^{t-h} is a forecast error and u_t is a zero-mean error term. To facilitate comparability, first three columns show β_0 from the regression, scaled by the standard deviation of the data outturns.
(c) 5% significance is denoted by **. 10% significance is denoted by *.

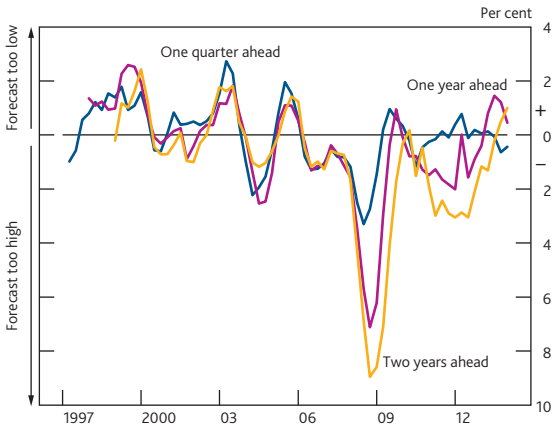
In general, the estimated bias coefficients tended to be larger for longer-horizon forecasts. That is consistent with forecast accuracy tending to decline as the forecast horizon increases (**Chart 3.1**).

For UK GDP growth and inflation forecasts there was no statistically significant evidence of bias at the 10% level at any forecast horizon. Nevertheless, relative to data outturns, inflation forecasts have tended to be low, especially at the two-year horizon, while two year ahead UK GDP growth forecasts have tended to be high. **Charts 3.6** and **3.7**, which plot the Bank's modal forecast errors for UK GDP growth and inflation over time, suggest that this pattern of over (under) prediction of two year ahead GDP growth (inflation) may have been more pronounced in the period from 2010 to 2014 than in the period prior to 2008.

While we do not formally assess the cause of these errors, we note that a tendency for GDP growth forecasts to be too high, and inflation forecasts to be too low, is consistent with unexpected cost shocks or weaker-than-expected supply growth. In the early stages of the post-crisis period, the MPC expected the boost to inflation from higher energy and import prices to be short-lived, and that productivity would rise steadily as the economy recovered. As set out in Hackworth, Radia and Roberts (2013), however, both energy costs, and

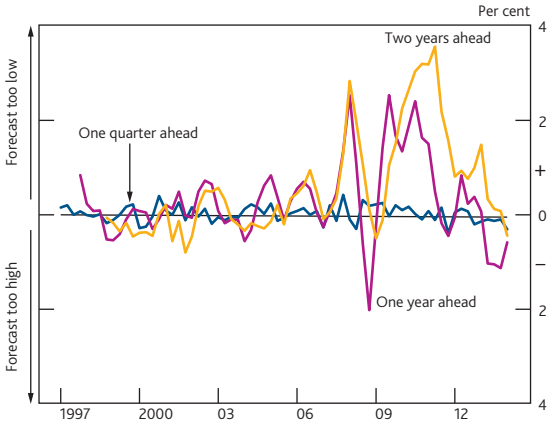
non-energy import costs, rose much more sharply than anticipated between 2010 and 2013, broadly accounting for the stronger-than-expected inflation outturns during that period. At the same time, productivity growth was much weaker than the MPC had initially anticipated: in 2010, the MPC had anticipated that whole-economy output per hour would rise by around 10 percentage points over the subsequent three years; in fact, it posted only a small rise. Barnett *et al* (2014) provide a detailed analysis of the likely causes of the weakness in UK productivity over the 2008 to 2013 period. They identify a series of persistent impairments associated with the financial crisis, including reduced investment in physical and intangible capital, impaired resource allocation, and unusually high firm-survival rates.

Chart 3.6 Bank modal UK GDP growth forecast errors(a)



(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3. Forecast errors computed against the latest vintage of data.

Chart 3.7 Bank modal inflation forecast errors(a)



(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3.

As with UK GDP growth and inflation forecasts, we failed to reject the hypothesis that forecasts have been unbiased at the 10% level for consumption and investment growth, as well as for household lending growth. That was the case at all forecast horizons. In contrast, the estimated biases in the unemployment rate and house price growth forecasts were statistically significant (at the 5% level) at all forecast horizons. Forecasts for the unemployment rate have tended to be too high, while those for house price inflation have tended to be too low. Indeed, the estimated (scaled) bias coefficient for forecasts for house price inflation was the largest in our sample at the two year ahead horizon. Like unemployment rate forecasts, wage growth forecasts have tended to be too high; that bias was significant at the 10% level for one year ahead forecasts and at the 5% level for two year ahead forecasts. US GDP growth forecasts have also tended to be too high, and that bias was significant at the 10% level for the one quarter ahead forecasts and at the 5% level for the two year ahead forecasts.

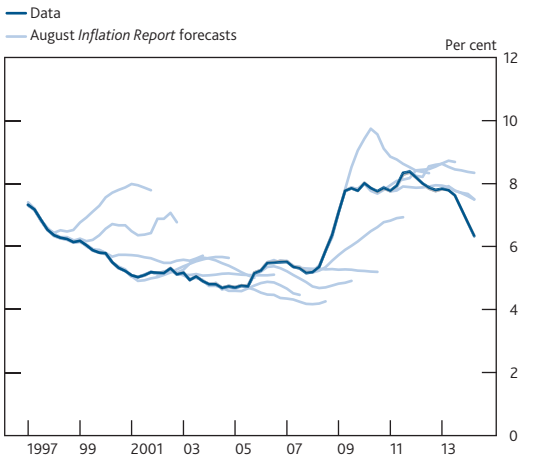
As noted above, we do not formally explore the root cause of these forecast errors, but we note that the biases in both unemployment rate and wage growth forecasts over the more recent period appear to have been associated with the surprising post-crisis weakness of productivity, widely discussed both inside and outside the Bank. For example, Broadbent (2012) and Barnett *et al* (2014) describe the possible impact of the financial crisis on investment and resource reallocation. Whatever the precise cause, UK productivity growth has been considerably weaker than the Bank had anticipated, which in itself would also be consistent with stronger employment growth (and so lower unemployment) than predicted for a given pace of demand growth, as well as with weaker wage inflation. The latter part of our forecast evaluation period was also characterised by participation rates in the UK labour market that were generally higher than implied by the MPC's central forecast projections (reflecting, for example, the abolition of the default retirement age, concerns about the adequacy of pension and saving income following the financial crisis and changes to the benefit regime).(1) *Ceteris paribus*, a higher labour force participation rate would also be associated with lower outturns for wage growth.

The early years of our sample, also a period where unemployment and wage outturns generally were lower than anticipated, were seen as another time of structural change in the labour market. Specifically, the equilibrium

(1) See, for example, the discussion on page 28 of the Bank's August 2014 *Inflation Report*.

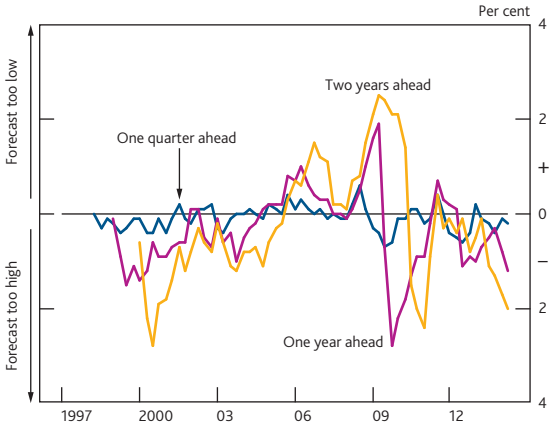
unemployment rate was thought to be declining, associated with a marked reduction in union coverage over the preceding decade, as well as successive reforms to the benefit regime (see, for example, Quintini and Nickell (2001)). For the unemployment rate, in particular, the early part of the sample period was associated with relatively large forecast errors (**Chart 3.8** (similar charts for other variables can be found in Annex 4); also **Chart 3.9**). Indeed, if we omit the earliest years of the sample period from our evaluation exercise (1998 to 2000), there was no statistically significant evidence of bias in Bank projections in the unemployment rate.

Chart 3.8 Bank unemployment rate forecasts and data outturns(a)



(a) For clarity, we only show August *Inflation Report* forecasts in this chart, but we do include quarterly unemployment rate forecasts from all *Inflation Reports* between May 1998 and May 2014 in our analysis.

Chart 3.9 Bank unemployment rate forecast errors(a)



(a) Sample period varies by forecast horizon, due to data availability. Forecast errors are computed using modal forecasts and the latest vintage of data. Unemployment forecasts made since August 2013 have been explicitly owned by the MPC. Prior to that, they were staff projections consistent with the MPC's modal forecasts for GDP and inflation.

The results of our unbiasedness tests were largely robust to including the worst years of the recent recession in the sample period. For most variables, including 2008 and 2009 did not alter the results in **Table 3.A** materially (see **Table A.4.D** in Annex 4). One notable difference was that the negative bias in two year ahead unemployment rate forecasts halved and became statistically insignificant at the 10% level. That reflects the large positive errors in unemployment rate forecasts seen in 2008 and 2009, when unemployment rose more quickly than expected. It seems likely that those positive errors were driven, at least in part, by the unexpectedly large falls in GDP in the early stages of the 2008–09 recession, rather than issues specific to unemployment forecast itself.

Consistent with the results for modal forecasts in **Table 3.A**, we failed to reject the hypothesis that forecasts have been unbiased at the 10% level for mean UK GDP growth and inflation forecasts, at all forecast horizons (**Table 3.B**).

We also compared modal GDP growth forecasts to real-time data instead of the more mature vintage of data. (**Table 3.B**). As noted above, we failed to reject the hypothesis of unbiasedness in UK GDP growth forecasts at the 10% significance level at all horizons when assessing forecasts against the latest available data. In contrast, UK GDP growth forecasts tended to be too high when assessed against real-time data. Those biases were statistically significant at the 5% level for one quarter ahead forecasts, and at the 10% level for two year ahead forecasts. These results are consistent with the MPC's stated aim to forecast mature vintages of data, incorporating expected revisions to early data vintages in their UK GDP forecasts.(1) In particular, early estimates of UK GDP growth data have tended to be revised up a little over time.(2) Unscaled bias coefficients are included in **Table A.4.F** in Annex 4.

(1) See, for example, the box on page 39 of the November 2007 *Inflation Report*.
(2) Official National Accounts data are subject to revision for a number of years following their initial release. These revisions partly reflect the receipt and processing of new information; they also partly reflect that the methods used to measure GDP growth are subject to continuous review. The Office for National Statistics publishes an annual *Blue Book* which contains detailed revisions to National Accounts data. Once data have been through two *Blue Books*, the convention has been to treat them as 'mature' estimates, although it remains possible that they will still be subject to further revision in future years, for example following methodological changes. For more, see, for example, Cunningham and Jeffery (2007).

There was statistically significant evidence of bias in Bank forecasts of US GDP growth at the one quarter and two year ahead horizons, when assessed against the latest available data, but that was not the case when we assessed those forecasts against real-time data (**Table 3.B**). The results were mixed for euro-area GDP growth forecasts; we found statistically significant evidence of bias only at the one quarter ahead horizon when assessed against the latest data and only at the two year ahead horizon when assessed against real-time data.

Table 3.B Unbiasedness of Bank modal and mean forecasts against alternative data vintages (full sample excluding 2008 and 2009)[a]

	Bias (scaled)[b] — positive (negative) coefficient means that forecasts have tended to be too low (high)			Statistical significance of unscaled bias coefficient, p-value[c]		
	One quarter ahead	One year ahead	Two years ahead	One quarter ahead	One year ahead	Two years ahead
UK GDP growth — mode	0.17	0.02	-0.39	0.31	0.95	0.29
UK GDP growth — mean	0.22	0.15	-0.17	0.21	0.57	0.56
Inflation — mode	0.01	0.37	0.75	0.67	0.18	0.22
Inflation — mean	-0.01	0.31	0.65	0.84	0.24	0.26
Modal projections						
UK GDP growth — latest data	0.17	0.02	-0.39	0.31	0.95	0.29
UK GDP growth — real-time data	-0.17	-0.36	-0.72	0.04**	0.19	0.09*
US GDP growth — latest data	-0.22	-0.24	-0.45	0.07*	0.33	0.02**
US GDP growth — real-time data	0.07	0.07	-0.13	0.42	0.75	0.48
Euro-area GDP growth — latest data	0.21	0.10	-0.30	0.00**	0.69	0.33
Euro-area GDP growth — real-time data	0.01	-0.11	-0.55	0.66	0.67	0.09*

(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3. See page 16–17 for a discussion of the use of 'real-time' and latest vintages of data to compute forecast errors.
(b) Results are from an OLS regression using HAC standard errors, as defined in Section 2.2.
(c) 5% significance is denoted by **. 10% significance is denoted by *.

3.3 Efficiency of Bank forecasts

As described in Section 2.2.3, we conducted three sets of efficiency tests; strong efficiency tests, joint tests of unbiasedness and weak efficiency, and tests of the predictability of forecast revisions. This subsection discusses the results from each set in turn.

3.3.1 Strong efficiency of Bank forecasts

Forecasts are deemed to be 'strongly' efficient if forecast errors are uncorrelated with information known at the time the forecasts were made (for more information on the tests we conducted, see Section 2.2.3). The results from our strong efficiency tests are shown in **Table 3.C** for Bank forecasts over the full sample, excluding 2008 and 2009. **Table 3.C** shows the estimated β_1 coefficients and associated p-values as defined in equation **(5)**. Results for the full sample including 2008 and 2009 are provided in Annex 4.

For one year ahead UK GDP growth and inflation forecasts we could not reject the hypothesis that forecast errors were unrelated to past forecast errors or past outturns. In the case of one quarter ahead GDP and inflation forecasts, only GDP forecasts were related to data outturns known at the time the forecasts were made (negatively related and significant at 10%). However, for two year ahead forecasts, both GDP and inflation forecast errors were positively and significantly (at the 5% level) related to data outturns known at the time the forecasts were made. In particular, when known GDP or inflation data were high, future outturns tended to be higher than the forecasts and vice versa. **Chart 3.10** indicates that, for GDP, the positive relationship between forecast errors and data outturns was driven by the pre-crisis period, while **Chart 3.11** suggests that, for inflation, the relationship has been stronger since the change of target from RPIX to CPI. For GDP growth, two year ahead forecast errors were also positively and significantly (at the 5% level) related to previous forecast errors known when the forecasts were made. These results were mostly robust to using mean forecasts instead of modal ones, and using real-time GDP data (see **Table A.4.H** in Annex 4). Overall, these results could indicate that the MPC did not build enough persistence from known shocks into its longer-horizon forecasts. We did not formally test this hypothesis, but we discuss other evidence later in this subsection that is consistent with it.

For the unemployment rate, the strong efficiency tests indicated that the Bank's forecasts could have been improved by placing more weight on past outturns; the null hypothesis of strong efficiency was rejected at all horizons at the 10% significance level (**Table 3.C**). The tests pointed to a negative relationship between the unemployment forecast errors and known data outturns, such that forecasts tended to be too high when the unemployment rate was known to be high and vice versa; which indicates that Bank forecasts tended to predict more inertia in the unemployment rate than actually occurred. The null hypothesis of strong efficiency was also rejected at the 5% significance level for one quarter ahead unemployment rate forecasts when we tested the information contained in past forecast errors.

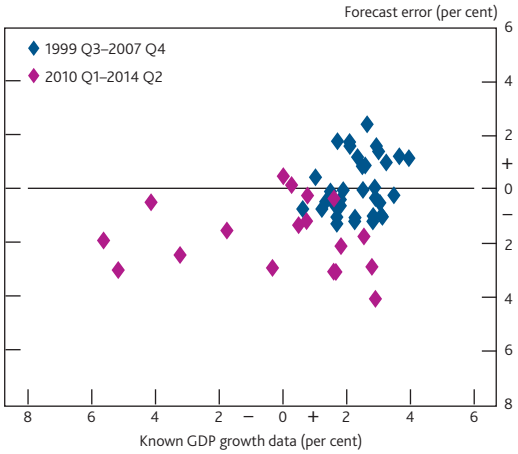
Table 3.C (Strong) efficiency of Bank modal forecasts (full sample excluding 2008 and 2009)[a]

β_1 (p-value) from regressions of forecast errors on:

	Previous forecast errors			Previous variable outturns		
	One quarter ahead	One year ahead	Two years ahead	One quarter ahead	One year ahead	Two years ahead
UK GDP growth	-0.05	0.04	0.15	-0.27	0.02	0.26
	(0.81)	(0.72)	(0.04**)	(0.06*)	(0.79)	(0.00**)
Inflation	0.03	0.04	0.25	-0.02	-0.03	0.30
	(0.80)	(0.82)	(0.15)	(0.37)	(0.70)	(0.01**)
Unemployment rate	0.39	0.00	-0.08	-0.05	-0.23	-0.32
	(0.00**)	(1.00)	(0.43)	(0.10*)	(0.05*)	(0.09*)
Consumption growth	-0.29	-0.04	0.40	-0.08	0.15	0.55
	(0.20)	(0.71)	(0.00**)	(0.36)	(0.01**)	(0.00**)
Investment growth	-0.21	-0.28	-0.28	-0.43	-0.43	-0.19
	(0.35)	(0.13)	(0.09*)	(0.00**)	(0.01**)	(0.22)
House price growth	-0.09	-0.28	0.18	-0.03	0.02	0.23
	(0.49)	(0.11)	(0.03**)	(0.40)	(0.89)	(0.00**)
Wage growth	0.08	-0.09	0.19	-0.04	0.19	0.41
	(0.62)	(0.66)	(0.38)	(0.61)	(0.30)	(0.00**)
US GDP growth	-0.06	-0.34	0.02	-0.31	-0.38	-0.03
	(0.74)	(0.03**)	(0.68)	(0.00**)	(0.00**)	(0.72)
Euro-area GDP growth	0.06	-0.13	-0.25	0.06	-0.25	-0.24
	(0.59)	(0.37)	(0.08*)	(0.27)	(0.01**)	(0.04**)
Household lending growth	-1.10	0.14	n.a.	-0.02	-0.24	n.a.
	(0.00**)	(0.36)		(0.66)	(0.00**)	
Corporate lending growth	-0.13	-0.40	n.a.	-0.62	-0.56	n.a.
	(0.39)	(0.11)		(0.01**)	(0.03**)	

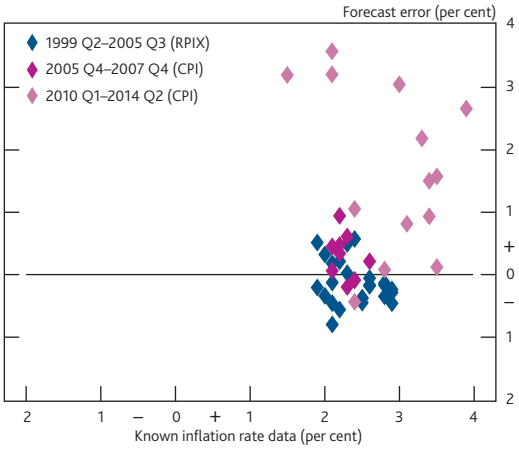
(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3. Two year ahead forecasts for household and corporate lending are not assessed because the sample sizes are very short. The table shows β_1 and the associated p-value (in brackets) from the following regression: $e_t^{t-h} = \beta_0 + \beta_1 z_{t-h} + u_t$ where e is a forecast error, as defined in equation **(3)** in Section 2, z is a variable that was known to the forecaster at time t-h (either a previous forecast error or a previous data outturn), and u is a zero-mean error term. 5% significance is denoted by **. 10% significance is denoted by *.

Chart 3.10 Bank modal two year ahead UK GDP growth forecast errors and known data outturns[a]



(a) Forecast errors are computed using the latest vintage of data.

Chart 3.11 Bank modal two year ahead inflation forecast errors and known data outturns[a]



(a) The MPC published its first CPI inflation forecast in February 2004.

Across the other variables tested, the results were mixed, but in general we found that the null hypothesis of strong efficiency tended to be rejected more often when we tested for the information contained in past outturns compared with tests that considered past forecast errors (**Table 3.C**).

As well as testing whether forecast errors were related to past forecast errors and past data outturns, we also tested whether UK GDP growth forecast errors were related to past world GDP growth outturns and contemporaneous Bank staff forecasts of world GDP growth. In our full sample excluding 2008 and 2009, we could not reject the null hypothesis of strong efficiency at any horizon or for any specification when we included past outturns of world GDP growth in the tests (**Table A.4.J** in Annex 4). But the tests did indicate a statistically significant (at the 10% level) relationship between one and two year ahead UK GDP growth forecast errors and contemporaneous world GDP growth forecasts (of the same horizon). That suggests the longer-horizon UK GDP forecasts could have been improved by more effectively incorporating information contained in staff analysis of global growth prospects. We note, however, that this result is not statistically significant for the one year ahead forecasts when assessed using the mean UK GDP growth projections rather than the modal projections. That suggests that for the one year ahead forecasts at least, the MPC have tended to incorporate information about global prospects into their wider forecast probability distribution.

3.3.2 Weak efficiency of Bank forecasts

Forecasts are said to be 'weakly' efficient if they could not have been systematically improved by scaling by, or adding, a constant value (see Section 2.2 for more detail on the tests we conducted). The results from our joint tests of unbiasedness and weak efficiency are shown in **Table 3.D** for Bank forecasts over the full sample, excluding 2008 and 2009. **Table 3.D** shows the estimated β_0 and β_1 coefficients as defined in equation **(6)**, and the p-values from the joint test of whether $\beta_0 = 0$ and $\beta_1 = 1$. Results for the full sample including 2008 and 2009, as well as results for different specifications (including for mean forecasts) are provided in Annex 4.

Table 3.D (Weak) efficiency of Bank modal forecasts (full sample excluding 2008 and 2009)(a)(b)

	One quarter ahead			One year ahead			Two years ahead		
	β_0	β_1	p-value from joint test: β_0=0, β_1=1	β_0	β_1	p-value from joint test: β_0=0, β_1=1	β_0	β_1	p-value from joint test: β_0=0, β_1=1
UK GDP growth	1.27	0.56	0.09*	2.02	0.23	0.00**	5.05	-0.83	0.00**
Inflation	0.08	0.97	0.50	1.64	0.39	0.00**	4.23	-0.82	0.00**
Unemployment rate	0.30	0.94	0.02**	1.39	0.72	0.00**	2.00	0.62	0.00**
Consumption growth	0.63	0.86	0.11	0.77	0.85	0.39	0.93	0.74	0.34
Investment growth	1.92	0.34	0.00**	3.43	-0.04	0.00**	4.00	-0.27	0.00**
House price growth	1.30	0.95	0.09*	5.83	0.76	0.00**	7.72	0.90	0.00**
Wage growth	0.25	0.92	0.51	0.51	0.77	0.01**	-3.07	1.50	0.00**
US GDP growth	0.61	0.68	0.00**	1.38	0.39	0.00**	0.57	0.67	0.07*
Euro-area GDP growth	0.13	1.11	0.00**	0.44	0.75	0.63	0.74	0.39	0.12
Household lending growth	0.16	0.99	0.90	0.03	0.77	0.00**	n.a.	n.a.	n.a.
Corporate lending growth	-1.16	0.53	0.00**	-2.32	0.16	0.00**	n.a.	n.a.	n.a.

(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3. Two year ahead forecasts for household and corporate lending are not assessed because the sample sizes are very short.
(b) The table shows estimated β_0 and β_1 coefficients from the following regression: $y_t = \beta_0 + \beta_1 y_t^{t-h} + u_t$ where u_t is a zero-mean error term, y_t is the outturn of variable y in period t and y_t^{t-h} is the forecast for variable y in period t made in period t–h. Under the null hypothesis, we jointly test whether $\beta_0 = 0$ and $\beta_1 = 1$, reporting the p-value on the F-statistic from the Wald Test for coefficient restrictions. 5% significance is denoted by **. 10% significance is denoted by *.

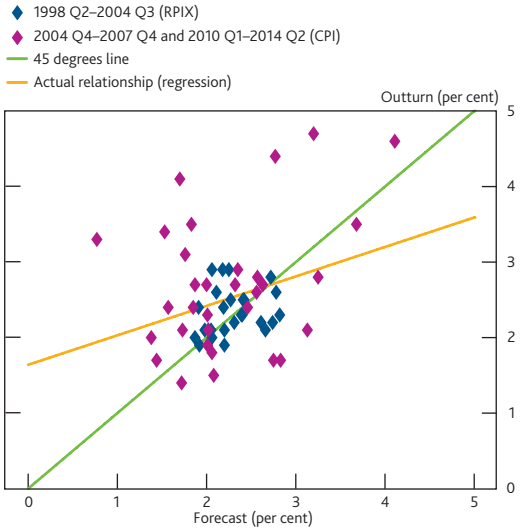
Across all three forecast horizons, we rejected the null hypothesis of unbiasedness and weak efficiency, mostly at the 5% significance level, for almost all of the variables assessed. The exceptions are consumption growth, where we could not reject the null hypothesis at the 10% level for any of the horizons; and euro-area GDP growth, where we could only reject the null hypothesis for one-quarter ahead forecasts. These results are largely robust to the inclusion of 2008 and 2009, and across different specifications (including using mean forecasts) (see **Tables A.4.L** and **A.4.M** in Annex 4).

Further investigation of the one year ahead inflation forecasts suggests that low inflation forecasts have been too low and high ones too high. If the forecasts were both unbiased and efficient, they would be clustered around the 45 degree line in **Chart 3.12**. The actual estimated relationship is shown by the orange line: inflation forecasts

that were around 3% or lower tended to be too low (ie the diamonds in **Chart 3.12** lie above the 45 degree line), while inflation forecasts that were above 3% tended to be too high.

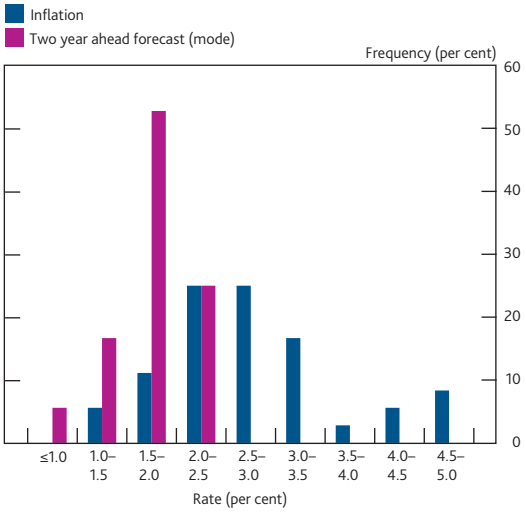
For two year ahead inflation forecasts, the rejection of the null hypothesis of weak efficiency appears to stem from the clustering of forecasts around, or a little below, the target. These forecasts have been much less dispersed than the data outturns, particularly for the CPI forecasts, as illustrated by **Chart 3.13** (this is less of a case for the preceding RPIX forecasts as shown in **Chart A.4.24** in Annex 4). The dispersion of two year ahead UK GDP growth forecasts has also been notably less than exhibited by the actual data outturns, with forecasts clustering close to the 1997–2007 average growth rate of GDP of around 3% (**Chart 3.14**). One potential explanation for the difference in distributions of forecasts and data outturns is that shocks occurred that the MPC did not foresee when it made its forecasts. Alternatively, the MPC could have underestimated the persistence of shocks known at the time it made those forecasts, such that its inflation and GDP growth forecasts have reverted towards the target or the mean too quickly. Interpretation of the clustering of the two year ahead inflation projections is complicated, however, by the fact that inflation is the Bank's target variable, and so the projections may reflect considerations of communication and monetary strategy.

Chart 3.12 Bank modal one year ahead inflation forecasts and data outturns[a]



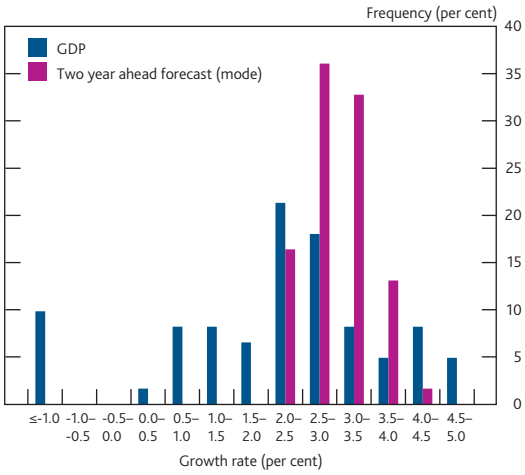
(a) The orange line shows the relationship between forecasts and outturns over the full sample period excluding 2008–09, using standard ordinary least squares regression.

Chart 3.13 Bank modal two year ahead CPI inflation forecasts and data outturns[a]



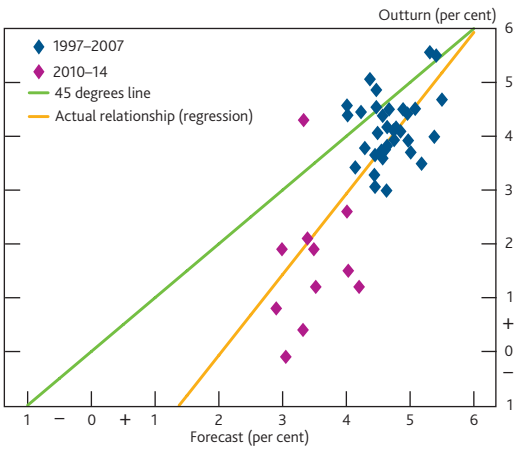
(a) The MPC published its first CPI inflation forecast in February 2004.

Chart 3.14 Bank modal two year ahead UK GDP growth forecasts and data outturns[a]



(a) Chart covers the period 1999 Q2–2014 Q2.

Chart 3.15 Bank two year ahead UK wage growth forecasts and data outturns[a]



(a) Annual growth rate of nominal wages. Measured using the whole-economy Average Earnings Index until May 2010 (the series was subsequently discontinued by the ONS). From August 2010, measured using Average Weekly Earnings. Diamonds show modal forecasts against the latest vintage of data outturns. The orange line shows the relationship between forecasts and outturns over the full sample period excluding 2008–09, using standard ordinary least squares regression.

As with inflation at the one-year horizon, we found that low unemployment rate forecasts have tended to be too low and high forecasts too high (at both the one and two-year horizons). However, this relationship does not appear to have been stable over time (**Charts A.4.25** and **A.4.26** in Annex 4).

As noted in Section 3.2, two year ahead wage growth forecasts have tended to be too high, and that bias was statistically significant at the 5% level. That bias means that we reject the null hypothesis of weak efficiency for these forecasts. In particular, it appears that the bias in these two year ahead forecasts is conditional on the level of wage growth (**Chart 3.15**); the extent of bias has tended to be larger when forecasts (and outturns) have been lower. That relationship has emerged in recent years as wage growth has been consistently lower than in the pre-crisis period.

3.3.3 Predictability of Bank forecast revisions

This subsection reports the results from our tests of whether past revisions to forecasts help to predict the final forecast revision before the data outturn is released (the tests are described in more detail in Section 2.2). We first assessed whether past revisions to forecasts for a given quarter help to predict the final forecast revision for the same quarter. Our regressions each included five consecutive quarters of past forecast revisions (with each of those past forecasts having a different horizon, because they were all forecasting the same point in the future); the estimated coefficients are shown in **Table 3.E**. We tested the statistical significance of individual β coefficients, as well as testing the joint hypothesis that all of the β coefficients were equal to zero. The relevant p-values from those tests are also shown in **Table 3.E**.

For a number of variables — specifically growth in UK GDP, consumption, investment and euro-area GDP — we could not reject the hypothesis that final forecast revisions have been unrelated to previous forecast revisions. For US GDP growth, we rejected the joint hypothesis that the β coefficients were all equal to zero, but none of the coefficients were individually significant at the 10% level.

Final revisions to inflation and unemployment rate forecasts have been significantly and negatively related to revisions made four or five quarters previously. The negative sign indicates that later revisions have tended to offset earlier revisions, which could point to some overreaction to information received when the earlier revisions were made. But more generally, the other estimated β coefficients for the remaining inflation and unemployment rate forecast revisions made in previous quarters were mostly positive, small, and not statistically significant at the 10% level. In contrast, negative serial correlation was evident across almost all of the estimated β coefficients when we assessed the predictability of final revisions to wage growth forecasts and to household and corporate lending growth forecasts. Those estimated β coefficients tended to be relatively large and were statistically significant at the 5% level in the one or two quarters preceding the final revisions.

We also assessed whether past revisions to forecasts for a given forecast horizon help to predict the final forecast revision for the same horizon. As with the first test of forecast revisions, our regressions each included the preceding five forecast revisions (with each of those past forecasts having the same horizon, such that they were all forecasting a different point in the future). And as before, we tested the statistical significance of individual β coefficients, as well as testing the (joint) hypothesis that all of the β coefficients were equal to zero. The estimated coefficients and relevant p-values for the one year ahead forecasts are shown in **Table 3.F**. The remaining results from those tests are shown in **Table A.4.O** in Annex 4.

For the one year ahead forecasts we rejected the joint hypothesis, mostly at the 5% significance level, that past revisions are unrelated to final forecast revisions for a number of variables; inflation, the unemployment rate, consumption and wage growth, and household and corporate lending growth. For one quarter ahead forecasts that set of variables was smaller with some differences; inflation, consumption and house price growth, and household lending growth. Generally, across both forecast horizons, the estimated β coefficients were small. An exception was the estimated β_1 coefficient for one year ahead inflation forecasts. This coefficient was large, positive and statistically significant at the 5% level, which could indicate a tendency to underreact to data news when revising one year ahead inflation forecasts. In contrast, for one year ahead wage growth forecasts three of the estimated β coefficients were negative and statistically significant (at the 10% level), which could indicate a tendency to overreact to data news when revising one year ahead wage growth forecasts. For the two year ahead forecasts, across all of the assessed variables, very few of the estimated β coefficients were statistically significant at the 10% level and we could not reject the null joint hypothesis that the β coefficients were all equal to zero.

Table 3.E Predictability of Bank forecast revisions when forecasting a single calendar quarter (full sample excluding 2008 and 2009)[(a)]

Modal forecasts:	α (p-value)	β_1 (p-value)	β_2 (p-value)	β_3 (p-value)	β_4 (p-value)	β_5 (p-value)	p-value from the joint test: β_i=0 $\forall i$=1…,5
UK GDP growth	-0.02	0.13	0.11	0.05	-0.04	-0.03	
	(0.67)	(0.28)	(0.20)	(0.71)	(0.79)	(0.82)	0.40
Inflation	0.03	0.05	0.11	0.14	-0.30	-0.19	
	(0.47)	(0.74)	(0.22)	(0.15)	(0.01**)	(0.14)	0.02**
Unemployment rate	-0.08	0.03	0.00	0.02	0.04	-0.17	
	(0.04**)	(0.76)	(0.99)	(0.77)	(0.57)	(0.00**)	0.02**
Consumption growth	0.10	-0.05	-0.12	-0.06	-0.03	0.14	
	(0.24)	(0.74)	(0.20)	(0.59)	(0.78)	(0.27)	0.58
Investment growth	-0.32	-0.16	-0.08	-0.06	0.00	0.12	
	(0.42)	(0.18)	(0.51)	(0.66)	(0.99)	(0.51)	0.73
House price growth	1.41	-0.03	-0.18	-0.03	-0.25	-0.12	
	(0.00**)	(0.79)	(0.01**)	(0.64)	(0.02**)	(0.23)	0.07*
Wage growth	-0.23	-0.17	-0.31	-0.16	-0.18	-0.29	
	(0.01**)	(0.08*)	(0.03**)	(0.19)	(0.18)	(0.04**)	0.05**
US GDP growth	-0.03	-0.03	-0.08	-0.14	-0.10	0.07	
	(0.72)	(0.81)	(0.57)	(0.13)	(0.41)	(0.45)	0.07*
Euro-area GDP growth	0.06	-0.02	0.07	-0.01	0.14	0.10	
	(0.31)	(0.88)	(0.33)	(0.96)	(0.21)	(0.52)	0.48
Household lending growth	-0.12	-0.44	-0.57	-0.49	-0.12	-0.26	
	(0.52)	(0.02**)	(0.11)	(0.13)	(0.65)	(0.07*)	0.01**
Corporate lending growth	-1.26	-0.04	-0.70	0.15	-0.66	-0.22	
	(0.03**)	(0.87)	(0.00**)	(0.55)	(0.02**)	(0.14)	0.04**
Mean forecasts:							
UK GDP growth	0.00	0.07	0.07	0.04	0.01	-0.08	
	(0.96)	(0.56)	(0.46)	(0.80)	(0.93)	(0.52)	0.76
Inflation	-0.01	0.09	0.14	0.08	-0.27	-0.18	
	(0.82)	(0.50)	(0.14)	(0.47)	(0.01**)	(0.13)	0.01**

(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3. Table shows results of the following regression: $R_t^{t-1} = \alpha + \beta_1 R_t^{t-2} + \beta_2 R_t^{t-3} + \beta_3 R_t^{t-4} + \beta_4 R_t^{t-5} + \beta_5 R_t^{t-6} + u_t$ where R_t^z is the revision to the forecast for period t made in period z, ie $R_t^z = y_t^z - y_t^{z-1}$ where y_t^z is the forecast for variable y in period t made in period z. Table shows beta coefficients, with the associated p-values in brackets. The last column shows the p-value on the F-statistic from the test that $\beta_i = 0$ for all i. 5% significance is denoted by **. 10% significance is denoted by *.

3.4 Accuracy of Bank forecasts compared with a simple forecasting model

As discussed in Section 2.2.4, we benchmarked the accuracy of Bank forecasts against a simple forecasting model, private sector forecasters, and other central banks. Comparisons with private sector forecasters and with other central banks are discussed in the following subsections. For the simple forecasting model, we used forecasts generated by random walks without drift. **Table 3.G** shows the RMSEs from those random walk processes relative to the RMSEs of the Bank forecasts, at all three forecast horizons.

For almost all variables, the Bank forecasts have been at least as accurate as the random-walk generated forecasts, and generally more accurate. For some variables, generally at the one quarter ahead horizon, the difference in RMSEs (with the Bank more accurate) was statistically significant at the 5% or 10% level. The exception was the Bank unemployment rate forecasts; these forecasts were less accurate than the random-walk forecasts at all horizons over the full sample. That result was statistically significant at the 10% level for the one quarter and two year ahead forecasts and at the 5% level for the one year ahead forecasts. Further investigation indicated that this result was driven in large part by the early part of the sample; when we excluded the 1998 to 2000 period, unemployment forecasts were still less accurate than the random-walk generated ones at all forecast horizons, but the differences were no longer statistically significant (at the 10% level). Excluding the pre-2001 period also changed the significance of the results for the wage growth forecasts; over the shortened sample the random-walk generated forecasts were significantly more accurate (at the 5% level) than Bank forecasts at the two-year horizon.

Table 3.F Predictability of Bank forecast revisions when forecasting one year ahead (full sample excluding 2008 and 2009)[(a)]

	α (p-value)	β_1 (p-value)	β_2 (p-value)	β_3 (p-value)	β_4 (p-value)	β_5 (p-value)	p-value from the joint test: β_i=0 $\forall i$=1…,5
Modal forecasts:							
UK GDP growth	-0.03	0.01	0.05	-0.08	-0.01	0.13	
	(0.67)	(0.96)	(0.74)	(0.59)	(0.93)	(0.23)	0.78
Inflation	0.11	0.42	-0.12	0.05	0.17	-0.13	
	(0.06*)	(0.01**)	(0.46)	(0.86)	(0.38)	(0.36)	0.02**
Unemployment rate	-0.17	-0.06	-0.16	-0.10	0.04	-0.13	
	(0.00**)	(0.59)	(0.10)	(0.37)	(0.68)	(0.14)	0.07*
Consumption growth	0.10	-0.11	-0.15	-0.10	-0.11	0.26	
	(0.33)	(0.46)	(0.37)	(0.48)	(0.40)	(0.01**)	0.04**
Investment growth	0.33	0.09	-0.03	0.05	-0.07	0.03	
	(0.40)	(0.51)	(0.78)	(0.68)	(0.62)	(0.87)	0.89
House price growth	2.54	0.06	-0.35	0.06	-0.27	-0.11	
	(0.00**)	(0.67)	(0.02**)	(0.68)	(0.06*)	(0.43)	0.15
Wage growth	-0.15	-0.33	0.00	-0.27	-0.25	0.18	
	(0.19)	(0.05*)	(0.98)	(0.05*)	(0.06*)	(0.27)	0.03**
US GDP growth	0.01	0.02	-0.08	-0.17	-0.07	0.16	
	(0.88)	(0.89)	(0.62)	(0.27)	(0.52)	(0.12)	0.22
Euro-area GDP growth	-0.15	0.18	-0.15	-0.12	0.05	-0.11	
	(0.26)	(0.20)	(0.31)	(0.33)	(0.66)	(0.32)	0.50
Household lending growth	0.13	-0.20	-0.10	-0.18	-0.02	-0.37	
	(0.37)	(0.23)	(0.65)	(0.36)	(0.86)	(0.04**)	0.01**
Corporate lending growth	-1.10	0.50	-0.45	0.33	-0.12	-0.48	
	(0.17)	(0.18)	(0.18)	(0.37)	(0.57)	(0.01**)	0.05**
Mean forecasts:							
UK GDP growth	-0.01	0.06	0.02	-0.24	0.01	0.06	
	(0.86)	(0.70)	(0.87)	(0.02**)	(0.93)	(0.38)	0.31
Inflation	0.08	0.43	-0.10	0.05	0.09	-0.08	
	(0.08)*	*(0.00**)*	*(0.55)*	*(0.75)*	*(0.47)*	*(0.59)*	0.04**

(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3. Table shows results of regressions 8a (for one quarter ahead forecasts), 8b (one year ahead forecasts) and 8c (two year ahead forecasts) using OLS with HAC standard errors:

$$R_t^{t-1} = \alpha + \beta_1 R_{t-1}^{t-2} + \beta_2 R_{t-2}^{t-3} + \beta_3 R_{t-3}^{t-4} + \beta_4 R_{t-4}^{t-5} + \beta_5 R_{t-5}^{t-6} + u_t \quad \textbf{(8a)}$$

$$R_t^{t-4} = \alpha + \beta_1 R_{t-1}^{t-5} + \beta_2 R_{t-2}^{t-6} + \beta_3 R_{t-3}^{t-7} + \beta_4 R_{t-4}^{t-8} + \beta_5 R_{t-5}^{t-9} + u_t \quad \textbf{(8b)}$$

$$R_t^{t-8} = \alpha + \beta_1 R_{t-1}^{t-9} + \beta_2 R_{t-2}^{t-10} + \beta_3 R_{t-3}^{t-11} + \beta_4 R_{t-4}^{t-12} + \beta_5 R_{t-5}^{t-13} + u_t \quad \textbf{(8c)}$$

where R_t^z is the revision to the forecast for period t made in period z,*i. e.* $R_t^z = y_t^z - y_t^{z-1}$ where y_t^z is the forecast for variable y in period t made in period z. Table shows beta coefficients, with the associated p-values in brackets. The last column shows the p-value on the F-statistic from the test that $\beta_i = 0$ for all i. 5% significance is denoted by **. 10% significance is denoted by *.

Assessing forecast accuracy of the Bank and random-walk forecasts using mean projections instead of modal ones, and using real-time data instead of the latest data, did not make any material difference to the results in **Table 3.G** (see **Table A.4.R** in Annex 4). The inclusion of 2008 and 2009 also made little difference to the results, except for US and euro-area GDP growth forecasts (see **Table A.4.Q** in Annex 4). For both these variables, the difference in RMSEs between the random-walk forecasts and the Bank forecasts was statistically significant at the 5% level for one quarter ahead forecasts and at the 10% level for one and two year ahead forecasts, with the Bank forecasts having been the more accurate.

3.5 Accuracy of Bank forecasts compared with UK private sector forecasters

In this subsection we compare the accuracy of Bank forecasts to the accuracy of forecasts submitted to the surveys of private sector forecasters conducted by Consensus Economics and by HMT (in the case of house prices). We conducted this exercise for most of the variables in our sample, except for: (i) US and euro-area GDP growth, where we benchmark Bank forecasts to those produced by the Fed and ECB in the next subsection, and (ii) household and corporate lending growth, where private sector forecasts for these variables are not commonly published. As discussed in Section 2.3, the results in this and the next subsection are likely to be less robust than

Table 3.G Accuracy of forecasts from random walk without drift processes relative to accuracy of Bank modal forecasts (full sample excluding 2008 and 2009)(a)

	Ratio of RMSEs: random-walk without drift forecast accuracy relative to Bank forecast accuracy(b)			p-value from Diebold-Mariano test of differences in RMSEs(c)		
	One quarter ahead	One year ahead	Two years ahead	One quarter ahead	One year ahead	Two years ahead
UK GDP growth	1.21	1.75	1.55	0.11	0.20	0.14
Inflation	2.34	1.02	0.90	0.00**	0.91	0.67
Unemployment rate	0.74	0.56	0.75	0.08*	0.02**	0.05*
Consumption growth	1.06	1.11	0.96	0.39	0.60	0.80
Investment growth	1.14	1.47	1.55	0.24	0.15	0.07*
House price growth	1.26	1.08	0.94	0.06*	0.58	0.67
Wage growth	1.32	0.97	0.81	0.05*	0.89	0.31
US GDP growth	1.23	1.83	2.35	0.05*	0.13	0.03**
Euro-area GDP growth	1.81	2.14	1.78	0.07*	0.15	0.15
Household lending growth	1.38	1.21	n.a.	0.01**	0.40	n.a.
Corporate lending growth	1.09	0.88	n.a.	0.58	0.66	n.a.

(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3. Two year ahead forecasts for household and corporate lending are not assessed because the sample sizes are very short.
(b) A ratio of more than 1 suggests that Bank forecasts were more accurate than those generated by a random-walk without drift process (black font); a value of less than 1 suggests that the random-walk generated forecasts were more accurate (grey font). To assess the statistical significance of these differences, we used a Diebold-Mariano test (final three columns of the table).
(c) P-values are from a Diebold-Mariano test as defined in Section 2.2.4, using HAC standard errors, as defined in Section 2.2.2. 5% significance is denoted by **. 10% significance is denoted by *.

in the rest of Section 3, because private sector (and other central bank) forecasts tend to be produced for calendar years, which means our sample period included a maximum of 17 observations for one year ahead forecasts and 16 observations for two year ahead forecasts.

Starting with those forecasts submitted to the Consensus Economics surveys, the Bank's one year ahead forecasts for GDP growth and inflation have tended to be more accurate than the majority of the private sector forecasters in our sample (**Table 3.H**). In contrast, Bank forecasts for the unemployment rate, consumption growth and investment growth have tended to be less accurate than the majority of private sector forecasters. Differences in accuracy between the Bank's forecasts and private sector forecasters have not tended to be significant (at the 10% level), although the Bank's one year ahead inflation forecasts have been statistically more accurate than those produced by 29% of the private sector forecasters.

Across all five of the variables that we compared with forecasts submitted to the Consensus Economics surveys, the Bank's two year ahead forecasts were less accurate than those produced by the majority of private sector forecasters in our sample. No private sector forecaster produced forecasts that were significantly less accurate than the Bank's two year ahead forecasts. But between 11% and 24% of our sample produced two year ahead forecasts for growth in GDP, consumption and investment that were significantly (at the 10% level) more accurate than the Bank's forecasts.

To highlight changes in relative forecast accuracy over time, we also compared Bank forecasts to private sector forecasts using (unscaled) RMSEs calculated over a four-year rolling window. **Charts 3.16–3.27** show the range and interquartile range of private sector forecasters' rolling RMSEs compared with those of the Bank. It should be noted that the charts mask variation in the performance of individual forecasters across time. For example, in all cases, the dotted line showing the lowest RMSE in a particular four-year window represents differing forecasters over time. There is no single private sector forecaster who is consistently the most (or least) accurate across the whole sample.

For all five variables considered here, there was a general deterioration in accuracy across private sector forecasts and the Bank forecasts associated with the recession of 2008 and 2009. That is particularly evident for the two year ahead forecasts. In most cases, forecast accuracy improved again once those years dropped out of the rolling window. The exceptions are two year ahead inflation forecasts and both horizons of wage growth forecasts where RMSEs, both for the Bank and private sector forecasters, remained relatively elevated in the most recent window from 2011 to 2014 (**Charts 3.19**, **3.26** and **3.27**). These findings come against a backdrop of a series of

Table 3.H Accuracy of Bank's modal one year and two year ahead calendar-year forecasts, relative to forecast accuracy of UK private sector forecasters, 1998–2014(a)

	Per cent UK private sector forecasters in this comparison:		
	Who are less accurate than the Bank	Who are significantly less accurate than the Bank (10% level)	Who are significantly more accurate than the Bank (10% level)
One year ahead forecasts:			
GDP growth	67	0	0
Inflation	95	29	0
Unemployment rate	29	0	0
Consumption growth	35	5	10
Investment growth	16	0	16
Two year ahead forecasts:			
GDP growth	23	0	14
Inflation	10	0	0
Unemployment rate	29	0	0
Consumption growth	32	0	11
Investment growth	24	0	24

(a) Sample includes 17 to 22 private sector forecasters, depending on the variable and forecast horizon. Accuracy is measured using RMSEs. To assess the statistical significance of differences in accuracy between forecasters, we used a Diebold-Mariano test as defined in Section 2.2.4, using HAC standard errors, as defined in Section 2.2.2. The percentage of private sector forecasters who are significantly less accurate than the central bank is a subset of the percentage who are less accurate; the percentage of significantly less and significantly more accurate private sector forecaster do not sum to 100 because in many cases, the difference in forecast accuracy between the central bank and the private sector forecaster was not statistically significant. The source of private sector forecasts is Consensus Economics. Full details of data sources are provided in Annex 3.

unexpected cost shocks (eg rising energy prices) during the 2011–14 period, which boosted import prices, and hence inflation, as well as a prolonged period of weak growth in nominal wages (possibly associated with impairments to productivity related to the financial crisis and/or with structural shifts that have boosted labour supply).(1) We note, however, that in the case of the wage growth forecasts, especially the two year ahead forecasts, we are cautious in drawing strong conclusions about recent forecast performance due to a structural break in the official wage data series; the Office for National Statistics discontinued the Average Earnings Index in September 2010, after replacing it with the alternative Average Weekly Earnings series in January 2010.(2) This series break means that we are missing calendar-year forecast errors for 2010 for both the one and two year ahead forecasts, and missing 2011 for the two year ahead forecasts.

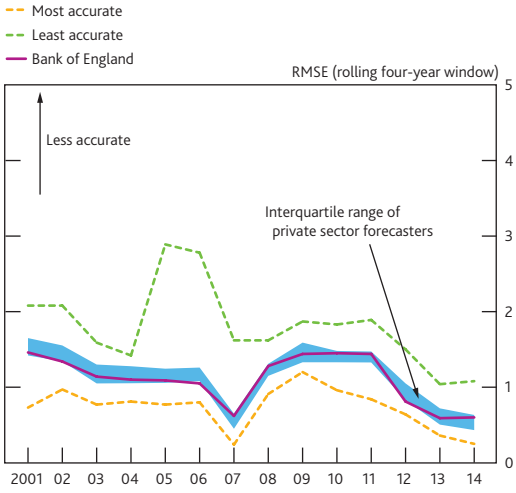
In most cases, the rolling RMSE of the Bank forecasts has remained within or close to the interquartile range of the private sector forecasters (the red line representing the rolling RMSE of Bank forecasts has typically been within, or close, to the blue swathe in **Charts 3.16** to **3.27**). Moreover, for inflation and wage growth, the one year ahead Bank of England forecasts have been more accurate than the majority of the private sector forecasts.

In contrast, the Bank's two year ahead inflation forecast accuracy decreased in the 2007–10 window, relative to the accuracy of private sector forecasts, and remained above the interquartile range of private sector forecast errors in the most recent 2011–14 window. Similarly, the relative accuracy of the Bank's two year ahead wage growth forecasts decreased in the 2009–12 window and these forecasts remained the least accurate subsequently. However, we are sceptical about the robustness of this result, given the series break in the official wage data described earlier in this section. For much of the sample period, the Bank's one and two year ahead unemployment forecasts were less accurate than many of those produced by the private sector, although the relative accuracy of the Bank forecasts did improve at the end of our sample (in 2013 and 2014). In general, the small samples involved in comparing Bank RMSEs to those of the private sector at different points of time mean that we cannot be confident that any apparent shifts in relative accuracy represent a systematic change.

Turning to house prices, **Table 3.I** compares RMSEs for one year ahead house price growth forecasts made by the Bank to those for private sector forecasts submitted to the HMT Survey of External Forecasters (as noted in Section 2.1, Consensus surveys do not include house price forecasts). The sample of private sector forecasts begins in 2009, providing a maximum of six observations. We apply a threshold of four forecasts for inclusion in our evaluation exercise, and only consider house price forecasts based on either the Halifax or Nationwide house

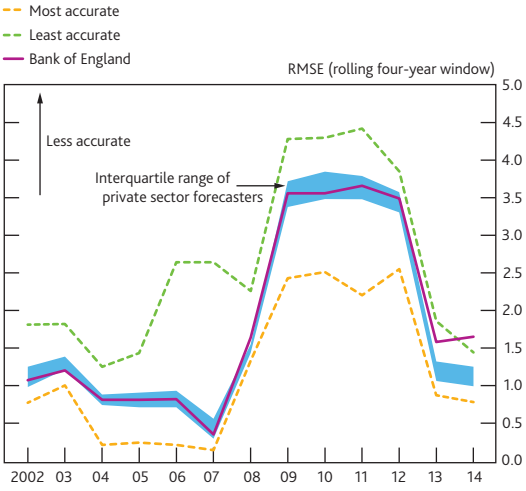
(1) See also the discussion on pages 30–31.
(2) See www.ons.gov.uk/ons/rel/aei/average-earnings-index/september-2010/withdrawal-of-average-earnings-index.pdf.

Chart 3.16 One year ahead UK GDP growth forecast accuracy (rolling four-year RMSEs)(a)



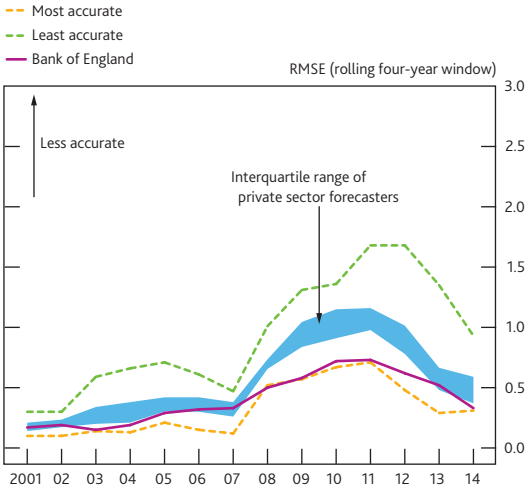
(a) Unscaled RMSEs calculated against the latest vintage of data outturns. Our sample includes 21 private sector forecasters. Dates on the horizontal axis refer to the end of the rolling window. For example, 2014 refers to 2011–14. The dates refer to the year that a forecast was made **for**, not necessarily the year it was made **in**. The source of private sector forecasts is Consensus Economics. Full details of data sources are provided in Annex 3.

Chart 3.17 Two year ahead UK GDP growth forecast accuracy (rolling four-year RMSEs)(a)



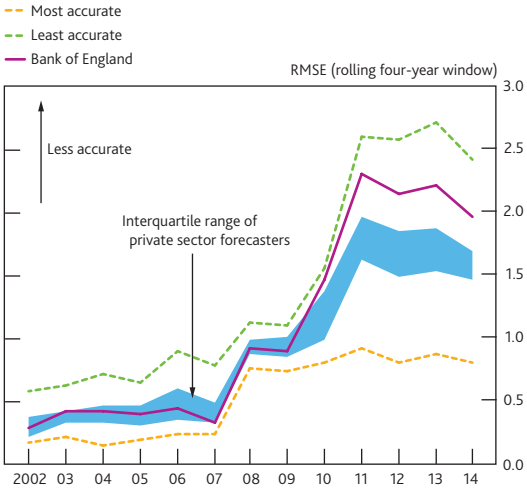
(a) See footnotes to **Chart 3.16**. For two year ahead GDP forecasts, our sample includes 22 private sector forecasters.

Chart 3.18 One year ahead UK inflation forecast accuracy (rolling four-year RMSEs)(a)



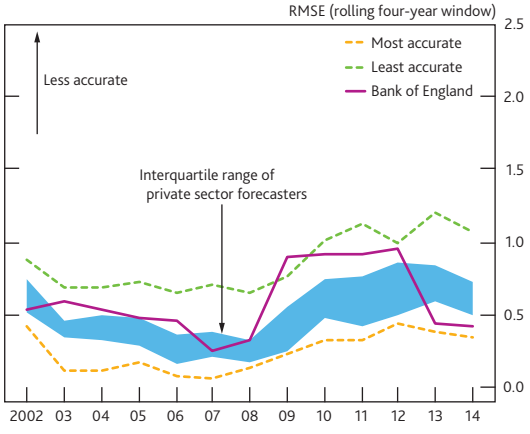
(a) See footnotes to **Chart 3.16**. Sample includes 21 private sector forecasters.

Chart 3.19 Two year ahead UK inflation forecast accuracy (rolling four-year RMSEs)(a)



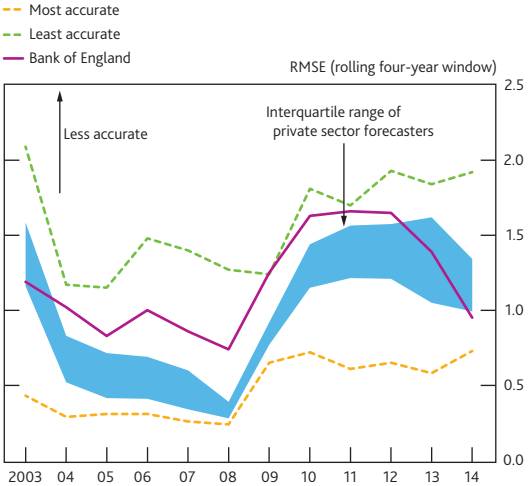
(a) See footnotes to **Chart 3.16**. Sample includes 20 private sector forecasters.

Chart 3.20 One year ahead UK unemployment rate forecast accuracy (rolling four-year RMSEs)(a)



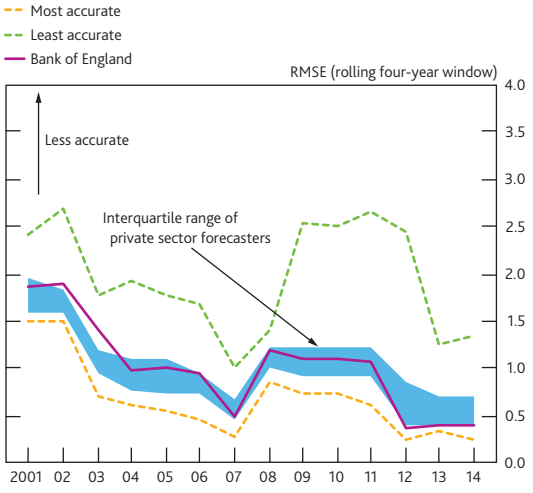
(a) See footnotes to **Chart 3.16**. Sample includes 17 private sector forecasters. For most of the period since the MPC's inception, the Bank has forecast the unemployment rate as measured by the Labour Force Survey. A small number of forecasts made prior to May 1998 used the claimant count measure. These have been excluded from our sample.

Chart 3.21 Two year ahead UK unemployment rate forecast accuracy (rolling four-year RMSEs)(a)



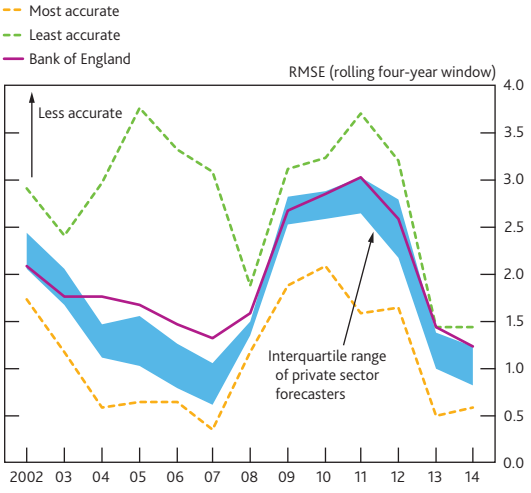
(a) See footnotes to **Charts 3.16** and **3.20**. Sample includes 17 private sector forecasters.

Chart 3.22 One year ahead UK consumption growth forecast accuracy (rolling four-year RMSEs)(a)



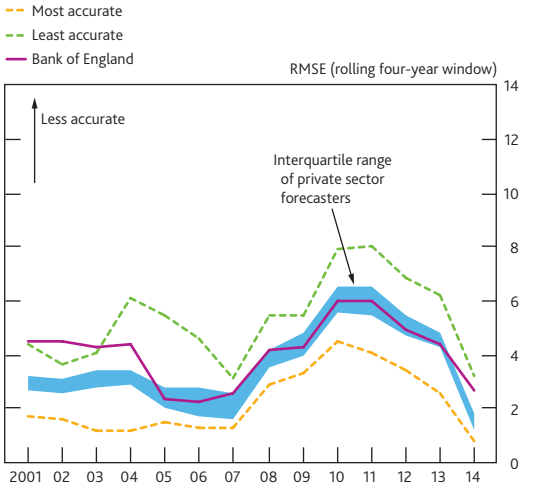
(a) See footnotes to **Chart 3.16**. Sample includes 20 private sector forecasters.

Chart 3.23 Two year ahead UK consumption growth forecast accuracy (rolling four-year RMSEs)(a)



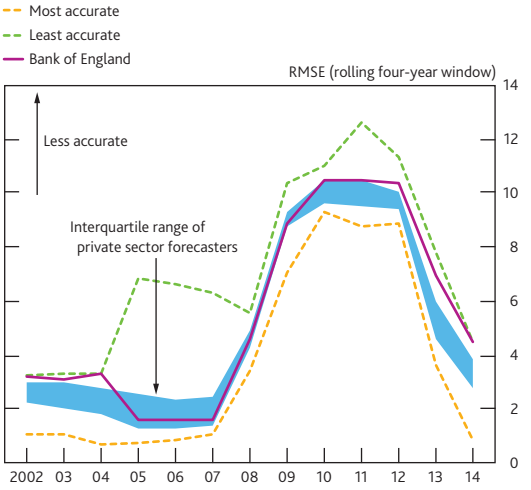
(a) See footnotes to **Chart 3.16**. Sample includes 19 private sector forecasters.

Chart 3.24 One year ahead UK investment growth forecast accuracy (rolling four-year RMSEs)(a)



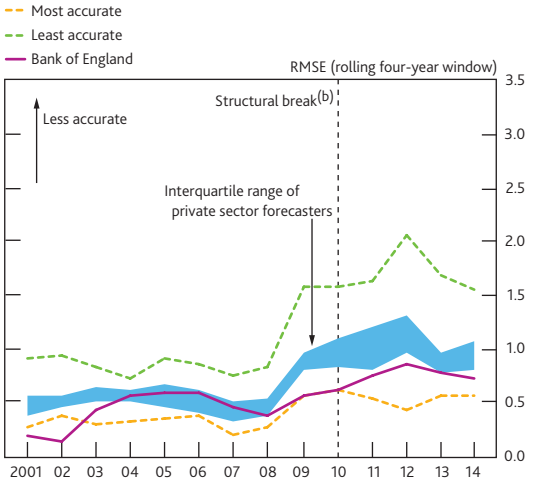
(a) See footnotes to **Chart 3.16**. Sample includes 19 private sector forecasters.

Chart 3.25 Two year ahead UK investment growth forecast accuracy (rolling four-year RMSEs)(a)



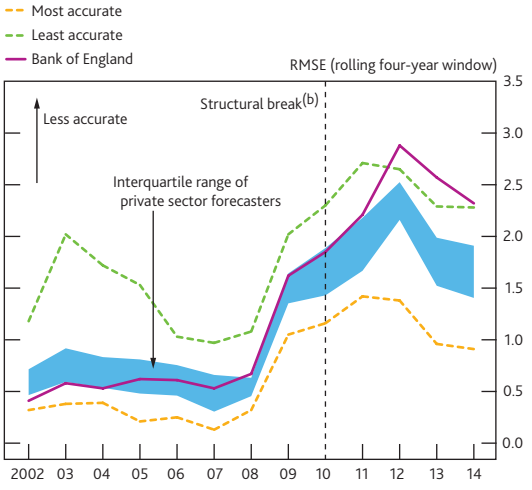
(a) See footnotes to **Chart 3.16**. Sample includes 17 private sector forecasters

Chart 3.26 One year ahead UK wage growth forecast accuracy (rolling four-year RMSEs)(a)



(a) See footnotes to **Chart 3.16**. Sample includes 16 private sector forecasters.
(b) The Office for National Statistics discontinued the Average Earnings Index in September 2010, after replacing it with the Average Weekly Earnings series in January 2010. This series break means that we are missing calendar-year forecast errors for 2010.

Chart 3.27 Two year ahead UK wage growth forecast accuracy (rolling four-year RMSEs)(a)



(a) See footnotes to **Chart 3.16**. Sample includes 18 private sector forecasters.
(b) The Office for National Statistics discontinued the Average Earnings Index in September 2010, after replacing it with the Average Weekly Earnings series in January 2010. This series break means that we are missing calendar-year forecast errors for 2010 and 2011.

Table 3.I Accuracy of one year ahead house price growth forecasts[a]

	RMSE		Absolute error	Average absolute error
Private sector forecasters[b]	2009–14	2010–14	2009	2010–14
1	5.5	3.1	11.5	2.4
2	6.4	4.3	11.6	3.8
3	6.7	4.9	11.5	4.4
4	7.8	4.3	16.5	3.6
5	8.9	5.0	17.3	3.9
6	12.2	8.5	21.3	7.5
Bank	**7.7**	**3.3**	**17.3**	**2.9**
Number of private sector forecasters with a higher RMSE or absolute error than the Bank	*3*	*5*	*1*	*5*

(a) Forecasts of annual house price growth are compared to the house price indices published by Nationwide and Halifax. The Bank forecasts the average of those house price indices. The other forecasters forecast one of the two indices. Each forecast is compared to the relevant house price index, or the average of the two indices as appropriate.
(b) Forecasters that report to the HMT Survey of External Forecasters, who forecast house prices based on either the Halifax or Nationwide house price indices, and who submitted at least 4 forecasts in the period 2009–14.

price indices, which were met by 6 individual forecasters. Given the very short sample period, we did not assess two year ahead forecasts.

Over the 2009 to 2014 period, the RMSE of the Bank's one year ahead house price inflation was similar to those of private sector forecasters in our sample. Excluding 2009 lowers the RMSES for all seven forecasters, particularly so for the Bank. Forecast errors for 2009 were markedly higher than those for other years; the absolute forecast error in 2009 on average across the private sector forecasters was around four times larger than the equivalent average absolute error over the 2010 to 2014 period; while for the Bank, the 2009 absolute forecast error was six times larger than the average across 2010 to 2014. Reflecting that, the Bank forecasts were relatively more accurate in the 2010 to 2014 period; on the RMSE measure Bank forecasts were more accurate than five of the six private sector forecasters in our sample. In contrast, the Bank's relative house price growth forecast accuracy deteriorates when the comparison period is expanded to include 2009, as Bank forecasts were more accurate than three of the six private sector forecasters.

3.6 International comparisons of forecast accuracy

This subsection describes our comparisons of forecast accuracy for five variables (see **Table 2.B** in Section 2.1) across the Bank and four other central banks; the European Central Bank, the US Federal Reserve, the Reserve Bank of New Zealand and the Swedish Riksbank. We conducted these comparisons in three ways, with accuracy being measured using RMSEs in each case:

- First, we compared central bank accuracy in forecasting domestic macroeconomic variables. For example, we compared the accuracy of Bank forecasts for UK GDP growth to the accuracy of ECB forecasts for euro-area GDP growth, the accuracy of Fed forecasts for US GDP growth, and so on. We used RMSEs scaled by the volatility of data outturns to control to some extent for the different economic shocks and economic structures characterising each economy. But scaled RMSEs cannot fully control for the different forecasting challenges facing central banks in different economies, leading us to our second set of comparisons.

- We next compared central bank forecasting accuracy relative to that of private sector forecasters in the same country for the same variable. For example, we compared the accuracy of Bank UK GDP forecasts relative to those of the UK private sector with the accuracy of ECB euro-area forecasts relative to those of the euro-area private sector. As discussed in Section 2.2.4, this should control for structural differences across economies as private sector forecasters face the same forecasting challenges as the central bank in the same country.

- In the final set of comparisons, we compared US and euro-area GDP growth forecasts made by the Bank to those made by the Fed (for the US) and the ECB (for the euro area). Although we compared forecasts for the same variable, we still used scaled RMSEs because we also compared across time periods.

Table 3.J Accuracy of one year and two year ahead calendar-year forecasts made by other central banks, relative to Bank forecast accuracy[a]

RMSE, scaled by standard deviation of data outturns, relative to scaled RMSE for Bank forecasts, 1998–2014[a]

	GDP growth		Inflation		Unemployment rate[b]		Consumption growth		Investment growth	
	One year ahead	Two years ahead	One year ahead	Two years ahead	One year ahead	Two years ahead	One year ahead	Two years ahead	One year ahead	Two years ahead
ECB[c]	0.8	1.1	0.7	0.6	n.a.	n.a.	1.2	1.3	0.5	1.0
Fed[d]	1.0	1.1	1.3	0.8	0.4	0.9	n.a.	n.a.	n.a.	n.a.
RBNZ	1.1	0.9	0.9	0.5	0.7	0.8	1.2	1.1	0.6	0.8
Riksbank[e]	1.0	0.9	1.0	0.7	n.a.	n.a.	1.8	1.5	0.7	0.9

RMSE, scaled by standard deviation of data outturns, relative to scaled RMSE for Bank forecasts, for one year ahead forecasts[a]

	GDP growth		Inflation		Unemployment rate[b]		Consumption growth		Investment growth	
	1998–2007	2010–14	1998–2007	2010-2014	1999–2007	2010–14	1998–2007	2010–14	1998–2007	2010–14
ECB[c]	0.4	0.7	6.4[f]	0.9	n.a.	n.a.	0.6	1.7	0.3	0.4
Fed[d]	0.7	1.3	1.6	1.5	0.2	0.5	n.a.	n.a.	n.a.	n.a.
RBNZ	0.5	1.6	1.0	0.7	0.2	2.9	0.6	3.9	0.4	0.6
Riksbank[e]	0.5	1.1	1.3	0.9	n.a.	n.a.	0.7	2.3	0.4	0.6

RMSE, scaled by standard deviation of data outturns, relative to scaled RMSE for Bank forecasts, for two year ahead forecasts[a]

	GDP growth		Inflation		Unemployment rate[b]		Consumption growth		Investment growth	
	1999–2007	2010–14	1999–2007	2010–14	2000–07	2010–14	1999–2007	2010–14	1999–2007	2010–14
ECB[c]	1.2	0.7	6.5[f]	0.5	n.a.	n.a.	1.2	0.9	0.8	0.7
Fed[d]	1.3	1.2	1.3	0.7	0.2	0.5	n.a.	n.a.	n.a.	n.a.
RBNZ	1.0	1.2	0.6	0.3	0.2	1.7	0.8	1.8	0.8	0.8
Riksbank[e]	0.8	0.6	0.6	0.6	n.a.	n.a.	0.7	0.8	0.7	0.7

(a) A ratio of more than 1 suggests that Bank forecasts were more accurate than those of other central banks (black font); a value of less than 1 suggests that Bank forecasts were less accurate than those of other central banks (grey font). Details of data sources are provided in Annex 3.
(b) A small number of forecasts made prior to May 1998 used the claimant count measure. These have been excluded from our sample.
(c) ECB forecasts are available from 2001 for one year ahead forecasts and 2002 for two year ahead forecasts. They are compared with Bank forecasts over the same period.
(d) Fed GDP and inflation forecasts relate to Q4 on Q4 percentage changes, which are compared with equivalent Bank forecasts; while Fed unemployment rate forecasts relate to Q4 of each year and these are also compared to equivalent Bank forecasts.
(e) Riksbank forecasts for consumption growth and investment growth are available from 2000 for one year ahead forecasts and 2001 for two year ahead forecasts. They are compared with Bank forecasts over the same period.
(f) The volatility of euro-area inflation was markedly lower in the 2001–07 period than seen in the other countries in our sample. As a result, the scaled RMSE for the ECB in that period was higher than for the other countries. The ECB's unscaled RMSEs in that period were much closer to those of the Bank. See **Table A.4.S** in Annex 4.

Accuracy of the Bank and four other central banks in forecasting domestic macroeconomic variables

Table 3.J compares (calendar year) forecasting accuracy for five domestic macroeconomic variables across the ECB, Fed, RBNZ and Riksbank. For each of those central banks and for each variable, the table shows RMSEs scaled by the standard deviation of the data outturns, relative to the Bank's own scaled RMSE. That means that if a ratio shown in **Table 3.J** is greater than 1, the Bank's forecasts have been more accurate than those of the peer central bank. If the ratio is less than 1, the Bank's forecasts have been less accurate. As noted above, we compared forecasts for domestic variables. For example, we compared Bank forecasts for UK GDP growth to ECB forecasts for euro-area GDP growth (scaled by the volatility of data outturns).

Over the full sample (1998–2014), these direct comparisons across central banks suggest that the Bank's UK GDP growth forecasts have tended to be broadly as accurate as domestic GDP growth forecasts from other central banks, at both the one and two year ahead horizons. The Bank's consumption growth forecasts have been more accurate across both forecast horizons. In contrast, for inflation, the unemployment rate and investment growth, Bank forecasts have tended to be less accurate than those from other central banks.

Comparing forecasting accuracy over time between the 1998–2007 and 2010–14 periods, the Bank's one year ahead GDP growth forecasts have tended to be relatively more accurate than those from other central banks. But for two year ahead GDP growth forecasts, the Bank's relative accuracy generally decreased. For inflation forecasts, the Bank's relative accuracy mostly decreased at both forecast horizons. In contrast, the relative accuracy of the Bank's unemployment rate and consumption growth forecasts mostly improved at both forecast horizons, while the relative accuracy of the Bank's investment forecasts was similar in both periods, at both forecast horizons.

As with the comparisons against private sector forecast performance, our international benchmarking exercise involved relatively small sample sizes, meaning that care should be taken when interpreting our results. Moreover, as set out in Section 2.2.4, while the scaling of RMSEs by the volatility of data outturns is a crude way of controlling for differences between countries (both differences in the nature of the economic shocks, and differences in the underlying structural characteristics of the economy), it is inevitably an imperfect way of conducting cross-country comparisons. In the next subsection, we provide an alternative and arguably superior set of cross-country comparisons, which compare the accuracy of domestic forecasts made by each central bank with the accuracy of forecasts for the same variable made by private sector forecasters in the same country; we additionally assess statistical significance of any differences in relative accuracy using Diebold-Mariano tests.

Accuracy of the Bank and four other central banks in forecasting domestic macroeconomic variables, compared with the accuracy of private sector forecasters

This subsection compares Bank forecast accuracy to that of other central banks, controlling for differences across economies by assessing how central bank forecasts have performed relative to forecasts made by private sector forecasters in the same economy.

Table 3.K shows the proportion of private sector forecasters in each economy whose forecasts were less accurate than the central bank's over the whole sample period. Accuracy is again measured using RMSEs. For the United Kingdom and euro-area, our samples of individual private sector forecasters range between 17 and 22, depending on the variable and forecast horizon. For New Zealand and Sweden the samples are smaller, from 7 to 12 forecasters, meaning that the results for these countries are less robust than for the United Kingdom and euro area. The private sector sample size for the US is similar to our samples for the United Kingdom and euro area, but the results are likely to be biased towards producing higher RMSEs for the Fed forecasts relative to the private sector forecasts. That is because, as discussed in Section 2.1, the Fed only publishes Q4 on Q4 forecasts, whereas the private sector forecasts are all on a calendar-year basis.

Table 3.K Accuracy of one year and two year ahead calendar-year forecasts made by central banks, relative to forecast accuracy of private sector forecasters, 1998–2014[a]

	% private sector forecasters in this comparison who are less accurate than the central bank				
	UK	US[b]	EA[c]	NZ	SW[d]
One year ahead forecasts					
GDP growth	67	0	38	91	57
Inflation	95	94	100	91	73
Unemployment rate	29	0	n.a.	80	n.a.
Consumption growth	35	n.a.	81	82	88
Investment growth	16	n.a.	86	91	100
Two year ahead forecasts					
GDP growth	23	16	16	58	100
Inflation	10	100	61	82	64
Unemployment rate	29	0	n.a.	55	n.a.
Consumption growth	32	n.a.	26	67	67
Investment growth	24	n.a.	63	17	67
Number of private sector forecasters	17–22	17–19	18–21	10–12	7–11

(a) For each central bank, their forecast accuracy (measured using RMSE) is compared with the accuracy of individual Consensus forecasters in the same country, forecasting the same variable. The source of private sector forecasts is Consensus Economics. Full details of data sources are provided in Annex 3.
(b) Fed forecasts relate to Q4 on Q4 percentage changes for GDP and inflation, and Q4 for the unemployment rate. They are compared with Consensus calendar-year forecasts. Consumption and investment growth forecasts are not published by the Fed.
(c) ECB forecasts are available from 2001 for one year ahead forecasts and 2002 for two year ahead forecasts. They are compared with Consensus forecasts over the same period. ECB unemployment forecasts have only been published since December 2013, so are excluded from this comparison.
(d) Unemployment forecasts are published by Riksbank but are not included in the Consensus survey, so a comparison is not possible.

The Bank's one year ahead inflation forecasts have been broadly as accurate, relative to private sector forecasts, as those produced by the other central banks. Interestingly, the Fed's one year ahead inflation forecasts were more accurate than those produced by 94% of the private sector forecasters, despite Q4 on Q4 forecasts being more affected by short-term data volatility than calendar-year forecasts. And the Bank's one year ahead GDP growth relative forecast accuracy has been greater than all the other central banks, except the RBNZ. In contrast, the Bank's one year ahead forecasts for the unemployment rate, consumption growth and investment growth have all been relatively less accurate than forecasts produced by the ECB, RBNZ and Riksbank; for these variables,

the Bank's forecasts have been more accurate than those produced by 16% to 35% of private sector forecasters, compared with 80% to 100% for the other three central banks. The Fed's one year ahead GDP and unemployment rate forecast accuracy appears to have been lower than all of the private sector forecasters in our sample, but as noted above, that is likely to underestimate the relative accuracy of the Fed's forecasts, because we cannot compare the forecasts on a comparable basis.

Across the central banks in our sample, the accuracy of two year ahead forecasts relative to private sector forecasts has, in most cases, been lower than the relative accuracy of the one year ahead forecasts. That was particularly the case for the Bank's inflation forecasts. The one year ahead forecasts were more accurate than forecast produced by 95% of the private sector forecasters in our sample, whereas for two year ahead forecasts the equivalent number was 10%. Consistent with the relative decrease in accuracy of the Bank's two year ahead inflation forecasts in **Chart 3.19**, this result was more pronounced in the post-crisis period.

The results in **Table 3.K** show the proportion of the private sector forecasters in each economy that have produced less accurate forecasts than the associated central bank forecasts, but they do not give any sense of how much less accurate those private sector forecasts have been. We conducted Diebold-Mariano tests, to assess whether differences in forecast accuracy between the central banks and private sector forecasters were statistically significant (at the 10% level). **Table 3.L** shows the proportion of private sector forecasters who produced forecasts significantly less or more accurate than the central bank forecasts. The small sample sizes involved mean that we cannot be fully confident in the robustness of these results. Nevertheless, **Table 3.L** still contains useful information.

In many cases, the forecast accuracy of private sector forecasters has not been statistically significantly different from that of central banks (as shown by the prevalence of zeros in **Table 3.L**). But there were some interesting differences in accuracy. Across all the central banks, except the Riksbank, their one year ahead inflation forecasts have been significantly more accurate than 12% to 62% of private sector forecasts, whereas no private sector forecaster has produced significantly more accurate forecasts than the central bank's forecasts. The pattern of results was broadly similar for the two year ahead inflation forecasts, although the Bank's two year ahead inflation forecasts have not been significantly more accurate than those produced by any private sector forecaster. More generally, the Bank's two year ahead forecasts have not been significantly more accurate than those produced by private sector forecasters for any of the five variables in our sample. In contrast, each of the other central banks has produced two year ahead forecasts significantly more accurate than those made by 5% to 36% of private sector forecasters for at least one variable.

Accuracy of the Bank, European Central Bank and US Federal Reserve in forecasting US and euro-area GDP growth

In our final set of international comparisons, we compared Bank forecasts for US and euro-area GDP growth to forecasts from the Fed (for US GDP) and from the ECB (for euro-area GDP). We found that Bank forecasts for Q4 on Q4 US GDP growth have tended to be as accurate as Fed ones at the one year ahead horizon (**Table 3.M**). At the two year ahead horizon, Bank forecasts have tended to be more accurate, although according to the Diebold-Mariano test described in Section 2.2, the difference in accuracy is only just significant at the 10% level when assessed over the full 2001–14 sample period. For both Bank and Fed forecasts, scaled RMSEs tended to be higher in the 2010–14 period than in the 2001–07 period. That reflects lower data volatility in the post-crisis period when US GDP growth was consistently weak by historical standards; unscaled RMSEs were the same or lower in the 2010–14 period compared to the earlier 2001–07 period for both central banks (see **Table A.4.T** in Annex 4).

The results in **Table 3.M** were robust to comparing Fed forecasts with Bank forecasts from November *Inflation Reports* instead of February ones. We compared Fed forecasts from the February Monetary Policy Reports (MPRs) and Bank forecasts from the February *Inflation Reports* because they are the first forecasts made when Q4 GDP data for the previous year is available. But since October 2007, the Fed has published both the biannual MPRs and the quarterly Summary of Economic Projections (SEPs). Since 2008, the projections published in the MPR in February have been identical to those in the winter SEPs in December or January, which means that the February MPR projections may be relatively out of date when compared with the February *Inflation Report* forecasts. As a robustness check, we also calculated the accuracy of Bank forecasts from November *Inflation Reports*. There is

Table 3.L Accuracy of one year and two year ahead calendar-year forecasts made by central banks, relative to forecast accuracy of private sector forecasters, 1998–2014, compared using a Diebold-Mariano test of forecast equivalence[a]

% private sector forecasters in this comparison who are significantly less accurate than the central bank (10% significance level)

	UK	US[b]	EA[c]	NZ	SW[d]
One year ahead forecasts					
GDP growth	0	0	0	18	14
Inflation	29	12	62	36	0
Unemployment rate	0	0	n.a.	0	n.a.
Consumption growth	5	n.a.	10	18	25
Investment growth	0	n.a.	14	36	0
Two year ahead forecasts					
GDP growth	0	0	5	0	33
Inflation	0	32	0	36	18
Unemployment rate	0	0	n.a.	0	n.a.
Consumption growth	0	n.a.	5	8	0
Investment growth	0	n.a.	11	0	22

% private sector forecasters in this comparison who are significantly more accurate than the central bank (10% significance level)

	UK	US	EA	NZ	SW
One year ahead forecasts					
GDP growth	0	71	19	0	0
Inflation	0	0	0	0	0
Unemployment rate	0	47	n.a.	0	n.a.
Consumption growth	10	n.a.	0	0	0
Investment growth	16	n.a.	0	0	0
Two year ahead forecasts					
GDP growth	14	0	5	8	0
Inflation	0	0	6	0	0
Unemployment rate	0	16	n.a.	9	n.a.
Consumption growth	11	n.a.	26	0	0
Investment growth	24	n.a.	0	0	0
Number of private sector forecasters	17–22	17–19	18–21	10–12	7–11

(a) For each central bank, their forecast accuracy (measured using RMSE) is compared with the accuracy of individual Consensus forecasters in the same country, forecasting the same variable. To assess the statistical significance of differences in accuracy between forecasters, we used a Diebold-Mariano test, using HAC standard errors. The percentage of private sector forecasters who are significantly less and significantly more accurate than the central bank does not sum to 100 because in many cases, the difference in forecast accuracy between the central bank and the private sector forecaster was not statistically significant. The source of private sector forecasts is Consensus Economics. Full details of data sources are provided in Annex 3.
(b) Fed forecasts relate to Q4 on Q4 percentage changes for GDP and inflation, and Q4 for the unemployment rate. They are compared with Consensus calendar-year forecasts. Consumption and investment growth forecasts are not published by the Fed.
(c) ECB forecasts are available from 2001 for one year ahead forecasts and 2002 for two year ahead forecasts. They are compared with Consensus forecasts over the same period. ECB unemployment forecasts have only been published since December 2013, so are excluded from this comparison.
(d) Unemployment forecasts are published by Riksbank but are not included in the Consensus survey, so a comparison is not possible.

Table 3.M Accuracy of one year and two year ahead US GDP growth forecasts made by the Bank and the Fed[a]

RMSEs scaled by the standard deviation of data outturns

	2001–14		2001–07		2010–14	
	One year ahead	Two years ahead	One year ahead	Two years ahead	One year ahead	Two years ahead
February Monetary Policy Reports (Fed)	0.9	1.2	1.0	1.2	1.6	2.3
February *Inflation Reports* (Bank)	0.9	1.0	1.0	0.9	1.7	1.4
November *Inflation Reports* (Bank)[b]	0.8	1.1	1.0	0.7	0.8	1.2

(a) Fed forecasts are for Q4 on Q4 growth, they are compared with equivalent Bank forecasts. Bank forecasts for US GDP growth are available from 2001. Fed two year ahead forecasts come from February Monetary Policy Reports from 2005 onwards and from July Monetary Policy Reports prior to 2005. See **Table A.3.C** in the Annex for more details. Forecast accuracy is measured by Root Mean Squared Errors, scaled by the standard deviation of data outturns. A higher scaled RMSE indicates a less accurate forecast after controlling for data volatility.
(b) Sample period for forecasts made in November *Inflation Reports* is 2002–14 for one year ahead forecasts and 2003–14 for two year ahead forecasts.

little difference in accuracy between the November and February *Inflation Report* forecasts, except for one year ahead forecasts in the 2010–14 period where the discrepancy predominantly reflects one overly optimistic upwards revision to the forecast for 2011 Q4 growth between the November 2010 and February 2011 *Reports*.

Turning to the euro-area, Bank forecasts for calendar-year euro-area GDP growth have tended to be as accurate as ECB ones at both the one year and two year ahead horizons (**Table 3.N**). Indeed, to one decimal place, the scaled RMSEs of forecasts made at the two central banks have been almost identical across forecast horizons and time periods. Comparing the 2001–07 and 2010–14 periods, scaled RMSEs tended to be higher in the post-crisis period for both central banks, but only for the two year ahead forecasts. The changes in scaled RMSEs between the two periods were less marked than for the US GDP growth forecasts. In this case, the rise in scaled RMSEs reflected a rise in unscaled RMSEs rather than a fall in data volatility (see **Table A.4.U** in Annex 4). Again we conducted a robustness check by comparing ECB forecasts published in March with Bank forecasts from February *Inflation Reports* as well as those from the May *Reports*. We found the accuracy of Bank forecasts of euro-area GDP growth has not tended to vary much between the February and May *Reports*.[1]

Table 3.N Accuracy of one year and two year ahead euro-area GDP growth forecasts made by the Bank and the European Central Bank[a]

	RMSEs scaled by the standard deviation of data outturns					
	2001–14		2001–07		2010–14	
	One year ahead	Two years ahead	One year ahead	Two years ahead	One year ahead	Two years ahead
March macroeconomic projections (ECB)	0.4	1.1	0.6	1.2	0.5	1.3
May *Inflation Reports* (Bank)	0.4	1.1	0.6	1.2	0.5	1.4
February *Inflation Reports* (Bank)	0.5	1.1	0.8	1.2	0.4	1.1

(a) Forecasts are for calendar-year growth. Forecast accuracy is measured by Root Mean Squared Errors, scaled by the standard deviation of data outturns. A higher scaled RMSE indicates a less accurate forecast after controlling for data volatility. ECB forecasts are available from 2001 for one year ahead forecasts and 2002 for two year ahead forecasts.

(1) ECB forecasts published in March are the first to be made after the estimate of euro-area GDP for Q4 of the previous year has become available. The appropriate comparison is with Bank forecasts from the May *Inflation Reports*, which are made before the Q1 euro-area GDP estimate has been published. But to check that this does not give the Bank a significant informational advantage we also calculated the accuracy of Bank forecasts from February *Reports*.

Box 4
Assessing the Bank's fan charts

The Bank publishes its forecasts of inflation and UK GDP growth as probability distributions — so-called 'fan charts'. As discussed in Section 2.3, the fans emphasise the inevitable uncertainty around the outlook for the economy. It is possible that by assessing the accuracy of Bank forecasts in terms of point estimates, as we have done throughout the rest of this evaluation, we underestimate the ability of forecasters to anticipate economic shocks, which they may have factored into the wider forecast probability distributions. However, as noted in Section 2.1, fan charts are only available over a sufficiently long period for inflation and UK GDP growth.

In this box, we present the results from statistical tests that we conducted to assess the accuracy of the UK GDP growth and inflation fan charts. These tests assess how close the distribution of data outturns appears to have been relative to the one implied by the fan charts. We found that for GDP growth forecasts the distribution of outturns has been consistent with the fan charts up to two or three quarters ahead, but for longer horizons the outturns appear to have come from a different distribution to the fan charts. For inflation, the tests generally suggested that outturns have come from a different distribution to the fan charts.

Methodology

Evaluating probability distributions is difficult because we never observe the true distribution, but only one realisation from it. A common approach is to calculate how likely it would be, with a given forecast density, to observe the realised value at a particular point in time. This probability is known as the Probability Integral Transform (PIT), which is then used to assess how close the forecast distribution is to the actual distribution of a variable.

The PITs of a forecast distribution are given by:

$$z_t = \int_{-\infty}^{x_t} f(u)du \qquad \textbf{(1a)}$$

where f*()* is the forecast density and x_t are the data outturns of the variable under consideration.

If a set of forecast densities offered a good approximation of the true underlying density, then the PITs should be evenly distributed over all the percentiles. In other words, we would expect an outturn to occur as frequently in practice as the forecast predicted in theory. Formally, this implies a uniform distribution of the PITs between 0 and 1, such that z_t has equal probability of being drawn anywhere between 0 and 1. Although it is possible to apply such tests of uniformity, in practice it can be more effective to transform the data points and instead test that they are normally distributed. We denoted this transformed series as z_t^*

The first test of normality that we conducted was a variant of the Berkowitz test (Berkowitz (2001)).(1) We performed a $\chi^2(2)$ likelihood ratio test where the errors ε_t in the regression

$$z_t^* = \rho z_{t-1}^* + \varepsilon_t \qquad \textbf{(1b)}$$

are normally distributed with zero mean and variance one.

The second test that we conducted came from Doornik and Hansen (2008). This complemented the Berkowitz test, which tests the first and second moments of the empirical distribution, by testing the $\{z_t^*\}_{i=1}^{N}$ series through the third and fourth moment of the empirical distribution.

Both of the tests that we conducted are tests of equality of distributions. We tested under the null hypothesis that the distribution of the outturns was equal to the one implied by the fan charts, against the alternative hypothesis that they were unequal. **Table 1** reports p-values from both of the tests. For short-horizon GDP growth forecasts, up to two or three quarters ahead, we could not reject the null hypothesis at the 5% significance level that the distribution of outturns has been consistent with the fan charts distributions. In contrast, for horizons beyond three quarters, we rejected the null hypothesis (at the 5% level), indicating that

(1) For more information on the details of this test, see Clements (2004), Elder *et al* (2005) and Boero, Smith and Wallis (2008). The Berkowitz test was performed with 2 degrees of freedom.

UK GDP growth outturns have tended to come from different distributions to the fan charts. For inflation forecasts, the tests suggested that the outturns have been consistent with the fan charts only at the two-quarter horizon; for all other horizons, at least one of the tests resulted in us rejecting the null hypothesis at the 5% significance level.

Table 1 Estimated p-values from tests of the Bank's fan charts[a]

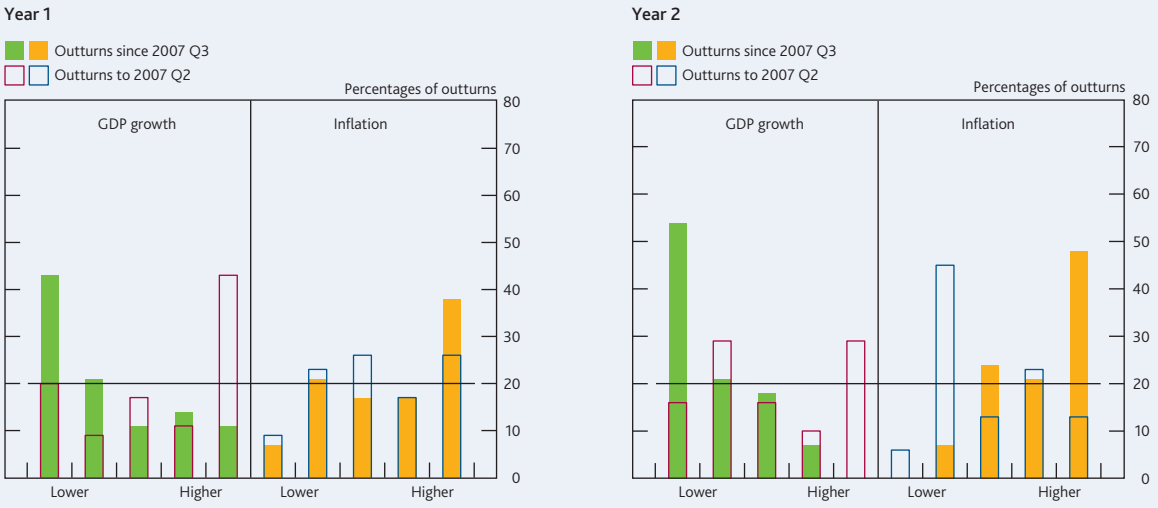
	Inflation		GDP	
Quarter	Berkowitz	Doornik and Hansen	Berkowitz	Doornik and Hansen
1	0.00**	0.00**	0.64	0.12
2	0.05*	0.62	0.07*	0.05*
3	0.00**	0.01**	0.08*	0.00**
4	0.00**	0.03**	0.00**	0.00**
5	0.00**	0.19	0.00**	0.00**
6	0.00**	0.07*	0.00**	0.00**
7	0.00**	0.00**	0.00**	0.00**
8	0.00**	0.00**	0.00**	0.00**
9	0.00**	0.00**	0.00**	0.00**

(a) Inflation forecasts include both RPIX and CPI forecasts. The sample includes inflation forecasts made between August 1997 and August 2014 and GDP forecasts made between November 1997 and August 2014. P-values shown in the table are based on the Berkowitz and Doornin Hansen tests as defined earlier in relation to equation **(1b)**. 5% significance is denoted by ** and 10% significance is denoted by *.

The results from the statistical tests presented here do not tell us how the distributions of inflation and GDP growth outturns have differed from the fan chart distributions. Further investigation indicated that inflation outturns have tended to be concentrated in the upper quintile of the fan chart distributions, and GDP growth outturns in the lower quintile, especially since mid-2007, as illustrated in **Charts A** and **B**.

Charts A and **B** show the proportion of inflation and GDP outturns in each quintile of the fan chart distributions at the one and two-year horizons. If the fan charts accurately described the uncertainty faced by the Bank and the samples were sufficiently large, then outturns would be expected to lie evenly across the fan chart distributions; in other words, 20% of outturns would be expected to lie within each quintile of the distribution — illustrated by the black solid lines. In fact, since the onset of the financial crisis there seems to have been a tendency to understate the probability of low outturns for GDP growth, and high outturns for inflation. That observation is not surprising given the wider economic context. The financial crisis, and its associated effects on credit conditions and uncertainty, triggered unexpectedly large and persistent declines in GDP growth, both in the United Kingdom and elsewhere. At the same time, the United Kingdom faced a series of unexpected cost shocks — including changes to VAT, increases in energy prices and rises in non-energy import costs — that drove inflation higher. For further discussion, see, for example, Hackworth, Radia and Roberts (2013).

Charts A and B Dispersion of inflation and UK GDP growth outturns across the fan charts[a]



(a) Dispersion of GDP growth and inflation outturns across the quintiles of the fan chart distributions. Calculated for the GDP growth fan charts published from February 1998 to August 2013 and for the inflation fan charts published from February 1998 to November 2013. The outturns for GDP growth and inflation are allocated to one of five buckets representing the quintiles of the fan chart from a year earlier. The modes of the fan chart distributions for GDP growth up to the August 2011 forecasts have been adjusted up by 0.3 percentage points, to reflect the effects of methodological changes implemented in the 2011 edition of the *Blue Book*. Inflation fan charts refer to RPIX inflation up to November 2003 and CPI inflation thereafter.

4 Conclusions and recommendations

This evaluation has assessed the Bank's forecasting performance along a number of dimensions, including benchmarking performance against a simple model, as well as against private sector forecasters and other central banks. This section summarises the key results for each forecast variable, discusses some common themes arising from our work and sets out some recommendations for improving particular aspects of Bank forecasting performance.

As set out in Section 2, for some parts of our evaluation exercise, in particular comparisons with private sector and other central bank forecasts, as well as the assessment of post-crisis performance, we had a relatively short backrun of data with which to work. This is important to bear in mind when interpreting our empirical results and forming recommendations. Nevertheless, there are sufficient common themes from the overall evaluation for some general conclusions to be drawn.

In setting out our conclusions and recommendations, we would note that this is necessarily a review of past performance. For two year ahead forecasts, for example, our sample included forecasts made in *Inflation Reports* up to and including 2013.(1) The Bank's forecast processes have changed in various ways since the 2012 Stockton Review. It is too soon to assess the impact of these changes, and there would therefore be merit in repeating an evaluation exercise of a similar nature in a few years' time.

4.1 Summary of empirical findings by variable

UK GDP growth

The Bank's UK GDP growth forecasts, especially the one quarter and one year ahead forecasts, generally performed well against our tests. There was no statistically significant evidence of bias at the 10% level at any forecast horizon (when assessing forecasts against the 'latest' vintage of data);(2) this finding was robust to using mean or modal projections. Neither could we reject the null hypothesis of strong efficiency at the one-year horizon. And, for the most part, final revisions to UK GDP growth forecasts appear to have been unrelated to previous revisions.

At the two-year horizon, however, we noted a tendency for forecasts of GDP growth to cluster at around the pre-recession (1997–2007) average growth rate of approximately 3%. Errors in two year ahead GDP forecasts were positively and significantly related to data outturns known when the forecasts were made; this finding was robust to the use of real-time data, rather than the latest vintage of data, and to the use of mean rather than modal forecasts.

The Bank's UK GDP growth forecasts tended to be broadly as accurate as forecasts of domestic GDP growth from other central banks, at both the one and two-year horizons. And one year ahead UK GDP forecasts were at least as accurate as those produced by the majority of the private sector institutions in our sample. As with other variables, the accuracy of the Bank's two year ahead UK GDP growth forecasts tended to be lower in the post-crisis period. That could be because the UK macroeconomy became more difficult to forecast in 2010–14 relative to the pre-2008 period, consistent with the generally lower level of accuracy of UK private sector forecasts in the latter part of our sample.

(1) As described in Section 2.1, we define forecast horizon in terms of data availability. So, for example, a two year ahead inflation forecast for calendar-year 2014 will have been made in February 2013, because that was the first *Inflation Report* where inflation data outturns to 2012 Q4 were available.
(2) We did find evidence of bias when evaluating GDP projections against real-time data, although that is consistent with MPC's stated aim of forecasting mature vintages of data, and the observation that, historically, there has been a tendency for real-time GDP data to be revised higher over time.

Our assessment of the performance of the Bank's GDP fan charts found that, up to around two to three quarters ahead, the distribution of GDP outturns was consistent with that implied by the fan charts. For longer horizons, however, our statistical tests suggested that the GDP outturns came from a different distribution to the fans.

Inflation

There was no statistically significant evidence of bias in the Bank's inflation projections at the 10% level at any forecast horizon. And we could not reject the null hypothesis of strong efficiency at the one-quarter and one-year horizons. We did, however, note a tendency to underreact to data news when revising one year ahead inflation forecasts.

We found statistically significant evidence of inefficiency at the two-year horizon for the inflation forecasts, where forecast errors were systematically related to data outturns known at the time the projections were made. More generally, we noted a systematic tendency for two year ahead inflation projections to cluster around, or a little below, the inflation target. As discussed below, one interpretation of this is a tendency for two year ahead forecasts to exhibit too little persistence following economic shocks.

In our benchmarking exercise, the one year ahead inflation forecasts were typically at least as accurate as those produced by the majority of the UK private sector institutions in our sample. And that performance relative to the private sector compared favourably to the relative performance of the other central banks (for example, the Bank's one year ahead inflation forecasts were significantly more accurate than those produced by 29% of UK private sector forecasters in our sample, while the equivalent estimates for the other four central banks range from 0% to 62%.) The two year ahead projections compared less favourably to private sector forecasters, particularly in the post-crisis period when the Bank's relative forecast accuracy declined, although our sample period is too short to determine whether this change in relative performance was systematic.

Unemployment rate and wage growth

Over our sample period as a whole, Bank forecasts for both the unemployment rate and wage growth were systematically too high; that bias was statistically significant at the one and two year ahead horizons for both variables. For unemployment, that finding in part reflected large forecast errors made in the early part of the sample period: when we excluded forecasts made for 1998–2000, we could not reject the hypothesis that unemployment rate forecasts (at all horizons) were unbiased (at the 5% or 10% significance level). Our observation of statistically significant bias in the wage growth forecasts was robust to the exclusion of the 1998–2000 period, however. In terms of efficiency, we found statistically significant evidence that forecast errors were related to past data outturns for unemployment rate projections (at all horizons), and for two year ahead forecasts for wage growth. In contrast, final revisions to forecasts for the unemployment rate and wage growth appear, for the most part, to have been unrelated to previous revisions, although our results indicated a tendency to overreact to data news when revising one year ahead wage growth forecasts.

In terms of comparisons with simple statistical benchmarks, the unemployment rate was the only variable tested where Bank forecasts (at all horizons) were consistently (and statistically significantly) less accurate than random walk generated forecasts. Again, this result was no longer statistically significant when we excluded forecasts made for 1998–2000, but we also note that using that shortened sample resulted in the random walk generated wage growth forecasts being significantly more accurate than Bank forecasts at the two-year horizon.

Despite the observations of bias noted above, Bank forecasts for wage growth typically compared favourably with those of UK private sector forecasters at the one-year horizon. Our empirical work points to a possible decline in the relative accuracy of two year ahead Bank wage growth forecasts in the 2010–14 period, but we are sceptical about the robustness of this finding because of the structural break in the official wage data series in 2010.

In contrast, Bank unemployment rate forecasts, at both the one and two year ahead horizons, were consistently less accurate than forecasts produced by the majority of the private sector forecasters in our sample, although there were improvements in the Bank's relative accuracy in 2013 and 2014. The Bank's unemployment rate forecasts have also been less accurate than the unemployment rate forecasts produced by the Fed and RBNZ for the United States and New Zealand respectively, although some improvement in the Bank's relative accuracy was apparent in the 2010–14 period compared with 2000–07.

Consumption and investment growth

As with the Bank's UK GDP growth forecasts, there was no statistically significant evidence of bias at the 10% level at any forecast horizon for either consumption or investment growth forecasts. For the most part, final revisions to forecasts appear to have been unrelated to previous revisions. But, at some horizons, there was statistically significant evidence that forecast errors were related to past data outturns or past forecast errors.

Over the full sample, and when compared with other central bank forecasts of domestic spending, the Bank's investment growth forecasts have tended to be less accurate, especially at the one-year horizon, while consumption growth forecasts have tended to be more accurate at both horizons. However, the Bank's forecast accuracy relative to private sector forecasts for both consumption and investment growth has been lower than the accuracy of other central banks relative to their respective private sector forecasting communities.

House price inflation

The Bank's house price inflation forecasts have typically been too low. Over the full sample, that bias was statistically significant at the 5% level at all horizons and was the largest bias of all the variables we tested at the two-year horizon. We also found that two year ahead house price growth forecast errors were related to past data outturns and past forecast errors; these results were significant at the 5% level. Despite this, the accuracy of the Bank's house price growth forecasts in recent years has compared favourably to that of private sector forecasts. The short sample of private sector forecasts available (2009–14) means that we are unable to draw strong conclusions from this comparison. But with the exception of 2009, the Bank's one year ahead forecasts have been among the most accurate.

US and euro-area GDP growth

Bank forecasts of US GDP growth have tended to be too high when evaluated against the latest vintage of data. At the one quarter ahead horizon that bias was significant at the 10% level; at the two-year horizon it was statistically significant at the 5% level. That bias was not statistically significant when Bank forecasts were evaluated against real-time data, however. For euro-area GDP growth forecasts, we found statistically significant evidence of bias only at the one quarter ahead horizon (for latest vintage data), and at the two-year horizon (for real-time data). In terms of efficiency, forecast errors were significantly related to past data outturns and past forecast errors for both US and euro-area GDP growth at some horizons. Nevertheless, the Bank's forecasts have been as accurate as those made by the Fed and the ECB.

Household and corporate lending growth

Compared with the other variables in our sample, the Bank's one year ahead corporate lending growth forecasts have been the least accurate. One year ahead household and corporate lending growth forecasts have both tended to be too high, but that bias was only statistically significant for corporate lending. We rejected the hypothesis of strong efficiency at the 5% level in half of the specifications tested, for both variables. We reiterate, however, that the relatively short samples of household and corporate lending growth forecasts available mean that we cannot be confident in the robustness of our results. Unlike other variables in our sample, we did not evaluate the performance of two year ahead projections for household and corporate lending, due to the relatively short sample periods under consideration (data were available from 2007 for household lending, and from 2009 for corporate lending).

4.2 Common themes

Our evaluation highlighted a number of areas where the Bank's forecasts performed well against our tests. These included the one year ahead forecasts for UK GDP growth and inflation, where we found no significant evidence of bias, and where the accuracy of Bank forecasts compared favourably to those of the private sector and the four other central banks included in our evaluation exercise. We also note that, although the Bank's one year ahead forecasts of wage growth have exhibited statistically significant bias (being systematically too high), when compared with UK private sector forecasters these Bank forecasts have been consistently among the most accurate in our sample.

However, our evaluation also highlighted two key areas of Bank forecasting, which merit further investigation by Bank staff and the MPC with a view to improving forecasting performance.

First, the performance of two year ahead forecasts contrasted with that of the shorter-horizon forecasts, especially for inflation and, to a lesser extent, UK GDP growth. Forecasting accuracy of the two year ahead forecasts apparently decreased across most of the macroeconomic variables in our sample in the 2010–14 period relative to the pre-2008 period, whereas the accuracy of shorter-horizon forecasts was generally little changed. The relatively small sample sizes mean that we need to interpret these results with caution, but the apparent generalised nature of the decrease in accuracy is nevertheless noteworthy. In part, this apparent decline in accuracy is likely to reflect the inherent difficulties of forecasting the UK economy in the post-crisis period, as illustrated in the broad-based deterioration in the accuracy of UK private sector forecasts. For two year ahead inflation, however, the accuracy of the Bank's forecasts also appeared to decline relative to that of the UK private sector forecasting community.

Identifying the drivers of the apparent decrease in forecasting accuracy of the Bank's two year ahead inflation forecasts is not straightforward, and a detailed assessment is outside the scope of this project. In particular, it is not clear how much reflects systematic forecast issues, as opposed to a few unusually inaccurate forecasts. Nevertheless, our evaluation has produced evidence that is consistent with too little persistence having been built into two year ahead inflation forecasts (and to some extent two year ahead UK GDP growth forecasts) following economic shocks.

We found, for example, that while shorter-horizon inflation forecast errors were unrelated to past data outturns, the two-year inflation forecast errors were positively and significantly (at the 5% level) related to past outturns. As such, two year ahead inflation forecasts tended to be too high when past outturns were low (we found the same pattern of results for UK GDP growth). We also noted a clustering of two year ahead inflation forecasts around, or a little below, target, and a clustering of two year ahead GDP growth forecasts around the 1997–2007 average growth rate. That could point to a systematic tendency to assume that the impact of shocks should have faded by the two-year horizon.

It is perhaps unsurprising that the Bank's two year ahead inflation forecasts cluster near to the target. It is generally accepted that monetary policy has its peak impact on inflation after 18 months to two years;[1] moreover, as inflation is the Bank's target variable, the interpretation of published two year ahead projections can be complicated by considerations of communication and monetary strategy. Indeed, clustering of two year ahead inflation forecasts is a common feature across the central banks in our sample (see **Charts 4.27** to **4.30** in Annex 4), especially those with narrower targets.[2] Nevertheless, the Bank may be able to improve its two year ahead forecasting accuracy by more actively challenging the assumption that inflation (and UK GDP growth) will most likely be brought back to near the target (around pre-crisis average growth) within two years. In this context, it is notable that a recent *Bank of England Working Paper* (Fawcett *et al* (2015)) found that a relatively judgement-free version of the Bank's central forecasting model, known as COMPASS, has tended to outperform the MPC's inflation forecasts at horizons longer than one year. (They also found that such longer-horizon forecasts for UK GDP growth from the Bank's suite of statistical models were more accurate than the MPC's forecasts.)[3] There would therefore potentially be merit in Bank staff and the MPC placing more weight on such statistical models when considering the likely persistence of shocks to inflation (and UK GDP growth); see 'Recommendations' below.

The second area of note is the Bank's forecasting record on unemployment. Of the variables we evaluated, the unemployment rate forecasts stood out as performing the least well against our statistical tests. Bank unemployment rate forecasts have been systematically too high, both over our sample as whole (although the bias was not statistically significant when we excluded forecasts prior to 2001), and in the post-crisis period. Our efficiency tests indicated that unemployment rate forecasts could have been improved by more effectively incorporating information from past outturns. And, over the sample period as a whole, the accuracy of the Bank's unemployment forecast did not compare favourably to that of a simple statistical rule, other private sector forecasters or other central banks in our sample. There has been some improvement in the Bank's accuracy

(1) See, for example, Carney (2015).
(2) Three of the central banks in our sample have implicitly or explicitly targeted an inflation rate close to 2%, while the Reserve Bank of New Zealand has targeted a band of 1%–3% since 2002 and a band of 0%–3% prior to that.
(3) Kapetanios, Labhard and Price (2008) describes the Bank's suite of statistical models.

relative to private sector forecasters and other central banks more recently, although it is too early to judge whether that improvement is systematic.

As discussed in Section 3.2, the historical difficulties with the Bank's forecasting of the unemployment rate, in particular the persistent bias, may have in part been symptomatic of structural change in the labour market. For example, the relatively large errors in unemployment forecasts in the early part of the sample period coincided with what now appears to have been a decline in the natural rate of unemployment (linked in turn, at least in part, to the marked decline in union coverage over the previous decade).(1) More recently, the Bank, along with other forecasters, did not foresee the degree to which labour productivity would be persistently impaired by the financial crisis and its repercussions. At least in the short run, weaker-than-expected productivity growth would typically be associated with higher employment (and lower unemployment) outturns than expected.

Persistent changes in variables such as the natural rate of unemployment or productivity growth are inevitably difficult to identify in real time. Nevertheless, as we discuss further below, making more systematic use of statistical indicators and models may be one way of providing effective 'early warning' signals to flag the risk of underlying change.

4.3 Recommendations

The IEO was established both to support Court in its oversight of the Bank's performance and to help embed a culture within the Bank of learning and self-reflection. In that context, we offer some reflections on changes to forecast processes and approach that could assist the Bank in its desire to deliver improvements in its forecasting performance. The recommendations set out below were based on the results presented in this paper, but also on discussions that the project team had with producers and users of the Bank's macroeconomic forecasts during the course of the forecast evaluation.

When setting out these recommendations, we recognise that there is a balance to be struck between forecasting models and approaches that help form a coherent economic narrative (such as structural models that incorporate a large degree of forecaster judgement) and forecasting models that are optimised for empirical performance (such as pure statistical models). As set out in Section 1, the purpose of forecasting within any central bank is far broader than the production of numerical point forecasts for macroeconomic variables alone. Consequently, while one theme of our recommendations is that there is a case for making more systematic use of statistical models in the Bank's forecast process, this should not detract from the importance of using structural models and forecaster judgement to construct a coherent and meaningful economic narrative for the Bank's projections.

When formulating our recommendations, we were also mindful of the need to strike an appropriate balance between providing concrete recommendations for change, but not taking an overly prescriptive and detailed approach. We have therefore identified four broad (and not necessarily mutually exclusive) areas for potential process improvement, and have, for each, set out some high-level initiatives that the Bank may wish to take forward. These four areas are: learning more from other forecasters and models; learning more systematically from the past; challenging conventions more; and providing more support for internal non-MPC users of Bank forecasts.

First, we note that, in some areas of our evaluation, the accuracy of a subset of Bank forecasts did not compare favourably to those of other forecasters, or to alternative modelling approaches. This was the case, for example, for some of the Bank's two year ahead projections in the post-crisis period (such as inflation), as well as, at least until very recently, unemployment. We therefore recommend the Bank take steps to **learn more from other forecasters and models**. Specifically, we recommend identifying other central banks and private sector institutions with a track record of producing more accurate forecasts than the Bank (recognising that their future forecasting performance may not be as good) and investigating whether there are any key differences in forecast approach that could be incorporated into the Bank's processes. Particular attention should be given to those forecast variables where the Bank's relative forecasting performance appears to have decreased in recent years.

(1) Additionally, as set out in Section 2, it is likely that the quality of internally held staff projections, such as those for the unemployment rate, has improved during our sample period, in part reflecting a greater degree of internal scrutiny, both by Bank staff and the MPC.

More generally, our work suggests a strong case for investigating (perhaps in consultation with external forecasters and/or academics) how best to analyse systematically alternative statistical models, and to use insights from those models to challenge Bank staff and MPC forecast judgements. That exercise may be especially worthwhile for longer-horizon forecasts and for variables where structural change may be occurring. This recommendation does not presume that new models should be developed — the Bank has a wide array of statistical models already at its disposal (such as those in the Bank's suite of models) — rather that there would be benefit from finding ways to more effectively use insights from statistical models in the forecasting process. As Professor James Mitchell, our external peer reviewer, notes in his statement (Annex 2): 'an active area of research is how best to use or combine the insights from many different models/data series/forecasts, especially when these differ in their forecast narrative'.

A recurring theme of our work was the value in **learning more systematically from the past**. Judging the degree to which an underlying economic 'shock' is likely to persist, or identifying in real time a structural shift in an economy, are inherently difficult challenges. But we identified scope to take greater steps to hard wire in 'early warning' signals into the Bank's forecasting processes and to ensure that institutional knowledge about problematic areas of past forecast performance is effectively maintained.

We note, for example, that the monitoring of absolute forecasting performance statistics (such as bias and inefficiency) has, historically, been episodic at the Bank. We therefore recommend more regular and routine monitoring of absolute forecasting performance statistics, and routine investigation into any apparent difficulties in forecast performance that these statistics suggest. We would recommend such monitoring to be reported to the Bank's Court on a regular basis.

Additionally, we see merit in Bank staff regularly comparing forecasts with historical tendencies towards bias (such as those for house price growth, the unemployment rate, and wage growth) with bias-adjusted profiles. The aim would be to provide another avenue for challenging forecast judgements, and to help to maintain institutional knowledge (for example, about historical difficulties in forecasting the unemployment rate).

As set out in the Summary, this evaluation exercise should be seen as complementary to the broader assessment of the MPC's forecasting capability conducted by Stockton (2012). A core theme of the Stockton Review was the need to take steps to **introduce greater challenge to forecast conventions** and Bank thinking. In the wake of the Stockton Review, considerable changes to forecast processes were made to improve both engagement with the external community and forecast transparency. The project team recognises that it is too soon to assess properly the results of these changes. Nevertheless, we note that some of our empirical findings are consistent with a systematic tendency to assume that the impact of economic shocks fades relatively quickly, and there may be merit in considering whether further challenge needs to be generated in this particular area.

Finally, we note that the Bank's powers and responsibilities have changed greatly in the post-crisis period, with the institution gaining significant new statutory responsibilities, including for macroprudential policy and microprudential supervision. Increasingly, the work of the Bank's macroeconomic forecasters — historically focussed on supporting the monetary policy process — is proving to be an important input into these broader areas of the Bank's work. One aim of the evaluation exercise was to assess the performance of forecasts such as the unemployment rate, house price inflation and lending growth, which are particularly pertinent to the work of the Bank's other policy committees — namely the FPC and the PRAB.

During the course of our work, we found that there was scope to **provide more support for non-MPC internal users of the Bank's forecasts** who work primarily outside the monetary policy area. Some members of the FPC and PRAB, as well as the staff supporting them, have not previously worked closely with the Bank's macroeconomic forecasts, and are not necessarily familiar with the way the forecasts are generated, nor the degree of oversight from the MPC (which is greater for the headline projections for GDP growth, inflation and the unemployment rate than for the other projections generated by Bank staff). To help these internal users employ Bank forecasts more effectively, we recommend that forecast staff should consider giving key forecasts a score based on past forecast performance and the degree of oversight by the MPC.

We also recommend that comprehensive evaluations such as this one, assessing both the absolute and relative performance of Bank forecasts, be undertaken at reasonably regular intervals(1) and that future evaluations include an assessment of whether past recommendations were successfully implemented.

Although this evaluation was comprehensive and broad in scope, there are a number of additional avenues that future work could consider investigating. With regards to the tests of strong efficiency, the set of explanatory variables examined could usefully be expanded. For example, for UK GDP growth the tests could also include business survey indicators, and for inflation, they could include the oil price and sterling exchange rate.(2) The range of statistical tests could also be expanded to include a test of non-decreasing forecast variance. As Patton and Timmermann (2007) explain, optimal forecasts should be such that the variance of the forecast error is non-decreasing in the forecast horizon. Otherwise, forecasters would not be updating their forecasts optimally in the light of new information that became available as the forecast horizon shortened.

In addition, even though we do not judge correlated forecast errors across variables to be a first-order issue for the interpretation of our results, future work in this area could formally account for such interdependencies, for example by testing different variables jointly in a Seemingly Unrelated Regression. Other potential extensions to our work could include measures of forecast predictability, beyond the variable's standard deviation (for example, signal to noise ratios). Moreover, the predictability properties of forecasts for a particular variable could be contrasted across different forecast horizons, since for a (locally) stationary variable there exists a maximum horizon beyond which forecasts can provide no more information than the unconditional mean of the same variable.(3)

(1) We do not recommend undertaking major evaluation exercises too frequently as a number of years of data are required to be able to judge with any confidence whether forecast performance has changed systematically, or whether the results are driven by one or two outliers.
(2) Hackworth, Radia and Roberts (2013) conducted such tests on forecasts up to May 2013 and showed, for example, that placing greater weight on business output surveys would have improved the accuracy of one quarter ahead GDP forecasts. That work could be updated and other variables could also be considered for inclusion in the strong efficiency tests.
(3) See Galbraith and Tkacz (2007).

Annex 1 Terms of reference and project governance

The remit for this project was discussed and approved by the Bank's Court of Directors, as set out in the minutes of the Court meeting on 29 September 2014.(1) Specifically, Court approved:

> *'a work programme that would give the Oversight Committee a better basis for evaluating the Bank's forecasting performance. The assessment would provide benchmarks for both MPC and FPC relevant forecast variables, including tests for bias, and comparisons with the forecast performance of other forecasters of the UK economy and central banks. This would provide a valuable resource to the Bank in its further work'.*

The analysis set out in this paper was carried out by an Evaluation Team — Lea Paterson (IEO Director), Nyssa Roberts and Magda Rutkowska,(2) reporting directly to the Chairman of Court. James Mitchell, Professor of Economic Modelling and Forecasting at Warwick Business School, provided a peer review of the project's methodological framework, as well as technical advice and support. A statement from Professor Mitchell is provided in Annex 2.

The IEO welcomes feedback on all aspects of its work and can be contacted on independentevaluation@bankofengland.co.uk.

(1) See www.bankofengland.co.uk/publications/minutes/Documents/court/court1409.pdf.
(2) Additional analytical support was provided by Nikoleta Anesti, Lai Wah Co, Manuela Lupton-Paez, Scott Simmons, Callum Taylor, Madeleine Warwick, Benjamin White and Daniel Woodgate. This analysis also benefited from helpful comments and suggestions from: Alina Barnett, James Bell, Venetia Bell, James Benford, Ben Broadbent, Phil Evans, Nicholas Fawcett, Andrew Haldane, Charlotte Hogg, Lena Koerber, Michael Kumhof, Gareth Ramsay, Martin Taylor, Martin Weale and Garry Young. Verina Oxley and Candy Stephens provided administrative support.

Annex 2 Statement from Professor James Mitchell

At the request of the IEO, James Mitchell, Professor of Economic Modelling and Forecasting at Warwick Business School, carried out an independent peer review of the empirical framework set out in this paper.

Specifically, Professor Mitchell was asked to comment on the quality and appropriateness of the IEO's analysis, taking into account the scope of the project as set out in its Terms of Reference (Annex 1), and any data constraints. He was additionally asked to comment on the fairness of the IEO's interpretation of its results.

A statement from Professor Mitchell summarising his views is set out overleaf. We would nevertheless note that the analysis and views contained in this paper, together with any errors herein, are the responsibility of the IEO, and not of Professor Mitchell, nor of the Bank more widely.

The IEO compensated Professor Mitchell for his time, in line with the Bank's standard rates for academic consultants.

Peer Review of IEO paper: evaluating forecast performance[1]

By James Mitchell, Professor of Economic Modelling and Forecasting,
Head of Economic Modelling and Forecasting Group, Warwick Business School,
University of Warwick

james.mitchell@wbs.ac.uk

2/10/15

I welcome publication of this Paper from the Bank of England's Independent Evaluation Office (IEO) evaluating, from a statistical perspective, the performance of a wide range of Bank forecasts. In exploiting internally held forecast data, as well as published forecasts, the IEO Paper extends the coverage and depth of previous studies of Bank forecast performance, which have typically focused on GDP growth and inflation alone.

The Paper provides, in my view, a fair and balanced statistical evaluation of the Bank's forecasts using a battery of appropriate forecast evaluation tests; sensible conclusions are drawn with worthy recommendations made to improve forecast performance in some areas. Of particular value in assessing the quality of the Bank's forecasts is the comparison against rival forecasts not just from a simple benchmark model but against individual forecasters from the private sector and other central banks.

The publication of this Paper with its comprehensive assessment of univariate forecast performance will, I both hope and expect, encourage further discussion about the production and use, as well as evaluation, of forecasts both within and outside the Bank. An open exchange of ideas can only serve to strengthen further the quality of and trust in the Bank forecasts on which many (different) economic and policymaking decisions are based.

Given that Bank forecasts do serve different purposes, I encourage the Bank both to produce and then publish probability or density forecasts more widely. For eight of the eleven variables considered in this Paper, point (specifically modal) forecasts only have been produced historically. The production and publication of density forecasts ('fan charts') for these eight variables too, ideally ensuring consistency with the headline 'fan charts' for GDP growth, inflation and unemployment which are currently published as agreed by the MPC, would enable a richer understanding of the risks associated with the forecasts; and it would facilitate improved decision-making for those users whose interests may not lie with the most likely (ie modal) outcome. Forecast and forecast error distributions need not be Gaussian and may vary over time; risks may be asymmetric. The production and *ex post* evaluation of density forecasts, ideally accommodating cross-variable interactions, provides a means of testing forecast accuracy free from specific assumptions about the forecast user's decision-making objective(s) or 'loss function'. Only when a density forecast is 'well-calibrated' (see Mitchell and Wallis (2011))[2] will it be preferred by all forecast users irrespective of their loss function. This is pertinent given that Bank density forecasts for GDP growth and inflation historically appear, except at short forecast horizons, to fail standard calibration tests. It may (or may not) be 'optimal' (or loss function minimising) for Bank point forecasts to be systematically optimistic or gloomy. Absent knowledge of the Bank's (economic) loss function we cannot tell. Despite the higher ('worse') RMSE estimates that follow over the evaluation sample analysed in this Paper, is the observed

(1) The IEO and its 'Evaluation Team' at the Bank (Lea Paterson, Nyssa Roberts and Magda Rutkowska) have provided clear, timely and comprehensive guidance to me about the nature and scope of their work and the role of the independent Peer Review within this. We have discussed the work as it has progressed; and the Evaluation Team have been open and transparent in sharing and discussing with me the evaluation methods, results and the interpretation of these, during the course of their preparation of this Paper.

(2) Mitchell, J and Wallis, K F (2011), 'Evaluating density forecasts: forecast combinations, model mixtures, calibration and sharpness', *Journal of Applied Econometrics*, Vol. 26, pages 1,023–40.

clustering of longer-horizon mean (as well as modal) inflation forecasts near to the 2% inflation target really an indication of relative forecast 'failure' if the Bank does not have a quadratic loss function in inflation alone and the underlying (conditional) density forecast is not well-calibrated? Given that the ranking of competing forecasts can, and in practice is often found to, vary with the chosen measure of forecast accuracy it is recommended, especially when forecasts are likely misspecified, that the Bank should seek to specify *ex ante* the loss function that should be used for univariate/multivariate forecast evaluation (see Patton (2015))[1]. While, if indeed feasible, this may complicate or indeed compromise other reasons why the Bank publishes its own forecasts, only then can forecast performance, *per se*, be evaluated impartially, free from second-guesses of the Bank's loss function.

The global financial crisis, and ensuing 'Great Recession', have highlighted the structural instabilities and structural breaks which confront all economic forecasters. Both absolute and relative (across forecasters and/or models) forecast performance does, inevitably, change over time; and the Bank has found evidence of a declining forecast performance for some variables post crisis. The forecasting profession has observed that some models (and data or indicators) work well when forecasting in some 'states' of the world, but often do not work well in others; and research continues on how best to identify this in 'real-time'. I therefore welcome the Paper's recommendation that the Bank looks to, like all forecasters should, 'learn more from other forecasters and models'. Econometric models vary in their robustness to different types of structural instability; and so consulting forecasts from many (different) models does offer one means of insuring you do not 'put all your eggs in one basket'; although an active area of research is how best to use or combine the insights from many different models/data-series/forecasts, especially when these differ in their forecast narrative. In turn, no doubt the valued-added experts provide when subjectively applying their judgement, to varying degrees, to model-based forecasts differs. Hence I endorse the recommendation that the Bank seeks to clarify the degree of 'ownership' of its different forecasts. The ideal of separating model based ('Staff') forecasts from 'Committee' or 'Bank' based forecasts — with an increased understanding of the conditioning assumptions — would let users of Bank forecasts identify the degree of judgement applied to the model(s) based forecast. A clearer appreciation of if and when model based forecasts depart from judgement forecasts may prove helpful, especially when forecasting at times of change, which is when forecasts, of course, are most needed.

(1) Patton, A J (2015), 'Evaluating and Comparing Possibly Misspecified Forecasts', Duke University, available at http://public.econ.duke.edu/~ap172/.

Annex 3 Additional detail on data and methodology used

Table A.3.A Details of data

Variable	Definition and source[a]	Sample periods (quarterly forecasts)[b]	Sample periods (calendar-year forecasts)[c]
UK GDP growth[d]	Annual growth rate of real GDP	One quarter ahead: 1997 Q3–2014 Q2 One year ahead: 1998 Q2–2014 Q2 Two years ahead: 1999 Q2–2014 Q2	Forecasts taken from February *IRs*: One year ahead: 1998–2014 Two years ahead: 1999–2014
Inflation	Annual growth rate of retail price inflation excluding mortgage interest payments (RPIX) until February 2004 when the MPC's target measure changed to consumer price inflation (CPI)	One quarter ahead: 1997 Q3–2014 Q3 One year ahead: 1998 Q2–2014 Q3 Two years ahead: 1999 Q2–2014 Q3	Forecasts taken from February *IRs*: One year ahead: 1998–2014 Two years ahead: 1999–2014
Unemployment rate[e]	Average rate of unemployment as measured by the ONS Labour Force Survey	One quarter ahead: 1998 Q2–2014 Q2 One year ahead: 1999 Q1–2014 Q2 Two years ahead: 2000 Q1–2014 Q2	Forecasts taken from February *IRs*: One year ahead: 1999–2014 Two years ahead: 2000–14
Consumption growth[d]	Annual growth rate of real private consumption including non-profit institutions serving households	One quarter ahead: 1997 Q2–2014 Q2 One year ahead: 1998 Q1–2014 Q2 Two years ahead: 1999 Q1–2014 Q2	Forecasts taken from May *IRs*: One year ahead: 1998–2014 Two years ahead: 1999–2014
Investment growth	Annual growth rate of real total investment	One quarter ahead: 1997 Q2–2014 Q2 One year ahead: 1998 Q1–2014 Q2 Two years ahead: 1999 Q1–2014 Q2	Forecasts taken from May *IRs*: One year ahead: 1998–2014 Two years ahead: 1999–2014
House price growth	Annual growth rate of nominal house prices. Measured using the DETR/ONS house price index prior to May 2002 and subsequently measured using the average of the Halifax and Nationwide house price indices	One quarter ahead: 1997 Q3–2014 Q3 One year ahead: 1998 Q2–2014 Q3 Two years ahead: 1999 Q2–2014 Q3	Forecasts taken from February *IRs*: One year ahead: 2009–14 Two years ahead: n.a.
Wage growth	Annual growth rate of nominal wages. Measured using the whole economy Average Earnings Index until May 2010 (the series was subsequently discontinued by the ONS). From August 2010, measured using Average Weekly Earnings Forecast made in August 2003 *Inflation Report* was for a different variable (wages and salaries) and is excluded from the evaluation	One quarter ahead: 1997 Q3–2014 Q2 One year ahead: 1998 Q2–2014 Q2, except for 2010 Q3–2011 Q1 due to a break in data series. Two years ahead: 1999 Q2–2014 Q2, except for 2010 Q3–2012 Q1 due to a break in data series.	Forecasts taken from February *IRs*: One year ahead: 1998–2014. Forecast error for 2010 is missing due to a break in data series Two years ahead: 1999–2014. Forecast errors for 2010 and 2011 are missing due to a break in data series
US GDP growth	Annual growth rate of real GDP in the United States Source: US Bureau of Economic Analysis	One quarter ahead: 2001 Q1–2014 Q3 One year ahead: 2001 Q4–2014 Q3 Two years ahead: 2002 Q4–2014 Q3	Forecasts taken from February *IRs*: One year ahead: 2001–14 Two years ahead: 2002–14
Euro-area GDP growth[f]	Annual growth rate of real GDP in the euro-area Source: Eurostat and OECD (for real-time data)	One quarter ahead: 2000 Q4–2014 Q2 One year ahead: 2001 Q3–2014 Q2 Two years ahead: 2002 Q3–2014 Q2	Forecasts taken from May *IRs*: One year ahead: 2001–14 Two years ahead: 2002–14
Household lending growth	Annual growth rate of total nominal household lending.	One quarter ahead: 2006 Q4–2014 Q2 One year ahead: 2007 Q3–2014 Q2 Two years ahead: n.a.	n.a.
Corporate lending growth[g]	Annual growth rate of total nominal corporate lending, excluding the effects of securitisation	One quarter ahead: 2009 Q1–2014 Q3 One year ahead: 2009 Q4–2014 Q3 Two years ahead: n.a.	n.a.

(a) Source of all UK data is the Office for National Statistics (ONS), unless otherwise stated.
(b) Forecasts published in the quarterly *Inflation Reports*.
(c) One year ahead calendar year forecasts are defined as forecasts for the year ending in Q4, which are published in the first *Inflation Report* after the publication of data outturns for Q4 of the previous year. Similarly for two years ahead calendar year forecasts.
(d) For UK GDP and consumption, when calculating forecast errors using the latest vintage of data, we made an adjustment to forecasts made before November 2011 to account for the unusual and large upwards revision to average growth of 0.3 percentage points per year, which reflected methodological change in the 2011 ONS *Blue Book*, on the grounds that methodological revisions are likely to be unpredictable, unlike revisions based on the addition of new information. For forecasts errors computed using real-time data, we adjusted those forecasts where the outturn was affected by the 2011 *Blue Book* revisions but the forecast was made without knowledge of this. Affected forecasts were: One quarter ahead forecasts for 2011 Q2 and Q3; one year ahead forecasts for 2011 Q2 to 2012 Q2; and two year ahead forecasts for 2011 Q2–2013 Q2. The box on page 20 of the November 2011 *Inflation Report* provides further discussion of the revisions in the 2011 *Blue Book*. We did not make such adjustments when comparing Bank forecasts with those made by UK private sector forecasters (because both sets of forecasters should have been equally affected by the revision) or with other central bank forecasts (because we are not sufficiently expert in foreign data methods to know what adjustments would be appropriate for other economies).
(e) Bank forecasts made in the August 1997, November 1997 and February 1998 *Inflation Reports* were for the claimant count measure of the unemployment rate (see page 27 of the May 1998 *Inflation Report*). We excluded these forecasts from our sample.
(f) The euro area expanded through a series of enlargements during the period covered in this report. For the purposes of benchmarking central banks forecasts against each other and against private sector forecasts, forecast errors are based on the floating country composition data, available from Eurostat. We also performed robustness checks using data for a fixed country composition (EA18). For the evaluation of Bank forecasts of euro-area GDP growth using the tests of unbiasedness, efficiency, and benchmarking against random walk forecasts, we also required real-time data, which were not available from Eurostat for the floating country composition; so for these tests, forecast errors are based on data for the fixed country composition (EA18).
(g) Data are break-adjusted, source is the Bank; see www.bankofengland.co.uk/statistics/Pages/iadb/notesiadb/break_adjusted.aspx for more details.

Table A.3.B European Central Bank (ECB) forecasts and euro-area data

Variable	Definition and source[a][b]	ECB forecasts used[c]	Sample periods
GDP growth	Calendar-year growth rate of real GDP[d]	ECB Staff Macroeconomic Projections, March exercises	One year ahead: 2001–14 Two years ahead: 2002–14
Inflation	Calendar-year growth rate of HICP[d]		
Consumption growth	Calendar-year growth rate of real private consumption including non-profit institutions serving households		
Investment growth	Calendar-year growth rate of real gross fixed capital formation		

(a) Source of all euro-area data is Eurostat, unless otherwise stated.
(b) See footnote f to **Table A.3.A**.
(c) One year ahead calendar year forecasts are defined as forecasts for the year ending in Q4, which are published after the publication of data outturns for Q4 of the previous year. Similarly for two years ahead calendar year forecasts.
(d) Prior to June 2013, projections for HICP inflation and real GDP growth were published in ranges. As a simplifying assumption, we used mid-points of these ranges as point forecasts used in the evaluation.

Table A.3.C US Federal Reserve (Fed) forecasts and US data

Variable	Definition and source	Fed forecasts used[a][b]	Sample periods
GDP growth	Q4-on-Q4 growth rate of real GDP Source: US Bureau of Economic Analysis	Economic projections of Federal Reserve Board members and Federal Reserve Bank presidents. One year ahead forecasts: February Monetary Policy Reports Two year ahead forecasts: February Monetary Policy Reports from 2005 onwards. July Monetary Policy Reports prior to 2005.[d]	Forecasts taken from February *IR*s: One year ahead: 1998–2014 Two years ahead: 1999–2014
Inflation	Q4-on-Q4 growth rate of consumer prices Forecast definition was CPI in 1998 and 1999, PCE from 2000–04, and core PCE thereafter[c] Source: US Bureau of Economic Analysis		
Unemployment rate	Average civilian unemployment rate in Q4 Source: US Bureau of Labor Statistics		

(a) One year ahead forecasts are defined as Q4 (percentage change on a year ago) forecasts which are published after the publication of data outturns for Q4 of the previous year. Similarly for two years ahead forecasts.
(b) For each Monetary Policy Report, there is a collection of forecasts made by FOMC participants. As a simplifying assumption, in order to evaluate point forecasts, we used mid-points of central tendency ranges (ranges which exclude the three highest and three lowest projections for each variable in each year).
(c) See, for example, footnote 1 on page 164 of the Monetary Policy Report to the congress March 2000 and page 3 of the Monetary Policy Report to Congress July 2004. Core PCE is the price index for personal consumption expenditures excluding food and energy.
(d) Two year ahead forecasts are not available in the February Monetary Policy Reports prior to 2005, so we used July forecasts as an approximation instead. As a result, we acknowledge that these forecasts are seven quarters ahead forecasts prior to 2005 and nine quarters ahead forecasts from 2005 onwards, in contrast to the eight quarters ahead forecasts that are evaluated across all the other central banks.

Table A.3.D Reserve Bank of New Zealand (RBNZ) forecasts and New Zealand data

Variable	Definition and source[a]	RBNZ forecasts used[b]	Sample periods
GDP growth	Calendar-year growth rate of real GDP (production measure)	June Monetary Policy Statements	One year ahead: 1998–2014 Two years ahead: 1999–2014
Inflation	Calendar-year growth rate of CPI	March Monetary Policy Statements	
Unemployment rate	Calendar-year average rate of unemployment		
Consumption growth	Calendar-year growth rate of real final consumption including non-profit organisations	June Monetary Policy Statements	
Investment growth	Calendar-year growth rate of real gross fixed capital formation, all sectors		

(a) Source of all New Zealand data is Statistics New Zealand.
(b) Forecasts published in the RBNZ Monetary Policy Statements are for March year (ie the year to Q1) rather than calendar year. We are grateful to RBNZ for providing the underlying data to us so that we could use calendar year forecasts in our forecast evaluation. One year ahead calendar year forecasts are defined as forecasts for the year ending in Q4, which are produced after the publication of data outturns for Q4 of the previous year. Similarly for two years ahead calendar year forecasts.

Table A.3.E Swedish Riksbank forecasts and Sweden data

Variable	Definition and source[a]	Riksbank forecasts used[b]	Sample periods
GDP growth	Calendar-year growth rate of real GDP	February, March or April Monetary Policy Reports	One year ahead: 1998–2014 Two years ahead: 1999–2014
Inflation	Calendar-year growth rate of CPI	February or March Monetary Policy Reports	
Consumption growth	Annual growth rate of real private consumption including non-profit institutions serving households	February, March or April Monetary Policy Reports	One year ahead: 2000–14 Two years ahead: 2001–14
Investment growth	Calendar-year growth rate of real gross fixed capital formation		

(a) Source of all Sweden data is Statistics Sweden.
(b) Dates of Riksbank MPRs have moved over time. One year ahead calendar year forecasts are defined as forecasts for the year ending in Q4, which are published after the publication of data outturns for Q4 of the previous year. Similarly for two years ahead calendar year forecasts.

Methodology

Estimating confidence intervals using the delta method

We estimate confidence intervals for the scaled RMSEs in **Charts 3.1–3.5** on pages 27–30 (and in the associated charts in Annex 4) using the delta method:(1)

$$\sqrt{\hat{\alpha}} \pm z \frac{\hat{\sigma}_{\alpha}}{2\sqrt{\hat{\alpha}}}$$

where $\hat{\alpha} = \frac{1}{T\hat{\sigma}_{y}^{2}} \sum_{t=1}^{T} (y_t - f_t)^2$

is the estimate of the Mean Squared Error scaled by the estimate of the variance of data outturns, $\hat{\sigma}_{y}^{2}$. Forecast errors are defined as the differences between the data outturns in period *t*, y_t, and the forecasts for period *t*, f_t; and we assume the variance of data outturns is constant across all *t*s.

$\hat{\sigma}_{\alpha}$ is the estimate of the standard error of $\hat{\alpha}$ and z is the appropriate critical value.

To derive the above confidence interval, we define the following transformation: $f(\alpha) = \sqrt{\alpha}$, apply the Central Limit Theorem, stating that:

$$\sqrt{T}(\hat{\alpha} - \alpha) \xrightarrow{d} N(0, \sigma_{\alpha}^{2})$$

and the delta method, stating that:

$$\sqrt{T}(f(\hat{\alpha}) - f(\alpha)) \xrightarrow{d} N(0, \sigma_{\alpha}^{2} f'(\alpha)^2)$$

(1) More details can be found in Greene (2011).

Annex 4 Additional empirical results

Table A.4.A Correlation of modal one quarter ahead forecast errors across variables, 1997–2014 excluding 2008 and 2009[a][b]

	Inflation	Unemployment rate	Consumption growth	Investment growth	House prices growth	Wages growth	Household lending growth	Corporate lending growth	US GDP growth	Euro-area GDP growth
UK GDP growth	-0.25	-0.17	0.50	0.26	0.30	-0.17	0.24	0.21	0.25	0.15
Inflation		0.18	-0.07	-0.12	-0.19	-0.08	-0.39	0.05	-0.06	0.04
Unemployment rate			-0.07	0.00	-0.18	0.05	-0.33	-0.33	-0.39	0.42
Consumption growth				-0.05	0.32	-0.11	0.02	0.20	0.15	-0.04
Investment growth					0.08	-0.14	-0.11	0.11	-0.03	-0.03
House prices growth						0.00	0.12	0.03	-0.01	-0.02
Wages growth							-0.06	-0.25	-0.13	0.29
Household lending growth								0.25	0.35	-0.10
Corporate lending growth									0.53	-0.13
US GDP growth										-0.28

(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A**.
(b) Forecast errors are computed using the latest vintage of data. Red shading indicates positive correlation and blue shading indicates negative correlation. Darker shading indicates stronger (positive or negative) correlation than lighter shading.

Table A.4.B Correlation of modal one year ahead forecast errors across variables, 1998–2014 excluding 2008 and 2009[a][b]

	Inflation	Unemployment rate	Consumption growth	Investment growth	House prices growth	Wages growth	Household lending growth	Corporate lending growth	US GDP growth	Euro-area GDP growth
UK GDP growth	-0.26	-0.41	0.58	0.09	0.48	0.06	-0.22	0.22	0.50	0.13
Inflation		-0.11	-0.38	0.35	-0.01	0.27	-0.30	0.05	0.19	0.34
Unemployment rate			-0.34	-0.33	-0.23	-0.20	0.17	-0.36	-0.69	-0.08
Consumption growth				-0.14	0.17	0.01	0.09	0.01	0.25	-0.31
Investment growth					0.40	-0.04	-0.18	0.04	0.32	0.42
House prices growth						0.21	-0.02	-0.25	0.29	0.12
Wages growth							-0.65	0.15	0.17	0.47
Household lending growth								-0.35	-0.17	-0.58
Corporate lending growth									0.22	0.12
US GDP growth										0.30

(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A**.
(b) Forecast errors are computed using the latest vintage of data. Red shading indicates positive correlation and blue shading indicates negative correlation. Darker shading indicates stronger (positive or negative) correlation than lighter shading.

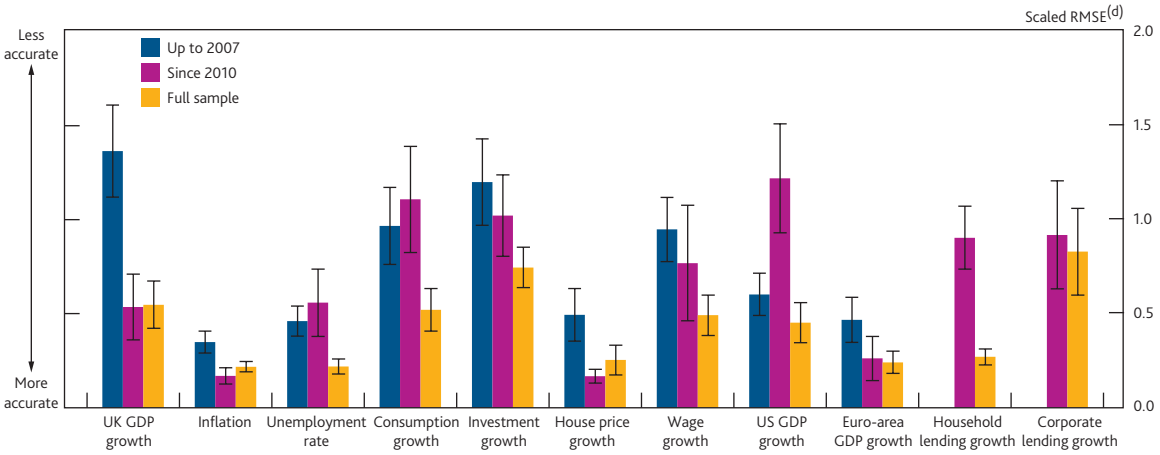
Table A.4.C Correlation of modal two year ahead forecast errors across variables, 1999–2014 excluding 2008 and 2009[a][b]

	Inflation	Unemployment rate	Consumption growth	Investment growth	House prices growth	Wages growth	Household lending growth	Corporate lending growth	US GDP growth	Euro-area GDP growth
UK GDP growth	-0.52	-0.27	0.65	0.24	0.38	0.15	0.10	0.62	0.23	0.09
Inflation		0.08	-0.70	0.24	-0.32	-0.44	-0.71	-0.20	-0.09	0.18
Unemployment rate			-0.29	0.01	0.22	0.04	0.26	-0.46	-0.26	0.16
Consumption growth				-0.04	0.32	0.24	0.56	0.06	0.27	-0.10
Investment growth					0.34	-0.11	-0.07	0.36	0.09	0.62
House prices growth						0.25	0.47	0.58	-0.12	0.14
Wages growth							-0.28	-0.22	-0.01	0.25
Household lending growth								-0.20	-0.01	-0.38
Corporate lending growth									0.43	0.46
US GDP growth										0.06

(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A**.
(b) Forecast errors are computed using the latest vintage of data. Red shading indicates positive correlation and blue shading indicates negative correlation. Darker shading indicates stronger (positive or negative) correlation than lighter shading.

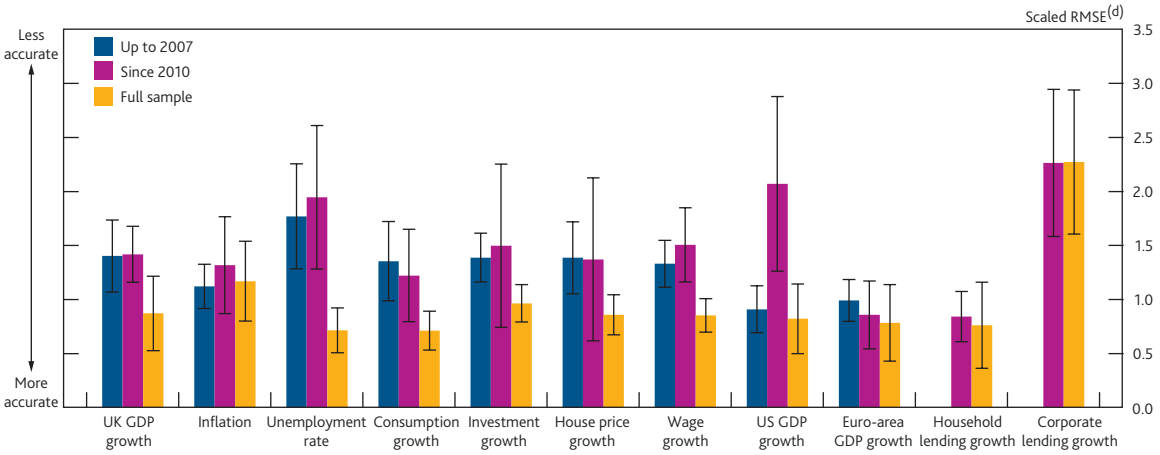
Accuracy of Bank forecasts

Chart A.4.1 Accuracy of Bank modal one quarter ahead forecasts, comparison across time(a)(b)(c)



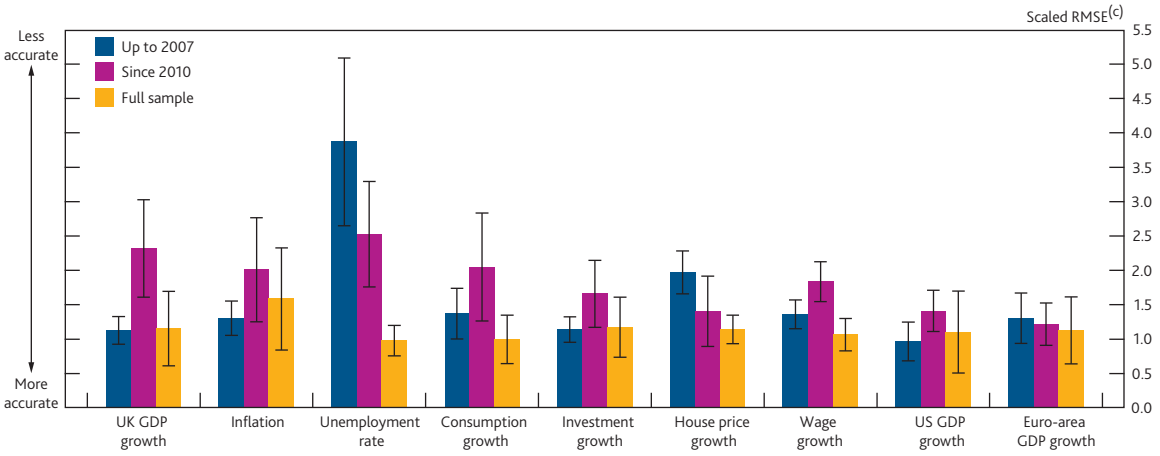
(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3.
(b) Forecasts for household and corporate lending up to 2007 are not assessed because data are not available.
(c) Forecast errors are computed using the latest vintage of data.
(d) Root mean squared errors scaled by the standard deviation of data outturns. Error bars indicate 90% confidence intervals for each variable in each period of time, conducted using the delta method (see Annex 3 for more details).

Chart A.4.2 Accuracy of Bank modal one year ahead forecasts, comparison across time(a)(b)(c)



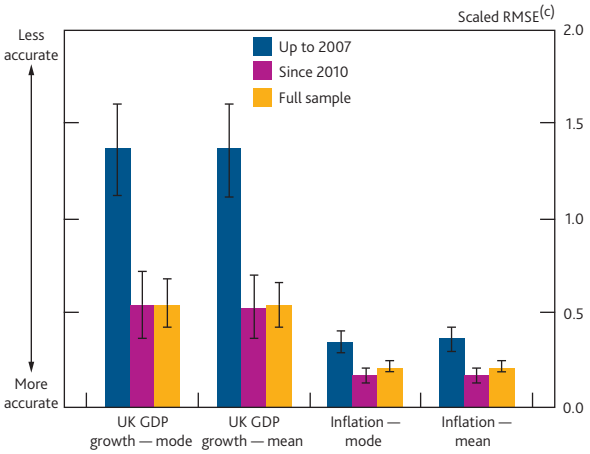
(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3.
(b) Forecasts for household and corporate lending up to 2007 are not assessed because data are not available.
(c) Forecast errors are computed using the latest vintage of data.
(d) Root mean squared errors scaled by the standard deviation of data outturns. Error bars indicate 90% confidence intervals for each variable in each period of time, conducted using the delta method (see Annex 3 for more details).

Chart A.4.3 Accuracy of Bank modal two year ahead forecasts, comparison across time(a)(b)



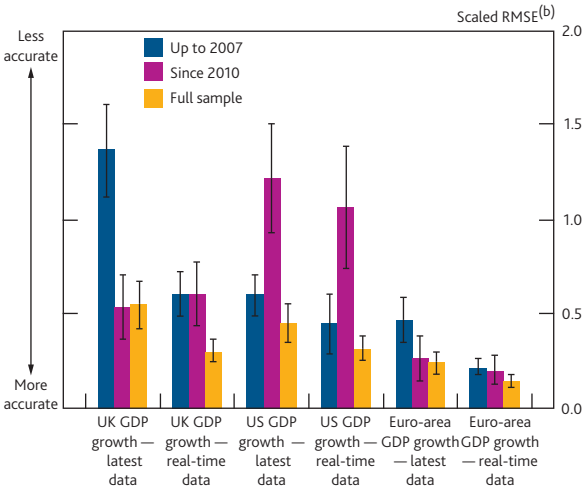
(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3.
(b) Forecast errors are computed using the latest vintage of data.
(c) Root mean squared errors scaled by the standard deviation of data outturns. Error bars indicate 90% confidence intervals for each variable in each period of time, conducted using the delta method (see Annex 3 for more details).

Chart A.4.4 Accuracy of Bank one quarter ahead forecasts, comparison of modal and mean forecasts[a][b]



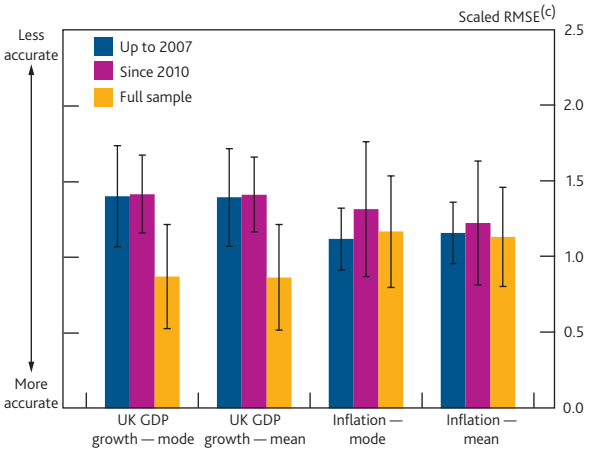
(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3.
(b) Forecast errors are computed using the latest vintage of data.
(c) Root mean squared errors scaled by the standard deviation of data outturns. Error bars indicate 90% confidence intervals for each variable in each period of time, conducted using the delta method (see Annex 3 for more details).

Chart A.4.5 Accuracy of Bank one quarter ahead forecasts, comparison of latest and real-time data[a]



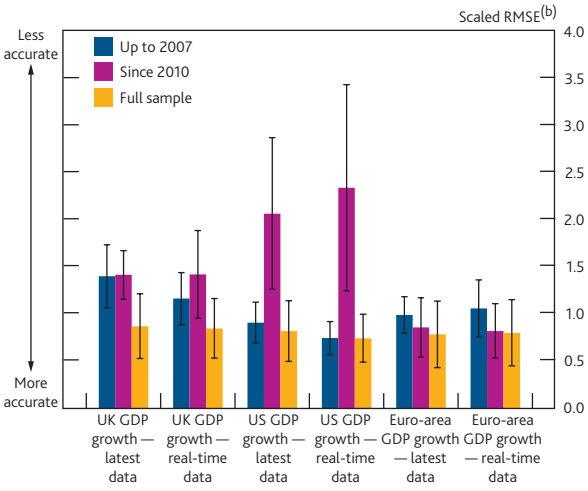
(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3.
(b) Root mean squared errors scaled by the standard deviation of data outturns. Error bars indicate 90% confidence intervals for each variable in each period of time, conducted using the delta method (see Annex 3 for more details).

Chart A.4.6 Accuracy of Bank one year ahead forecasts, comparison of modal and mean forecasts[a][b]



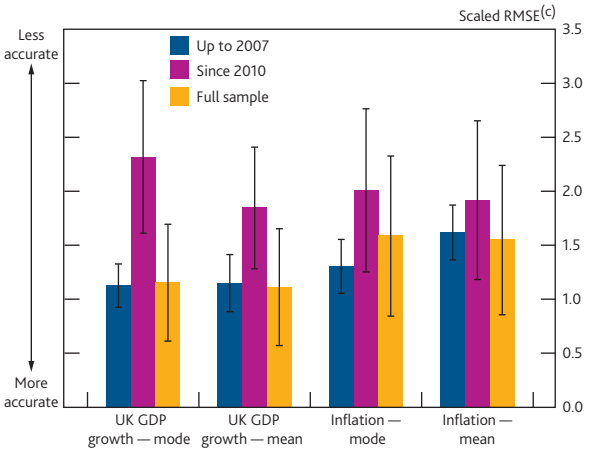
(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3.
(b) Forecast errors are computed using the latest vintage of data.
(c) Root mean squared errors scaled by the standard deviation of data outturns. Error bars indicate 90% confidence intervals for each variable in each period of time, conducted using the delta method (see Annex 3 for more details).

Chart A.4.7 Accuracy of Bank one year ahead forecasts, comparison of latest and real-time data[a]



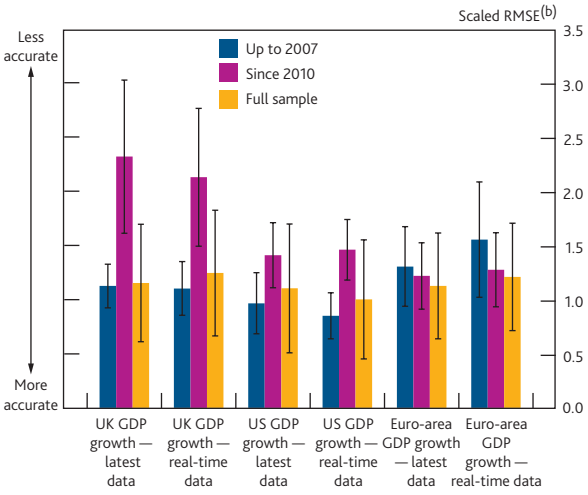
(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3.
(b) Root mean squared errors scaled by the standard deviation of data outturns. Error bars indicate 90% confidence intervals for each variable in each period of time, conducted using the delta method (see Annex 3 for more details).

Chart A.4.8 Accuracy of Bank two year ahead forecasts, comparison of modal and mean forecasts[a][b]



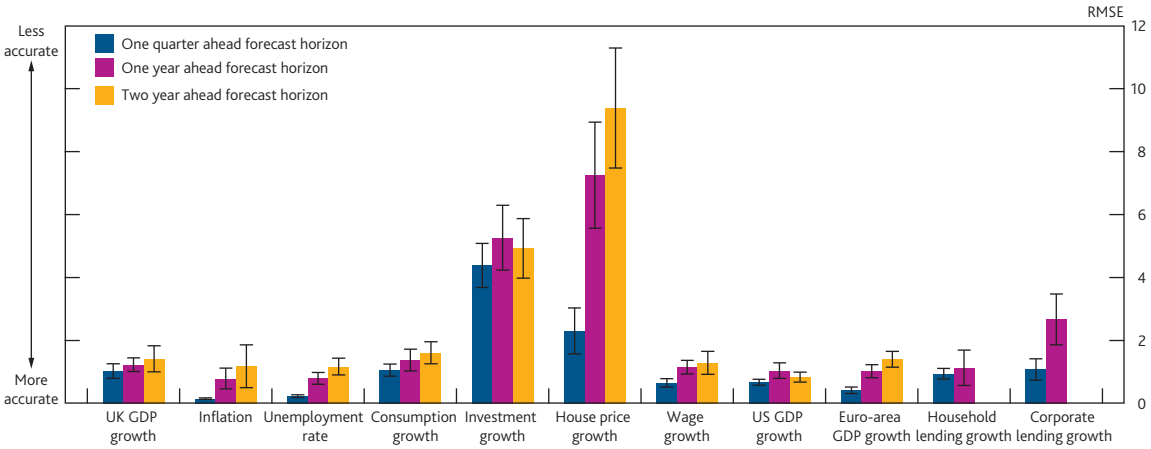
(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3.
(b) Forecast errors are computed using the latest vintage of data.
(c) Root mean squared errors scaled by the standard deviation of data outturns. Error bars indicate 90% confidence intervals for each variable in each period of time, conducted using the delta method (see Annex 3 for more details).

Chart A.4.9 Accuracy of Bank two year ahead forecasts, comparison of latest and real-time data[a]



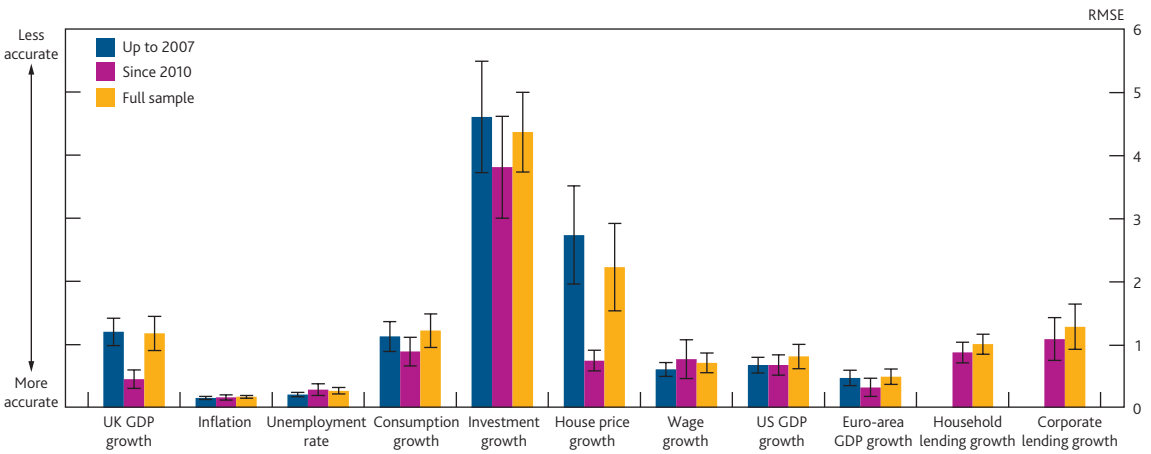
(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3.
(b) Root mean squared errors scaled by the standard deviation of data outturns. Error bars indicate 90% confidence intervals for each variable in each period of time, conducted using the delta method (see Annex 3 for more details).

Chart A.4.10 Accuracy of Bank forecasts using unscaled RMSEs (full sample excluding 2008 and 2009)[a][b][c]



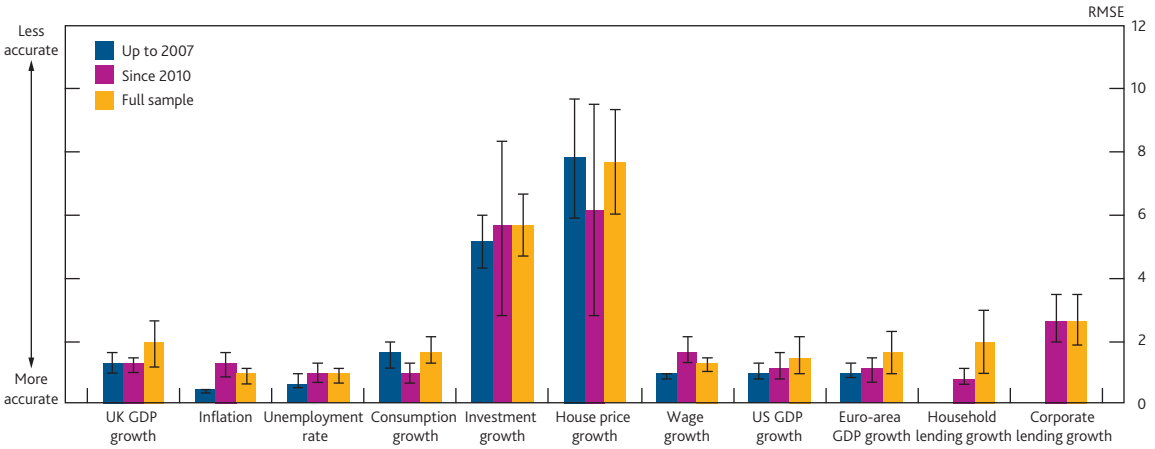
(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3.
(b) Two year ahead forecasts for households and corporate lending are not assessed because the sample sizes are very short.
(c) Forecast errors are computed using the latest vintage of data.

Chart A.4.11 Accuracy of Bank modal one quarter ahead forecasts using unscaled RMSEs, comparison across time[a][b][c]



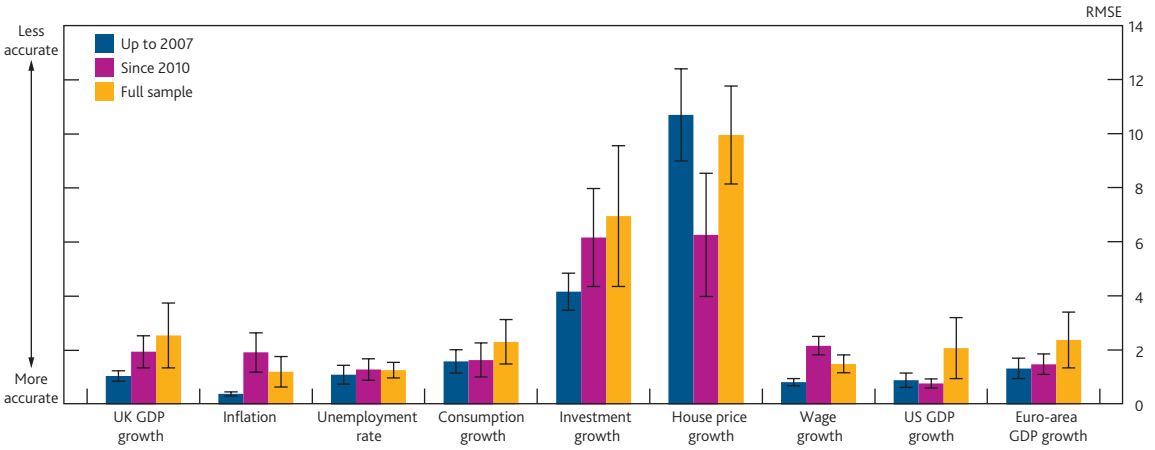
(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3.
(b) Forecasts for household and corporate lending up to 2007 are not available.
(c) Forecast errors are computed using the latest vintage of data.

Chart A.4.12 Accuracy of Bank modal one year ahead forecasts using unscaled RMSEs, comparison across time[a][b][c]



(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3.
(b) Forecasts for household and corporate lending up to 2007 are not available.
(c) Forecast errors are computed using the latest vintage of data.

Chart A.4.13 Accuracy of Bank modal two year ahead forecasts using unscaled RMSEs, comparison across time[a][b][c]



(a) Sample period varies by variable and forecast horizon, due to data availability. For details see **Table A.3.A** in Annex 3.
(b) Two year ahead forecasts for households and corporate lending are not assessed because the sample sizes are very short.
(c) Forecast errors are computed using the latest vintage of data.

Unbiasedness of Bank forecasts

Table A.4.D Unbiasedness of Bank modal forecasts (full sample including 2008 and 2009)[a]

	Bias — positive (negative) coefficient means that forecasts have tended to be too low (high)			Bias (scaled) — positive (negative) coefficient means that forecasts have tended to be too low (high)			Statistical significance of unscaled bias coefficient, p-value		
	One quarter ahead	One year ahead	Two year ahead	One quarter ahead	One year ahead	Two year ahead	One quarter ahead	One year ahead	Two year ahead
UK GDP growth	0.01	-0.42	-1.07	0.00	-0.19	-0.48	0.98	0.42	0.15
Inflation	0.02	0.29	0.57	0.03	0.38	0.75	0.34	0.09*	0.12
Unemployment rate	-0.09	-0.34	-0.26	-0.08	-0.27	-0.20	0.03**	0.07*	0.53
Consumption growth	0.06	0.11	-0.30	0.03	0.05	-0.13	0.82	0.80	0.72
Investment growth	-0.50	-1.23	-2.25	-0.09	-0.21	-0.38	0.50	0.25	0.06*
House price growth	0.78	3.83	5.26	0.09	0.43	0.60	0.06*	0.00**	0.02**
Wage growth	-0.05	-0.51	-1.04	-0.03	-0.35	-0.74	0.60	0.02**	0.00**
US GDP growth	-0.33	-0.56	-1.03	-0.18	-0.31	-0.55	0.02**	0.06*	0.06*
Euro-area GDP growth	0.15	-0.24	-1.01	0.07	-0.12	-0.48	0.20	0.41	0.01**
Household lending growth	-0.15	-1.03	n.a.	-0.04	-0.40	n.a.	0.63	0.12	n.a.
Corporate lending growth	0.04	-1.19	n.a.	0.02	-1.04	n.a.	0.88	0.07*	n.a.

(a) See footnotes to **Table 3.A** on page 30.

Table A.4.E Unbiasedness of Bank modal forecasts (full sample excluding 2008 and 2009)[a]

	Bias — positive (negative) coefficient means that forecasts have tended to be too low (high)			Bias (scaled) — positive (negative) coefficient means that forecasts have tended to be too low (high)			Statistical significance of unscaled bias coefficient, p-value		
	One quarter ahead	One year ahead	Two years ahead	One quarter ahead	One year ahead	Two years ahead	One quarter ahead	One year ahead	Two years ahead
UK GDP growth	0.19	0.02	-0.43	0.17	0.02	-0.39	0.31	0.95	0.29
Inflation	0.01	0.26	0.53	0.01	0.37	0.75	0.67	0.18	0.22
Unemployment rate	-0.09	-0.41	-0.54	-0.07	-0.33	-0.41	0.04**	0.04**	0.05**
Consumption growth	0.24	0.41	0.30	0.14	0.24	0.18	0.20	0.20	0.56
Investment growth	-0.31	-0.46	-0.66	-0.08	-0.12	-0.18	0.69	0.65	0.49
House price growth	0.90	4.73	7.61	0.13	0.69	1.14	0.04**	0.00**	0.00**
Wage growth	-0.03	-0.40	-0.87	-0.02	-0.29	-0.65	0.74	0.07*	0.00**
US GDP growth	-0.21	-0.21	-0.37	-0.22	-0.24	-0.45	0.07*	0.33	0.02**
Euro-area GDP growth	0.27	0.12	-0.39	0.21	0.10	-0.30	0.00**	0.69	0.33
Household lending growth	0.12	-0.44	n.a.	0.03	-0.17	n.a.	0.66	0.25	n.a.
Corporate lending growth	0.05	-1.30	n.a.	0.04	-1.10	n.a.	0.85	0.07*	n.a.

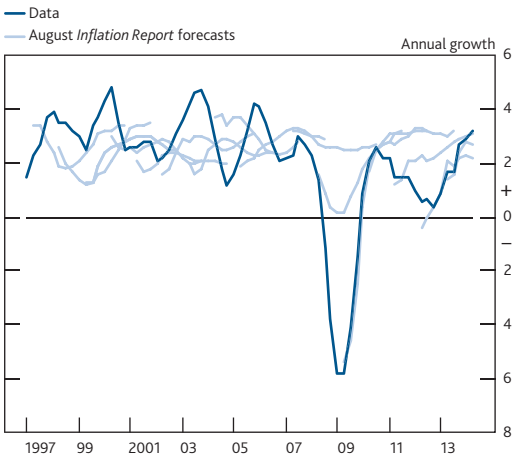
(a) See footnotes to **Table 3.A** on page 30.

Table A.4.F Unbiasedness of Bank forecasts, comparing forecast errors calculated using modal and mean forecasts, and alternative data vintages (full sample excluding 2008 and 2009)[a]

	Bias — positive (negative) coefficient means that forecasts have tended to be too low (high)			Bias (scaled) — positive (negative) coefficient means that forecasts have tended to be too low (high)			Statistical significance of unscaled bias coefficient, p-value		
	One quarter ahead	One year ahead	Two years ahead	One quarter ahead	One year ahead	Two years ahead	One quarter ahead	One year ahead	Two years ahead
UK GDP growth — mode	0.19	0.02	-0.43	0.17	0.02	-0.39	0.31	0.95	0.29
UK GDP growth — mean	0.23	0.16	-0.19	0.22	0.15	-0.17	0.21	0.57	0.56
Inflation — mode	0.01	0.26	0.53	0.01	0.37	0.75	0.67	0.18	0.22
Inflation — mean	0.00	0.22	0.46	-0.01	0.31	0.65	0.84	0.24	0.26
Modal projections:									
UK GDP growth — latest data	0.19	0.02	-0.43	0.17	0.02	-0.39	0.31	0.95	0.29
UK GDP growth — real-time data	-0.18	-0.37	-0.76	-0.17	-0.36	-0.72	0.04**	0.19	0.09*
US GDP growth — latest data	-0.21	-0.21	-0.37	-0.22	-0.24	-0.45	0.07*	0.33	0.02**
US GDP growth — real-time data	0.07	0.07	-0.12	0.07	0.07	-0.13	0.42	0.75	0.48
Euro-area GDP growth — latest data	0.27	0.12	-0.39	0.21	0.01	-0.03	0.00**	0.69	0.33
Euro-area GDP growth — real-time data	0.01	-0.12	-0.62	0.01	-0.11	-0.55	0.66	0.67	0.09*

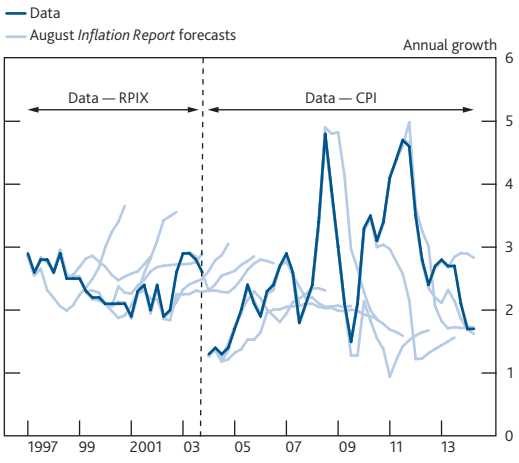
(a) See footnotes to **Table 3.A** on page 30.

Chart A.4.14 Bank modal UK GDP growth forecasts and data outturns(a)



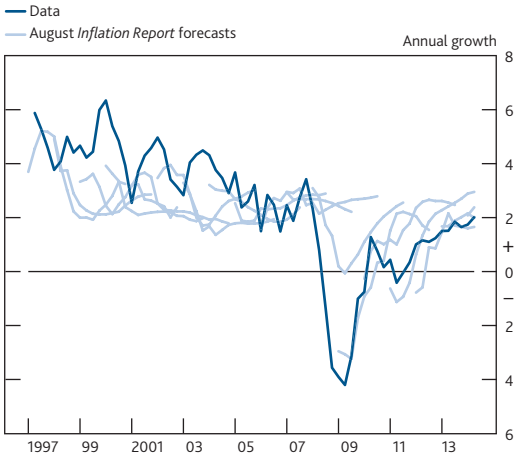
(a) For clarity, we only show August *Inflation Report* forecasts in this chart, but we do include GDP growth forecasts from all *Inflation Reports* between August 1997 and August 2014 in our analysis. Past data revisions mean that the starting points of forecasts may differ to the latest vintage shown in the dark blue line.

Chart A.4.15 Bank modal UK inflation forecasts and data outturns(a)



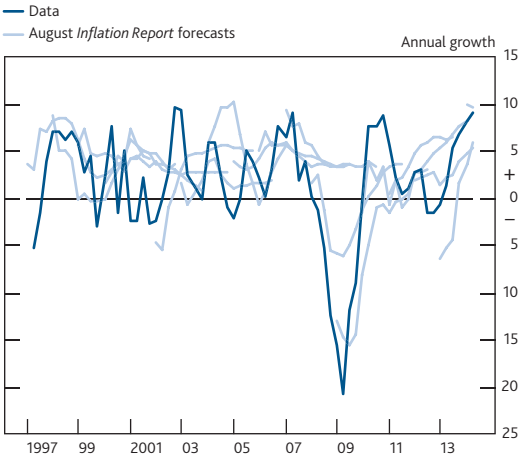
(a) For clarity, we only show August *Inflation Report* forecasts in this chart, but we do include inflation forecasts from all *Inflation Reports* between August 1997 and August 2014 in our analysis.

Chart A.4.16 Bank UK consumption growth forecasts and data outturns(a)



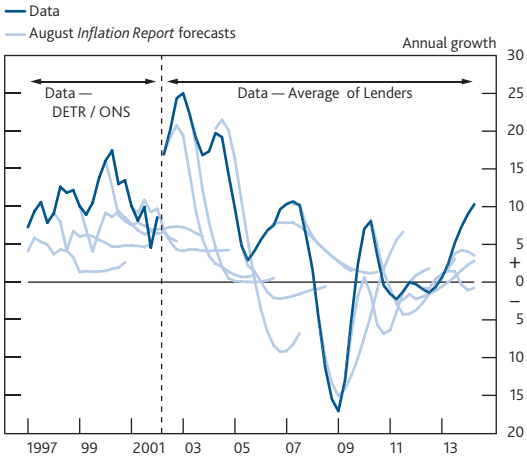
(a) For clarity, we only show August *Inflation Report* forecasts in this chart, but we do include GDP growth forecasts from all *Inflation Reports* between August 1997 and August 2014 in our analysis. Chart shows modal forecasts. Past data revisions mean that the starting points of forecasts may differ to the latest vintage shown in the dark blue line.

Chart A.4.17 Bank UK investment growth forecasts and data outturns(a)



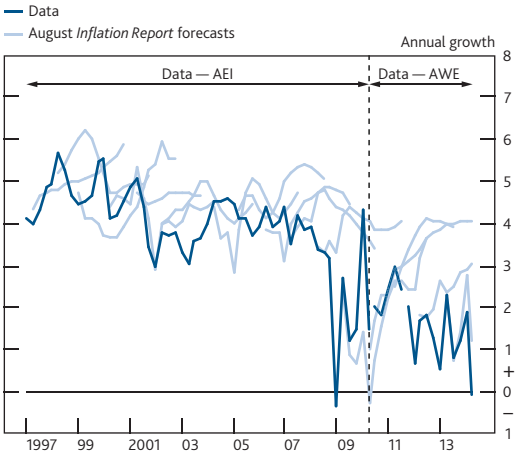
(a) For clarity, we only show August *Inflation Report* forecasts in this chart, but we do include inflation forecasts from all *Inflation Reports* between August 1997 and August 2014 in our analysis. Chart shows modal forecasts. Past data revisions mean that the starting points of forecasts may differ to the latest vintage shown in the dark blue line.

Chart A.4.18 Bank UK house price inflation forecasts and data outturns(a)



(a) For clarity, we only show August *Inflation Report* forecasts in this chart, but we do include house price growth forecasts from all *Inflation Reports* between August 1997 and August 2014 in our analysis. Chart shows modal forecasts. House price growth is measured using the DETR/ONS house price index prior to May 2002 and subsequently measured using the average of the Halifax and Nationwide house price indices.

Chart A.4.19 Bank UK wage growth forecasts and data outturns(a)

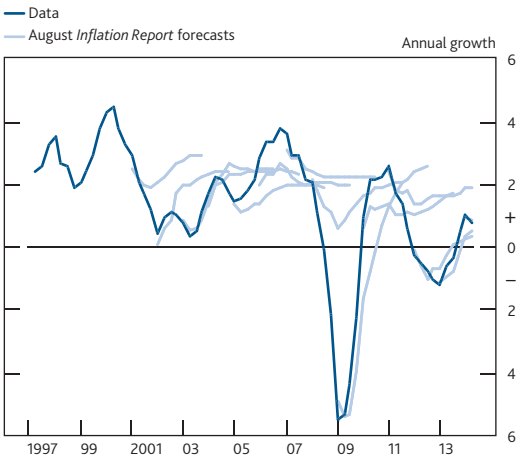


(a) For clarity, we only show August *Inflation Report* forecasts in this chart, but we do include wage growth forecasts from all *Inflation Reports* between August 1997 and August 2014 in our analysis. Chart shows modal forecasts. The Office for National Statistics discontinued the Average Earnings Index in September 2010, after replacing it with the alternative Average Weekly Earnings series in January 2010.

Chart A.4.20 Bank US GDP growth forecasts and data outturns(a)

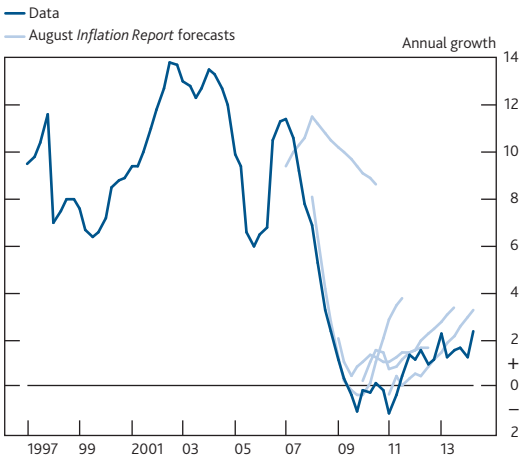
(a) For clarity, we only show August *Inflation Report* forecasts in this chart, but we do include US GDP growth forecasts from all *Inflation Reports* between February 2001 and August 2014 in our analysis. Chart shows modal forecasts. Past data revisions mean that the starting points of forecasts may differ to the latest vintage shown in the dark blue line.

Chart A.4.21 Bank euro-area GDP growth forecasts and data outturns(a)



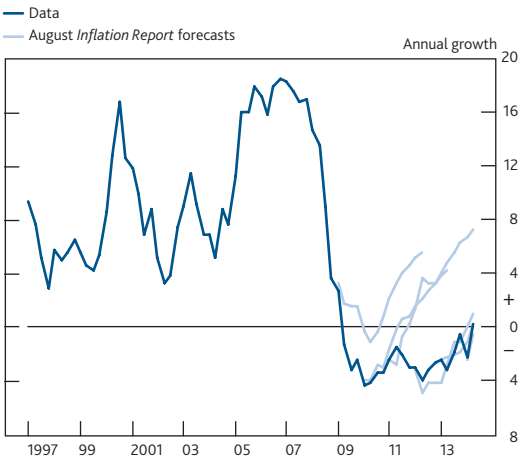
(a) For clarity, we only show August *Inflation Report* forecasts in this chart, but we do include euro-area GDP growth forecasts from all *Inflation Reports* between February 2001 and August 2014 in our analysis. Chart shows modal forecasts. Past data revisions mean that the starting points of forecasts may differ to the latest vintage shown in the dark blue line.

Chart A.4.22 Bank UK household lending growth forecasts and data outturns(a)



(a) For clarity, we only show August *Inflation Report* forecasts in this chart, but we do include household lending growth forecasts from all *Inflation Reports* between February 2007 and August 2014 in our analysis. Chart shows modal forecasts. Past data revisions mean that the starting points of forecasts may differ to the latest vintage shown in the dark blue line.

Chart A.4.23 Bank UK corporate lending growth forecasts and data outturns(a)



(a) For clarity, we only show August *Inflation Report* forecasts in this chart, but we do include corporate lending growth forecasts from all *Inflation Reports* between February 2009 and August 2014 in our analysis. Chart shows modal forecasts. Past data revisions mean that the starting points of forecasts may differ to the latest vintage shown in the dark blue line.

Efficiency of Bank of England forecasts

Table A.4.G Strong efficiency of Bank forecasts (full sample including 2008 and 2009)[a]

	β_1 (p-value) from the regression of forecast errors on **past forecast errors**			β_1 (p-value) from the regression of forecast errors on **past variable outturns**		
	One quarter ahead	One year ahead	Two years ahead	One quarter ahead	One year ahead	Two years ahead
UK GDP growth	0.49	0.00	-0.04	-0.04	-0.07	0.03
	(0.26)	(0.98)	(0.78)	(0.63)	(0.49)	(0.89)
Inflation	-0.01	-0.08	0.26	-0.03	-0.11	0.32
	(0.96)	(0.72)	(0.15)	(0.19)	(0.28)	(0.02**)
Unemployment rate	0.49	-0.02	0.10	-0.08	-0.29	-0.43
	(0.00**)	(0.91)	(0.53)	(0.02**)	(0.03**)	(0.02**)
Consumption growth	-0.13	-0.12	0.31	-0.02	0.08	0.45
	(0.55)	(0.48)	(0.12)	(0.81)	(0.45)	(0.01**)
Investment growth	-0.13	-0.21	-0.46	-0.23	-0.42	-0.41
	(0.52)	(0.25)	(0.03**)	(0.06*)	(0.01**)	(0.08*)
House price growth	-0.02	-0.15	-0.08	0.00	-0.05	0.19
	(0.90)	(0.42)	(0.70)	(0.89)	(0.77)	(0.15)
Wage growth	0.02	-0.09	0.19	-0.06	0.19	0.36
	(0.89)	(0.65)	(0.26)	(0.38)	(0.30)	(0.00**)
US GDP growth	0.00	-0.27	-0.04	-0.06	-0.36	-0.11
	(0.99)	(0.05*)	(0.72)	(0.46)	(0.00**)	(0.28)
Euro-area GDP growth	0.25	-0.18	-0.57	0.09	-0.36	-0.53
	(0.20)	(0.25)	(0.04**)	(0.00**)	(0.00**)	(0.02**)
Household lending growth	-0.80	0.05	n.a.	-0.06	-0.29	n.a.
	(0.00**)	(0.73)		(0.37)	(0.00**)	
Corporate lending growth	-0.27	-0.40	n.a.	-0.22	-0.16	n.a.
	(0.14)	(0.11)		(0.26)	(0.46)	

(a) See footnotes to **Table 3.C** on page 34.

Table A.4.H Strong efficiency of Bank forecasts, comparing forecast errors calculated using modal and mean forecasts, and alternative data vintages (full sample excluding 2008 and 2009)[a]

	β_1 (p-value) from the regression of forecast errors on **past forecast errors**			β_1 (p-value) from the regression of forecast errors on **past variable outturns**		
	One quarter ahead	One year ahead	Two years ahead	One quarter ahead	One year ahead	Two years ahead
UK GDP growth — mode	-0.05	0.04	0.15	-0.27	0.02	0.26
	(0.81)	(0.72)	(0.04**)	(0.06*)	(0.79)	(0.00**)
UK GDP growth — mean	-0.07	-0.12	0.06	-0.25	-0.07	0.14
	(0.74)	(0.32)	(0.40)	(0.07*)	(0.31)	(0.03**)
Inflation — mode	0.03	0.04	0.25	-0.02	-0.03	0.30
	(0.80)	(0.82)	(0.15)	(0.37)	(0.70)	(0.01**)
Inflation — mean	-0.01	0.01	0.21	-0.04	-0.07	0.24
	(0.97)	(0.95)	(0.20)	(0.17)	(0.40)	(0.06*)
Modal projections:						
UK GDP growth — latest data	-0.05	0.04	0.15	-0.27	0.02	0.26
	(0.81)	(0.72)	(0.04**)	(0.06*)	(0.79)	(0.00**)
UK GDP growth — real-time data	0.18	0.07	0.18	0.00	-0.01	0.27
	(0.37)	(0.55)	(0.11)	(0.98)	(0.78)	(0.00**)
US GDP growth — latest data	-0.06	-0.34	0.02	-0.31	-0.38	-0.03
	(0.74)	(0.03**)	(0.68)	(0.00**)	(0.00**)	(0.72)
US GDP growth — real-time data	-0.03	-0.37	0.04	-0.21	-0.38	0.03
	(0.88)	(0.03**)	(0.53)	(0.01**)	(0.00**)	(0.56)
Euro-area GDP growth — latest data	0.06	-0.13	-0.25	0.06	-0.25	-0.24
	(0.59)	(0.37)	(0.08*)	(0.27)	(0.01**)	(0.04**)
Euro-area GDP growth — real-time data	0.06	-0.12	-0.24	-0.04	-0.28	-0.25
	(0.67)	(0.38)	(0.05**)	(0.06*)	(0.00**)	(0.01**)

(a) See footnotes to **Table 3.C** on page 34.

Table A.4.I Strong efficiency of Bank forecasts, comparing forecast errors calculated using modal and mean forecasts, and alternative data vintages (full sample including 2008 and 2009)[a]

	β_1 (p-value) from the regression of forecast errors on **past forecast errors**			β_1 (p-value) from the regression of forecast errors on **past variable outturns**		
	One quarter ahead	One year ahead	Two years ahead	One quarter ahead	One year ahead	Two years ahead
UK GDP growth — mode	0.49	0.00	-0.04	-0.04	-0.07	0.03
	(0.26)	(0.98)	(0.78)	(0.63)	(0.49)	(0.89)
UK GDP growth — mean	0.49	-0.13	-0.12	-0.04	-0.16	-0.09
	(0.26)	(0.40)	(0.41)	(0.61)	(0.15)	(0.63)
Inflation — mode	-0.01	-0.08	0.26	-0.03	-0.11	0.32
	(0.96)	(0.72)	(0.15)	(0.19)	(0.28)	(0.02**)
Inflation — mean	-0.08	-0.10	0.23	-0.05	-0.13	0.25
	(0.50)	(0.61)	(0.20)	(0.06*)	(0.17)	(0.08*)
Modal projections:						
UK GDP growth — latest data	0.49	0.00	-0.04	-0.04	-0.07	0.03
	(0.26)	(0.98)	(0.78)	(0.63)	(0.49)	(0.89)
UK GDP growth — real-time data	0.34	0.08	0.02	0.05	-0.03	0.07
	(0.07*)	(0.51)	(0.89)	(0.12)	(0.51)	(0.68)
US GDP growth — latest data	0.00	-0.27	-0.04	-0.06	-0.36	-0.11
	(0.99)	(0.05*)	(0.72)	(0.46)	(0.00**)	(0.28)
US GDP growth — real-time data	0.02	-0.32	0.02	-0.06	-0.36	-0.02
	(0.93)	(0.05**)	(0.87)	(0.15)	(0.00**)	(0.87)
Euro-area GDP growth — latest data	0.25	-0.18	-0.57	0.09	-0.36	-0.53
	(0.20)	(0.25)	(0.04**)	(0.00**)	(0.00**)	(0.02**)
Euro-area GDP growth — real-time data	-0.02	-0.16	-0.50	-0.02	-0.36	-0.48
	(0.92)	(0.28)	(0.03**)	(0.20)	(0.00**)	(0.01**)

(a) See footnotes to **Table 3.C** on page 34.

Table A.4.J Strong efficiency of Bank forecasts using world GDP growth data and forecasts, comparing forecast errors calculated using modal and mean forecasts, and alternative data vintages (full sample excluding 2008 and 2009)[a]

	β_1 (p-value) from the regression of forecast errors on **past world GDP growth outturns**			β_1 (p-value) from the regression of forecast errors on **contemporaneous world GDP growth forecasts**		
	One quarter ahead	One year ahead	Two years ahead	One quarter ahead	One year ahead	Two years ahead
UK GDP growth — mode	-0.11	-0.03	0.00	-0.08	0.49	0.95
	(0.22)	(0.72)	(0.99)	(0.74)	(0.10*)	(0.07*)
UK GDP growth — mean	-0.11	-0.08	0.08	-0.08	0.51	1.11
	(0.21)	(0.39)	(0.24)	(0.71)	(0.17)	(0.07*)
Modal projections:						
UK GDP growth — latest data	-0.11	-0.03	0.00	-0.08	0.49	0.95
	(0.22)	(0.72)	(0.99)	(0.74)	(0.10*)	(0.07*)
UK GDP growth — real-time data	-0.12	-0.06	0.03	-0.01	0.46	0.46
	(0.12)	(0.46)	(0.83)	(0.93)	(0.06*)	(0.20)

(a) See footnotes to **Table 3.C** on page 34.

Table A.4.K Strong efficiency of Bank forecasts using world GDP growth data and forecasts, comparing forecast errors calculated using modal and mean forecasts, and alternative data vintages (full sample including 2008 and 2009)[a]

	β_1 (p-value) from the regression of forecast errors on **past world GDP growth outturns**			β_1 (p-value) from the regression of forecast errors on **contemporaneous world GDP growth forecasts**		
	One quarter ahead	One year ahead	Two years ahead	One quarter ahead	One year ahead	Two years ahead
UK GDP growth — mode	-0.38	-0.29	-0.41	-0.14	-0.56	0.45
	(0.02**)	(0.14)	(0.07*)	(0.72)	(0.38)	(0.67)
UK GDP growth — mean	-0.37	-0.39	-0.27	-0.16	-0.95	0.69
	(0.02**)	(0.08*)	(0.21)	(0.66)	(0.31)	(0.42)
Modal projections:						
UK GDP growth — latest data	-0.38	-0.29	-0.41	-0.14	-0.56	0.45
	(0.02**)	(0.14)	(0.07*)	(0.72)	(0.38)	(0.67)
UK GDP growth — real-time data	-0.16	-0.24	-0.30	-0.08	-0.42	0.18
	(0.00**)	(0.06*)	(0.06*)	(0.50)	(0.40)	(0.82)

(a) See footnotes to **Table 3.C** on page 34.

Table A.4.L Weak efficiency of Bank forecasts (full sample including 2008 and 2009)[a]

	One quarter ahead			One year ahead			Two year ahead		
	β_0 (p-value)	β_1 (p-value)	p-value from the joint test: $\beta_0 = 0, \beta_1 = 1$	β_0 (p-value)	β_1 (p-value)	p-value from the joint test: $\beta_0 = 0, \beta_1 = 1$	β_0 (p-value)	β_1 (p-value)	p-value from the joint test: $\beta_0 = 0, \beta_1 = 1$
UK GDP growth	-0.04	1.02	0.98	-0.56	1.06	0.46	4.38	-0.82	0.00**
Inflation	0.13	0.96	0.09*	1.83	0.33	0.00**	4.31	-0.84	0.00**
Unemployment rate	0.47	0.91	0.00**	1.85	0.66	0.00**	2.81	0.52	0.00**
Consumption growth	0.06	1.00	0.97	-0.04	1.07	0.83	0.82	0.54	0.33
Investment growth	0.05	0.74	0.06*	0.01	0.58	0.00**	2.20	-0.22	0.00**
House price growth	0.84	0.99	0.13	4.24	0.87	0.02**	5.75	0.62	0.01**
Wage growth	0.19	0.93	0.47	0.25	0.8	0.01**	-3.54	1.57	0.00**
US GDP growth	-0.26	0.97	0.00**	-0.28	0.88	0.00**	-0.72	0.89	0.15
Euro-area GDP growth	0.03	1.14	0.00**	-0.19	0.95	0.73	0.46	0.21	0.12
Household lending growth	-0.03	0.96	0.80	0.01	0.63	0.00**	n.a.	n.a.	n.a.
Corporate lending growth	-0.89	0.60	0.02**	-2.31	0.16	0.00**	n.a.	n.a.	n.a.

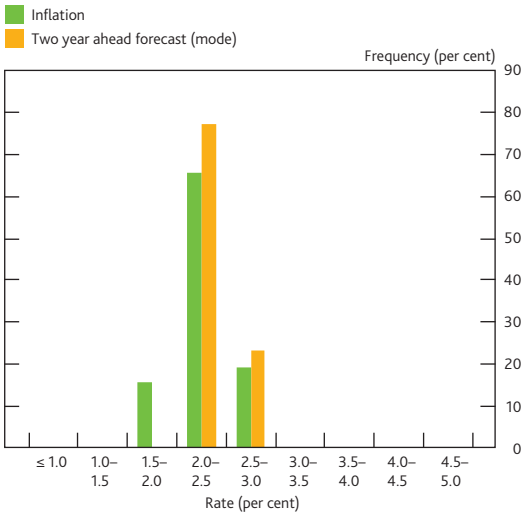
(a) See footnotes to **Table 3.D** on page 35.

Table A.4.M Weak efficiency of Bank forecasts, comparing forecast errors calculated using modal and mean forecasts, and alternative data vintages (full sample excluding 2008 and 2009)[a]

	One quarter ahead			One year ahead			Two year ahead		
	β_0 (p-value)	β_1 (p-value)	p-value from the joint test: $\beta_0 = 0, \beta_1 = 1$	β_0 (p-value)	β_1 (p-value)	p-value from the joint test: $\beta_0 = 0, \beta_1 = 1$	β_0 (p-value)	β_1 (p-value)	p-value from the joint test: $\beta_0 = 0, \beta_1 = 1$
UK GDP growth — mode	1.27	0.56	0.09*	2.02	0.23	0.00**	5.05	-0.83	0.00**
UK GDP growth — mean	1.25	0.58	0.11	1.88	0.30	0.00**	3.84	-0.46	0.00**
Inflation — mode	0.08	0.97	0.50	1.64	0.39	0.00**	4.23	-0.82	0.00**
Inflation — mean	0.10	0.96	0.17	1.47	0.46	0.00**	4.05	-0.71	0.00**
Modal projections:									
UK GDP growth — latest data	1.27	0.56	0.09*	2.02	0.23	0.00**	5.05	-0.83	0.00**
UK GDP growth — real-time data	-0.12	0.97	0.05*	1.10	0.38	0.00**	4.52	-0.91	0.00**
US GDP growth — latest data	0.61	0.68	0.00**	1.38	0.39	0.00**	0.57	0.67	0.07*
US GDP growth — real-time data	0.62	0.78	0.01**	1.55	0.44	0.02**	0.11	0.92	0.80
Euro-area GDP growth — latest data	0.13	1.11	0.00**	0.44	0.75	0.63	0.74	0.39	0.12
Euro-area GDP growth — real-time data	0.04	0.98	0.61	0.34	0.64	0.38	0.76	0.25	0.01**

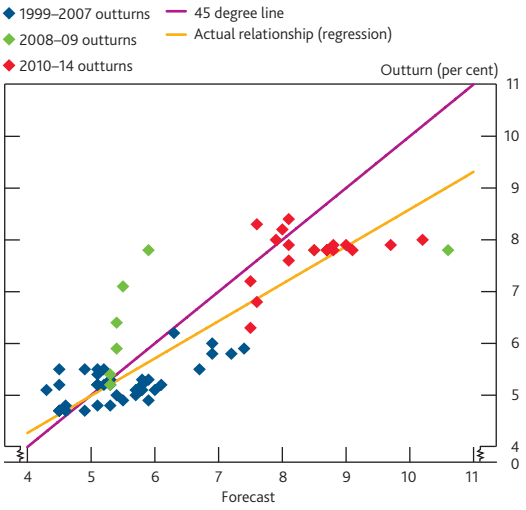
(a) See footnotes to **Table 3.D** on page 35.

Chart A.4.24 Bank two year ahead RPIX inflation forecasts and data outturns[a]



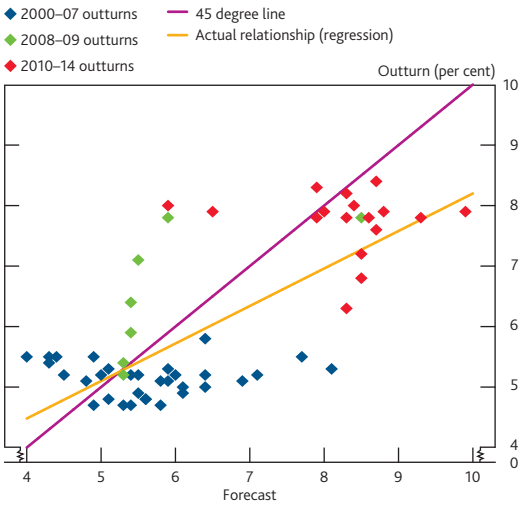
(a) Chart covers the period of two year ahead RPIX forecasts, 1999 Q2–2005 Q3.

Chart A.4.25 Bank one year ahead unemployment rate forecasts and data outturns[a]



(a) See footnote to **Chart 3.12** on page 36.

Chart A.4.26 Bank two year ahead unemployment rate forecasts and data outturns[a]



(a) See footnote to **Chart 3.12** on page 36.

Table A.4.N Predictability of Bank forecasts revisions, when forecasting a fixed horizon (full sample including 2008 and 2009)[a]

	α (p-value)	β_1 (p-value)	β_2 (p-value)	β_3 (p-value)	$\beta_5$4(p-value)	β_5 (p-value)	p-value from the joint test: $\beta_i = 0$ $\forall i = 1, \ldots, 5$
				One quarter ahead			
UK GDP growth (mode)	-0.07	0.38	0.13	0.05	-0.20	0.05	
	(0.16)	(0.03**)	(0.31)	(0.74)	(0.13)	(0.67)	0.03**
UK GDP growth (mean)	-0.05	0.31	0.11	0.00	-0.21	0.08	
	(0.32)	(0.08*)	(0.38)	(0.99)	(0.10)	(0.54)	0.24
Inflation (mode)	0.07	0.10	-0.21	-0.03	-0.07	-0.05	
	(0.17)	(0.55)	(0.36)	(0.87)	(0.66)	(0.75)	0.74
Inflation (mean)	0.01	0.15	-0.12	0.06	-0.09	0.09	
	(0.92)	(0.31)	(0.57)	(0.71)	(0.54)	(0.57)	0.92
Unemployment rate	-0.08	0.20	0.02	0.01	0.01	-0.14	
	(0.07*)	(0.09*)	(0.89)	(0.94)	(0.89)	(0.11)	0.35
Consumption growth	0.06	-0.03	0.09	-0.03	-0.21	0.20	
	(0.44)	(0.79)	(0.46)	(0.82)	(0.11)	(0.08*)	0.16
Investment growth	-0.50	-0.10	-0.06	0.00	0.08	-0.20	
	(0.17)	(0.45)	(0.62)	(1.00)	(0.42)	(0.14)	0.43
House price growth	0.76	0.31	-0.30	0.11	-0.23	0.09	
	(0.11)	(0.05*)	(0.04**)	(0.36)	(0.10*)	(0.43)	0.11
Wage growth	-0.17	-0.19	-0.12	-0.02	-0.27	-0.23	
	(0.02**)	(0.16)	(0.32)	(0.84)	(0.04**)	(0.12)	0.18
US GDP growth	-0.07	0.16	-0.32	-0.03	-0.09	-0.08	
	(0.38)	(0.28)	(0.05*)	(0.81)	(0.51)	(0.53)	0.39
Euro-area GDP growth	-0.03	0.31	-0.14	-0.01	-0.03	0.04	
	(0.62)	(0.17)	(0.44)	(0.97)	(0.82)	(0.79)	0.72
Household lending growth	-0.22	-0.10	-0.20	-0.25	-0.32	0.05	
	(0.24)	(0.44)	(0.39)	(0.19)	(0.06*)	(0.82)	0.20
Corporate lending growth	-0.52	0.17	-0.30	0.40	-0.26	-0.21	
	(0.25)	(0.46)	(0.43)	(0.29)	(0.44)	(0.31)	0.28
				One year ahead			
UK GDP growth (mode)	-0.08	0.24	0.11	-0.06	0.01	0.10	
	(0.22)	(0.16)	(0.39)	(0.74)	(0.94)	(0.40)	0.45
UK GDP growth (mean)	-0.07	0.35	0.08	-0.23	0.10	0.05	
	(0.24)	(0.12)	(0.60)	(0.05*)	(0.37)	(0.50)	0.15
Inflation (mode)	0.12	0.48	-0.27	0.04	0.14	-0.16	
	(0.02**)	(0.00**)	(0.10)	(0.88)	(0.45)	(0.21)	0.00**
Inflation (mean)	0.11	0.49	-0.26	0.04	0.07	-0.10	
	(0.02**)	(0.00**)	(0.14)	(0.76)	(0.58)	(0.42)	0.00**
Unemployment rate	-0.05	0.30	-0.20	-0.05	0.08	-0.13	
	(0.52)	(0.13)	(0.11)	(0.65)	(0.43)	(0.21)	0.29
Consumption growth	0.00	0.04	-0.03	-0.05	-0.12	0.31	
	(1.00)	(0.79)	(0.86)	(0.77)	(0.36)	(0.00**)	0.06*
Investment growth	0.03	0.16	0.01	0.08	-0.08	-0.03	
	(0.93)	(0.19)	(0.89)	(0.41)	(0.53)	(0.88)	0.38
House price growth	1.09	0.22	-0.18	0.19	-0.19	-0.04	
	(0.12)	(0.14)	(0.29)	(0.21)	(0.24)	(0.77)	0.26
Wage growth	-0.20	-0.37	0.03	-0.27	-0.20	0.08	
	(0.05*)	(0.01**)	(0.83)	(0.05*)	(0.12)	(0.57)	0.02**
US GDP growth	-0.08	0.17	0.00	-0.10	0.01	0.19	
	(0.35)	(0.38)	(0.99)	(0.54)	(0.95)	(0.07*)	0.17
Euro-area GDP growth	-0.19	0.55	-0.23	-0.06	0.08	-0.13	
	(0.06*)	(0.02**)	(0.29)	(0.67)	(0.57)	(0.26)	0.03**
Household lending growth	0.04	-0.07	0.23	-0.27	-0.10	-0.04	
	(0.79)	(0.73)	(0.15)	(0.03**)	(0.47)	(0.83)	0.07*
Corporate lending growth	-1.10	0.50	-0.45	0.33	-0.12	-0.48	
	(0.17)	(0.18)	(0.18)	(0.37)	(0.57)	(0.01**)	0.05**

	α (p-value)	β_1 (p-value)	β_2 (p-value)	β_3 (p-value)	β_4 (p-value)	β_5 (p-value)	p-value from the joint test: $\beta_i = 0$ $\forall i = 1, \ldots, 5$
				Two year ahead			
UK GDP growth (mode)	-0.06	-0.30	-0.23	0.04	-0.01	0.21	0.22
	(0.23)	(0.05*)	(0.14)	(0.76)	(0.92)	(0.12)	
UK GDP growth (mean)	-0.05	-0.15	-0.09	-0.05	0.09	-0.01	0.70
	(0.29)	(0.30)	(0.64)	(0.72)	(0.46)	(0.97)	
Inflation (mode)	-0.07	-0.04	-0.16	-0.30	-0.01	-0.04	0.59
	(0.18)	(0.77)	(0.32)	(0.11)	(0.95)	(0.84)	
Inflation (mean)	-0.08	-0.01	-0.42	-0.17	-0.12	-0.03	0.26
	(0.12)	(0.96)	(0.02**)	(0.22)	(0.40)	(0.81)	
Unemployment rate	0.02	0.05	-0.01	-0.10	-0.06	0.07	0.90
	(0.78)	(0.77)	(0.92)	(0.49)	(0.59)	(0.62)	
Consumption growth	-0.14	0.10	0.00	-0.13	0.09	-0.01	0.76
	(0.02**)	(0.39)	(0.99)	(0.39)	(0.54)	(0.96)	
Investment growth	-0.47	-0.17	-0.09	-0.06	0.05	-0.14	0.45
	(0.04**)	(0.25)	(0.56)	(0.65)	(0.62)	(0.28)	
House price growth	-0.24	-0.12	0.11	0.28	0.04	-0.09	0.38
	(0.41)	(0.51)	(0.49)	(0.13)	(0.73)	(0.52)	
Wage growth	-0.11	-0.18	0.10	-0.20	0.12	0.07	0.24
	(0.15)	(0.28)	(0.45)	(0.23)	(0.36)	(0.62)	
US GDP growth	-0.10	-0.35	-0.01	0.11	0.09	0.10	0.55
	(0.20)	(0.26)	(0.94)	(0.29)	(0.55)	(0.61)	
Euro-area GDP growth	-0.12	0.19	0.08	-0.30	-0.04	0.02	0.40
	(0.01**)	(0.31)	(0.66)	(0.15)	(0.77)	(0.89)	

(a) See footnotes to **Table 3.F** on page 39.

Table A.4.O Predictability of Bank forecast revisions when forecasting one quarter and two year ahead (full sample excluding 2008 and 2009)[a]

	α (p-value)	β_1 (p-value)	β_2 (p-value)	β_3 (p-value)	β_4 (p-value)	β_5 (p-value)	p-value from the joint test: $\beta_i = 0$ $\forall i = 1, \ldots, 5$
	One quarter ahead						
Modal forecasts							
UK GDP growth	-0.03	0.17	0.13	0.11	-0.16	0.04	
	(0.54)	(0.22)	(0.32)	(0.54)	(0.23)	(0.68)	0.46
Inflation	0.03	0.09	0.05	0.17	0.04	-0.24	
	(0.48)	(0.55)	(0.69)	(0.13)	(0.77)	(0.03**)	0.02**
Unemployment rate	-0.07	0.11	0.12	0.03	0.04	-0.14	
	(0.06*)	(0.38)	(0.33)	(0.79)	(0.60)	(0.11)	0.29
Consumption growth	0.09	0.03	0.01	-0.01	-0.25	0.28	
	(0.29)	(0.83)	(0.92)	(0.97)	(0.07*)	(0.02**)	0.05*
Investment growth	-0.43	-0.11	-0.03	-0.03	0.06	-0.19	
	(0.25)	(0.45)	(0.78)	(0.80)	(0.59)	(0.17)	0.56
House price growth	1.07	0.09	-0.32	0.03	-0.21	0.07	
	(0.01**)	(0.56)	(0.00**)	(0.76)	(0.09*)	(0.54)	0.03**
Wage growth	-0.15	-0.21	-0.15	-0.05	-0.28	-0.25	
	(0.06*)	(0.11)	(0.26)	(0.61)	(0.05**)	(0.12)	0.20
US GDP growth	-0.03	0.00	-0.17	-0.04	-0.10	-0.04	
	(0.66)	(0.98)	(0.24)	(0.77)	(0.42)	(0.76)	0.82
Euro-area GDP growth	0.01	-0.07	0.02	0.17	0.07	0.04	
	(0.78)	(0.61)	(0.84)	(0.16)	(0.67)	(0.69)	0.66
Household lending growth	-0.05	-0.10	-0.40	-0.05	-0.16	0.14	
	(0.71)	(0.60)	(0.02**)	(0.79)	(0.26)	(0.33)	0.08*
Corporate lending growth	-0.52	0.17	-0.30	0.40	-0.26	-0.21	
	(0.25)	(0.46)	(0.43)	(0.29)	(0.44)	(0.31)	0.28
Memo: mean forecasts							
UK GDP growth	-0.01	0.14	0.11	0.07	-0.19	0.08	
	(0.91)	(0.34)	(0.40)	(0.66)	(0.15)	(0.50)	0.59
Inflation	-0.02	0.23	0.05	0.14	0.00	-0.07	
	(0.67)	(0.17)	(0.69)	(0.23)	(0.99)	(0.61)	0.22
	Two years ahead						
Modal forecasts							
UK GDP growth	-0.05	-0.28	-0.23	0.06	-0.01	0.21	
	(0.38)	(0.09*)	(0.13)	(0.64)	(0.93)	(0.14)	0.31
Inflation	-0.08	-0.06	-0.17	-0.30	-0.02	-0.05	
	(0.17)	(0.70)	(0.32)	(0.13)	(0.93)	(0.83)	0.64
Unemployment rate	0.00	0.06	-0.04	-0.08	-0.07	0.08	
	(0.99)	(0.72)	(0.82)	(0.61)	(0.52)	(0.58)	0.88
Consumption growth	-0.14	0.11	-0.03	-0.12	0.10	-0.02	
	(0.04**)	(0.45)	(0.83)	(0.45)	(0.51)	(0.91)	0.79
Investment growth	-0.48	-0.18	-0.08	-0.06	0.04	-0.14	
	(0.07*)	(0.25)	(0.60)	(0.65)	(0.68)	(0.27)	0.49
House price growth	-0.07	-0.13	0.16	0.36	0.08	-0.10	
	(0.85)	(0.49)	(0.42)	(0.14)	(0.59)	(0.60)	0.36
Wage growth	-0.12	-0.17	0.11	-0.14	0.11	-0.01	
	(0.17)	(0.33)	(0.48)	(0.44)	(0.45)	(0.97)	0.44
US GDP growth	-0.07	-0.37	0.00	0.12	0.09	0.10	
	(0.42)	(0.25)	(0.99)	(0.24)	(0.56)	(0.63)	0.43
Euro-area GDP growth	-0.14	0.15	0.08	-0.31	-0.05	0.02	
	(0.02**)	(0.46)	(0.71)	(0.14)	(0.70)	(0.91)	0.51
Memo: mean forecasts							
UK GDP growth	-0.04	-0.15	-0.10	-0.03	0.09	-0.01	
	(0.46)	(0.34)	(0.62)	(0.85)	(0.47)	(0.96)	0.76
Inflation	-0.09	-0.01	-0.44	-0.17	-0.15	-0.03	
	(0.13)	(0.94)	(0.02**)	(0.26)	(0.35)	(0.82)	0.24

(a) See footnotes to **Table 3.F** on page 39.

Table A.4.P Predictability of Bank forecasts revisions, when forecasting a single calendar quarter (full sample including 2008 and 2009)[a]

	α (p-value)	β_1 (p-value)	β_2 (p-value)	β_3 (p-value)	β_4 (p-value)	β_5 (p-value)	p-value from the joint test: $\beta_i = 0$ $\forall i = 1, \ldots, 5$
UK GDP growth — mode	-0.04	0.30	0.09	0.07	0.02	0.00	
	(0.43)	(0.07*)	(0.24)	(0.55)	(0.86)	(0.99)	0.05**
UK GDP growth — mean	-0.03	0.25	0.04	0.04	0.08	-0.07	
	(0.56)	(0.11)	(0.68)	(0.68)	(0.42)	(0.56)	0.45
Inflation — mode	0.09	0.15	-0.14	0.00	-0.20	-0.22	
	(0.06*)	(0.27)	(0.45)	(0.97)	(0.15)	(0.11)	0.01**
Inflation — mean	0.04	0.14	-0.11	0.00	-0.20	-0.20	
	(0.38)	(0.29)	(0.58)	(0.97)	(0.06*)	(0.06*)	0.00**
Unemployment rate	-0.10	0.05	-0.12	-0.04	-0.04	-0.17	
	(0.03**)	(0.55)	(0.26)	(0.50)	(0.58)	(0.00**)	0.00**
Consumption growth	0.08	-0.07	-0.04	0.00	0.03	0.18	
	(0.28)	(0.56)	(0.71)	(0.95)	(0.73)	(0.08*)	0.34
Investment growth	-0.41	-0.15	-0.09	-0.05	0.03	0.12	
	(0.26)	(0.16)	(0.41)	(0.67)	(0.86)	(0.48)	0.70
House price growth	0.82	0.15	-0.13	0.06	-0.21	-0.12	
	(0.04**)	(0.12)	(0.13)	(0.38)	(0.06*)	(0.29)	0.11
Wage growth	-0.25	-0.13	-0.28	-0.12	-0.17	-0.25	
	(0.00**)	(0.20)	(0.03**)	(0.19)	(0.18)	(0.06*)	0.12
US GDP growth	-0.04	0.11	-0.13	-0.10	0.01	0.18	
	(0.52)	(0.41)	(0.31)	(0.43)	(0.94)	(0.17)	0.05**
Euro-area GDP growth	0.00	0.32	-0.10	-0.10	0.18	0.00	
	(0.98)	(0.07*)	(0.49)	(0.34)	(0.14)	(1.00)	0.18
Household lending growth	-0.24	-0.21	-0.25	-0.16	-0.03	-0.05	
	(0.22)	(0.13)	(0.08*)	(0.09*)	(0.85)	(0.70)	0.03**
Corporate lending growth	-1.26	-0.04	-0.70	0.15	-0.66	-0.22	
	(0.03**)	(0.87)	(0.00**)	(0.55)	(0.02**)	(0.14)	0.04**

(a) See footnotes to **Table 3.E** on page 38.

Accuracy of Bank forecasts compared with a simple forecasting model

Table A.4.Q Accuracy of forecasts from random walk without drift processes relative to accuracy of Bank modal forecasts (full sample including 2008 and 2009)[a]

	Ratio of RMSEs: random-walk without drift forecast accuracy relative to Bank forecast accuracy			p-value from Diebold-Mariano test of differences in RMSEs		
	One quarter ahead	One year ahead	Two years ahead	One quarter ahead	One year ahead	Two years ahead
UK GDP growth	1.31	1.46	1.20	0.10	0.13	0.11
Inflation	2.66	1.13	0.88	0.01**	0.39	0.55
Unemployment rate	0.84	0.74	0.81	0.21	0.07*	0.05*
Consumption growth	1.18	1.36	1.04	0.13	0.25	0.72
Investment growth	1.16	1.48	1.36	0.14	0.08*	0.02**
House price growth	1.52	1.20	1.06	0.02**	0.24	0.69
Wage growth	1.46	1.00	0.87	0.05*	1.00	0.23
US GDP growth	1.42	1.56	1.25	0.05**	0.07*	0.08*
Euro-area GDP growth	2.09	1.67	1.35	0.05**	0.09*	0.06*
Household lending growth	1.58	1.65	n.a.	0.03**	0.19	n.a.
Corporate lending growth	1.18	1.02	n.a.	0.28	0.95	n.a.

(a) See footnotes to **Table 3.G** on page 40.

Table A.4.R Accuracy of forecasts from random walk without drift processes relative to accuracy of Bank forecasts (full sample excluding 2008 and 2009), assessed using modal and mean projections, and latest and real-time data[a]

	Ratio of RMSEs: random-walk without drift forecast accuracy relative to Bank forecast accuracy			p-value from Diebold-Mariano test of differences in RMSEs		
	One quarter ahead	One year ahead	Two years ahead	One quarter ahead	One year ahead	Two years ahead
UK GDP growth — mode	1.21	1.75	1.55	0.11	0.20	0.14
UK GDP growth — mean	1.21	1.76	1.75	0.10	0.19	0.14
Inflation — mode	2.34	1.02	0.90	0.00**	0.91	0.67
Inflation — mean	2.32	1.08	0.92	0.00**	0.70	0.73
Modal projections:						
UK GDP growth — latest data	1.21	1.75	1.55	0.11	0.20	0.14
UK GDP growth — real-time data	1.42	1.81	1.32	0.16	0.18	0.17
US GDP growth — latest data	1.23	1.83	2.35	0.05*	0.13	0.03**
US GDP growth — real-time data	1.39	1.93	2.58	0.05*	0.11	0.04**
Euro-area GDP growth — latest data	1.81	2.14	1.78	0.07*	0.15	0.15
Euro-area GDP growth — real-time data	3.01	2.23	1.71	0.04**	0.16	0.22

(a) See footnotes to **Table 3.G** on page 40.

International comparisons of forecast accuracy

Table A.4.S Accuracy of one year and two year ahead calendar year forecasts made by other central banks, relative to Bank forecast accuracy[a]

RMSE relative to RMSE for Bank forecasts, 1998–2014

	GDP growth		Inflation		Unemployment rate		Consumption growth		Investment growth	
	One year ahead	Two years ahead	One year ahead	Two years ahead	One year ahead	Two years ahead	One year ahead	Two years ahead	One year ahead	Two years ahead
ECB	0.8	1.0	0.7	0.6	n.a.	n.a.	0.9	0.7	0.4	0.7
Fed	0.9	0.9	0.8	0.5	0.7	1.2	n.a.	n.a.	n.a.	n.a.
RBNZ	1.0	0.8	1.2	0.8	0.7	0.6	1.0	0.9	0.9	1.1
Riksbank	1.3	1.2	1.5	1.0	n.a.	n.a.	1.2	0.8	0.7	1.0

RMSE relative to RMSE for Bank forecasts, for one year ahead forecasts

	GDP growth		Inflation		Unemployment rate[a]		Consumption growth		Investment growth	
	1998–2007	2010–14	1998–2007	2010–14	1999–2007	2010–14	1998–2007	2010–14	1998–2007	2010–14
ECB	1.0	1.1	1.1	0.8	n.a.	n.a.	1.0	1.9	0.3	0.3
Fed	1.0	0.9	2.1	0.4	0.6	0.7	n.a.	n.a.	n.a.	n.a.
RBNZ	1.2	1.4	2.5	0.8	0.6	1.2	1.1	1.7	1.1	0.7
Riksbank	1.0	3.1	3.1	1.1	n.a.	n.a.	0.9	2.5	0.8	0.7

RMSE relative to RMSE for Bank forecasts, for two year ahead forecasts

	GDP growth		Inflation		Unemployment rate[a]		Consumption growth		Investment growth	
	1999–2007	2010–14	1999–2007	2010–14	2000–07	2010–14	1999–2007	2010–14	1999–2007	2010–14
ECB	1.9	1.1	1.4	0.4	n.a.	n.a.	0.7	1.0	1.6	0.6
Fed	1.6	0.8	2.2	0.2	0.6	0.7	n.a.	n.a.	n.a.	n.a.
RBNZ	1.3	1.0	2.9	0.4	0.6	0.7	1.4	0.8	2.4	1.0
Riksbank	1.5	1.5	2.3	0.7	n.a.	n.a.	0.9	0.8	1.4	0.9

(a) See footnotes to **Table 3.J** on page 45.

Table A.4.T Accuracy of one year and two year ahead US GDP growth forecasts made by the Bank and the Fed[a]

RMSEs

	2001–14		2001–07		2010–14	
	One year ahead	Two years ahead	One year ahead	Two years ahead	One year ahead	Two years ahead
February Monetary Policy Reports (Fed)	1.6	2.2	1.3	1.6	1.0	1.5
February *Inflation Reports* (Bank)	1.6	1.8	1.3	0.9	1.1	0.9
November *Inflation Reports* (Bank)	1.5	2.0	0.9	0.7	0.5	0.7

(a) See footnotes to **Table 3.M** on page 48.

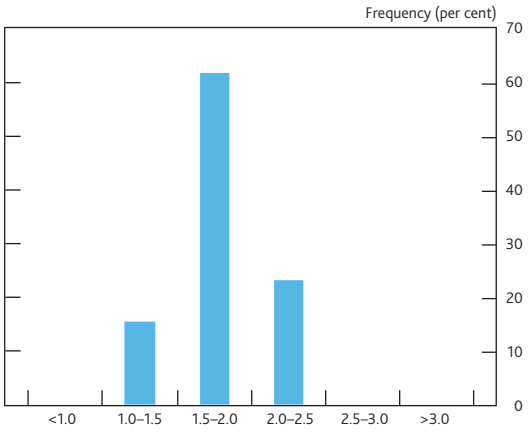
Table A.4.U Accuracy of one year and two year ahead euro-area GDP growth forecasts made by the Bank and the European Central Bank[a]

RMSEs

	2001–14		2001–07		2010–14	
	One year ahead	Two years ahead	One year ahead	Two years ahead	One year ahead	Two years ahead
March macroeconomic projections (ECB)	0.8	2.1	0.8	1.4	0.6	1.6
May *Inflation Reports* (Bank)	0.8	2.1	0.9	1.4	0.6	1.8
February *Inflation Reports* (Bank)	1.0	2.1	1.2	1.4	0.5	1.4

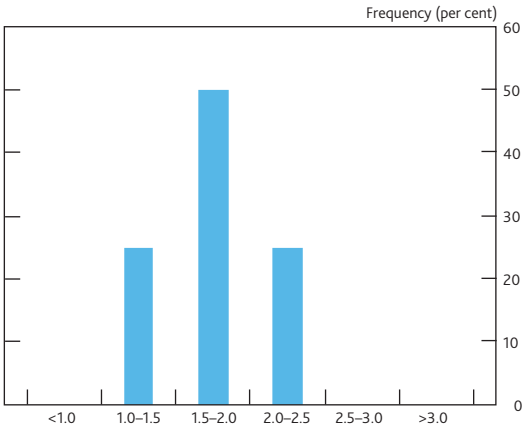
(a) See footnotes to **Table 3.N** on page 49.

Chart A.4.27 Distribution of ECB's two year ahead inflation forecasts[a]



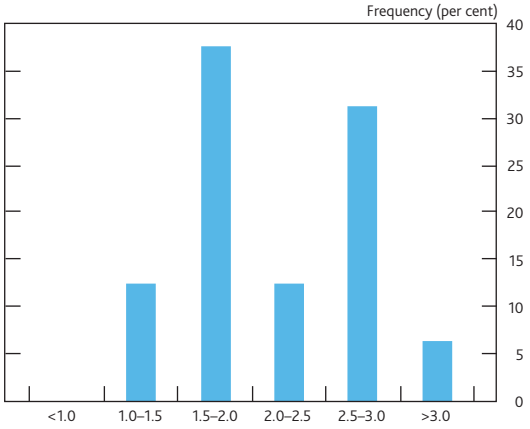
(a) Calendar year HICP inflation. 2002–14, with one forecast from each year. See **Table A.3.B** for more details of the forecasts included in this chart.

Chart A.4.28 Distribution of Fed's two year ahead inflation forecasts[a]



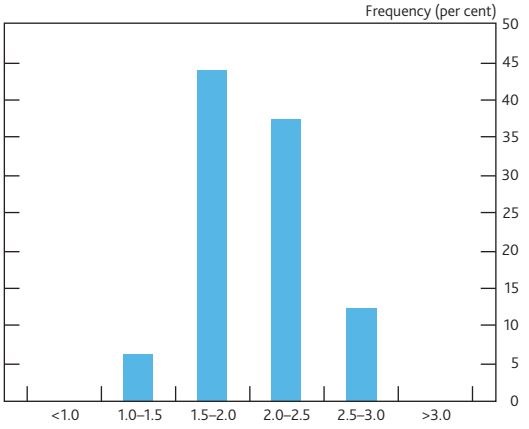
(a) Q4-on-Q4 inflation; the forecast definition has changed over time, see **Table A.3.C**. 1999–2014, with one forecast from each year. See **Table A.3.C** for more details of the forecasts included in this chart.

Chart A.4.29 Distribution of RBNZ's two year ahead inflation forecasts[a]



(a) Calendar year CPI inflation. 1999–2014, with one forecast from each year. See **Table A.3.D** for more details of the forecasts included in this chart.

Chart A.4.30 Distribution of Riksbank's two year ahead inflation forecasts[a]



(a) Calendar year CPI inflation. 1999–2014, with one forecast from each year. See **Table A.3.E** for more details of the forecasts included in this chart.

References

Abreu, I (2011), 'International organisations' vs. private analysts' forecasts: an evaluation', *Working Papers 20-2011*, Economics and Research Department, Banco de Portugal.

Alessi, L, Ghysels, E, Onorante, L, Peach, R and Potter, S (2014), 'Central bank macroeconomic forecasting during the global financial crisis: the European Central Bank and Federal Reserve Bank of New York experiences', *Journal of Business & Economic Statistics*, Vol. 32, No. 4, pages 483–500.

Amisano, G and Giacomini, R (2007), 'Comparing density forecasts via weighted likelihood ratio tests', *Journal of Business & Economic Statistics*, Vol. 25, No. 2, pages 177–90.

Andersson, M K, Karlsson, G and Svensson, J (2007), 'The Riksbank's forecasting performance', *Working Paper Series 218*, Sveriges Riksbank.

Andrews, D W K (1991), 'Heteroskedasticity and autocorrelation consistent covariance matrix estimation', *Econometrica*, Vol. 59, No. 3, pages 817–58.

Andrews, D W K and Monahan, J C (1992), 'An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator', *Econometrica*, Vol. 60, No. 4, pages 953–66.

Bank of England (1999), 'Economic models at the Bank of England', available at http://webarchive.nationalarchives.gov.uk/20111203050222/http:/www.bankofengland.co.uk/publications/other/beqm/modcobook.htm.

Bank of England (2000), 'Economic models at the Bank of England: September 2000 update', available at http://webarchive.nationalarchives.gov.uk/20111203050222/http:/www.bankofengland.co.uk/publications/other/beqm/modcobook.htm.

Barnett, A, Batten, S, Chiu, A, Franklin, J and Sebastiá-Barriel, M (2014), 'The UK productivity puzzle', *Bank of England Quarterly Bulletin*, Vol. 54, No. 2, pages 114–28, available at www.bankofengland.co.uk/publications/Documents/quarterlybulletin/2014/qb14q201.pdf.

Bates, J M and Granger, C W J (1969), 'The combination of forecasts', *Operational Research Quarterly*, Vol. 20, No. 4, pages 451–68.

Bean, C (2012), 'Central banking in boom and slump', the JSG Wilson Lecture in Economics, University of Hull

Bean, C and Jenkinson, N (2001), 'The formulation of monetary policy at the Bank of England', *Bank of England Quarterly Bulletin*, Winter, pages 434–41, available at www.bankofengland.co.uk/archive/Documents/historicpubs/qb/2001/qb010403.pdf

Benito, A, Whitley, J and Young, G (2001), 'Analysing corporate and household sector balance sheets', *Bank of England Financial Stability Review*, Issue 11, December, pages 160–74, available at www.bankofengland.co.uk/archive/Documents/historicpubs/fsr/2001/fsrfull0112.pdf.

Berkowitz, J (2001), 'Testing density forecasts, with applications to risk management', *Journal of Business & Economic Statistics*, Vol. 19, No. 4, pages 465–74.

Blix, M, Wadefjord, J, Wienecke, U and Ådahl, M (2001), 'How Good is the forecasting performance of major institutions?', *Sveriges Riksbank Economic Review*, No. 3, pages 38–68.

Boero, G, Smith, J and Wallis, K F (2008), 'Evaluating a three-dimensional panel of point forecasts: the Bank of England Survey of External Forecasters', *International Journal of Forecasting*, Vol. 24, No. 3, pages 354–67.

Broadbent, B (2012), 'Productivity and the allocation of resources', available at www.bankofengland.co.uk/publications/Documents/speeches/2012/speech599.pdf.

Broadbent, B (2013), 'Forecast errors', available at www.bankofengland.co.uk/publications/Documents/speeches/2013/speech653.pdf.

Burgess, S, Fernandez-Corugedo, E, Groth, C, Harrison, R, Monti, F, Theodoridis, K, Waldron, M (2013), 'The Bank of England's forecasting platform: COMPASS, MAPS, EASE and the suite of models', *Bank of England Working Paper No. 471*, available at www.bankofengland.co.uk/research/Documents/workingpapers/2013/wp471.pdf.

Carney, M (2014), 'One Misson. One Bank. Promoting the good of the people of the United Kingdom', available at www.bankofengland.co.uk/publications/Documents/speeches/2014/speech715.pdf.

Carney, M (2015), 'From Lincoln to Lothbury: Magna Carta and the Bank of England', available at www.bankofengland.co.uk/publications/Documents/Speeches/2015/speech832.pdf.

Clark, T and McCracken, M (2013), 'Advances in forecast evaluation', in Elliott, G and Timmermann, A (eds), *Handbook of Economic Forecasting*, Vol. 2, Part B, pages 1,107–201.

Clements, M (1997), 'Evaluating the rationality of fixed-event forecasts', *Journal of Forecasting*, Vol. 16, No. 4, pages 225–39.

Clements, M (2004), 'Evaluating the Bank of England density forecasts of inflation', *The Economic Journal*, Vol. 114 (498), pages 844–66.

Clements, M and Hendry, D (1998), *Forecasting Economic Time Series*, Cambridge University Press, Cambridge.

Clements, M and Hendry, D (2004), 'Pooling of Forecasts, *Econometrics Journal*, Vol. 7, pages 1–31.

Croushore, D (2012), 'Forecast Bias in two dimensions', *Working Paper No. 12-9*, Federal Reserve Bank of Philadelphia, Philadelphia.

Cunningham, A and Jeffery, C (2007), 'Extracting a better signal from uncertain data', *Bank of England Quarterly Bulletin*, Vol. 47, No. 3, pages 364–75, available at www.bankofengland.co.uk/publications/Documents/quarterlybulletin/qb070301.pdf.

De Bondt, W F M and Bange, M M (1992), 'Inflation forecast errors and time variation in term premia', *The Journal of Financial and Quantitative Analysis*, Vol. 27, No. 4, pages 479–96.

Delbrück, F, Dunstan, A, Hargreaves, D, Lienert, A, Pepper, H and Sleeman, C (2008), 'The evolution of the forecasting and policy system (FPS) at the Reserve Bank of New Zealand', *Discussion Paper Series DP2008/19*, Reserve Bank of New Zealand.

Den Haan, W J and Levin, A T (1996), 'A practitioner's guide to robust covariance matrix estimation', *Technical Working Paper Series 197*, National Bureau of Economic Research.

Diebold, F X and Lopez, J A (1996), 'Forecast evaluation and combination', in Maddala, G S and Rao, C R (eds), *Handbook of Statistics*, Vol. 14, pages 241–68.

Diebold, F X and Mariano, R S (1995), 'Comparing predictive accuracy', *Journal of Business & Economic Statistics*, Vol. 13, No. 3, pages 253–63.

Doornik, J A and Hansen, H (2008), 'An omnibus test for univariate and multivariate normality', *Oxford Bulletin of Economics and Statistics*, Vol. 70, No. S1, pages 927–39.

Eitrheim, Ø, Husebø, T A and Nymoen, R (2004), 'Empirical comparisons of inflation models' forecast accuracy', in Clements, M P and Hendry, D F (eds), *A Companion to Economic Forecasting*, Blackwell Publishing.

Elder, R, Kapetanios, G, Taylor, T and Yates, T (2005), 'Assessing the MPC's fan charts', *Bank of England Quarterly Bulletin*, Autumn, pages 326–48, available at www.bankofengland.co.uk/publications/Documents/quarterlybulletin/qb050301.pdf.

Elliott, G, Komunjer, I and Timmermann, A (2005), 'Estimation and testing of forecast rationality under flexible loss', The *Review of Economic Studies*, Vol. 72, No. 4, pages 1,107–125.

Elliott, G, Komunjer, I and Timmermann, A (2008), 'Biases in macroeconomic forecasts: irrationality or asymmetric loss?', *Journal of the European Economic Association*, Vol. 6, No. 1, pages 122–57.

El-Shagi, M, Giesen, S and Jung, A (2014), 'Does the Federal Reserve staff still beat private forecasters?', *Working Paper Series 1635*, European Central Bank.

Ericsson, N R (2001), 'Forecast uncertainty in economic modelling', *International Finance Discussion Papers 697*, Board of Governors of the Federal Reserve System.

European Central Bank (2013), 'An assessment of Eurosystem staff macroeconomic projections', *Monthly Bulletin*, May.

Faust, J, Rogers, J H and Wright, J H (2005), 'News and noise in G-7 GDP announcements', *Journal of Money, Credit, and Banking*, Vol. 37, No. 3, pages 403–19.

Faust, J and Wright, J H (2008), 'Efficient forecast tests for conditional policy forecasts', *Journal of Econometrics*, Vol. 146, No. 2, pages 293–03.

Fawcett, N, Körber, L, Masolo, R M and Waldron, M (2015), 'Evaluating UK point and density forecasts from an estimated DSGE model: the role of off-model information over the financial crisis', *Bank of England Working Paper No. 538*, available at www.bankofengland.co.uk/research/Documents/workingpapers/2015/swp538.pdf.

Freedman, C (2014), 'An evaluation of commissioned studies assessing the accuracy of IMF forecasts', *Background Paper BP/14/02*, Independent Evaluation Office of the International Monetary Fund, Washington.

Galbraith, J W and Tkacz, G (2007), 'Forecast content and content horizons for some important macroeconomic time series', *Canadian Journal of Economics*, Vol. 40, No. 3, 935–53.

George, E A J (1999), Foreword to *Economic Models at the Bank of England*, available at http://webarchive.nationalarchives.gov.uk/20111203050222/http:/www.bankofengland.co.uk/publications/other/beqm/modcobook.htm.

Genberg, H A J and Martinez, A (2014), 'On the accuracy and efficiency of IMF forecasts: a survey and some extensions', *Background Paper BP/14/04*, Independent Evaluation Office of the International Monetary Fund, Washington.

Giacomini, R and Rossi, B (2010), 'Forecast comparisons in unstable environments', *Journal of Applied Econometrics*, Vol. 25, No. 4, pages 595–20.

Gneiting, T (2011), 'Making and evaluating point forecasts', *Journal of the American Statistical Association*, Vol. 106 (494), pages 746–62.

Greene, W H (2011), *Econometric Analysis*, 7th Edition, Prentice-Hall, New Jersey.

Greenspan, A (2005), 'Reflections on central banking', remarks delivered at a symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming.

Groen, J J J, Kapetanios, G and Price, S (2009), 'A real time evaluation of Bank of England forecasts of inflation and growth', *International Journal of Forecasting*, Vol. 25, No. 1, pages 74–80.

Hackworth, C, Radia, A and Roberts, N (2013), 'Understanding the MPC's forecast performance since mid-2010', *Bank of England Quarterly Bulletin*, Vol. 53, No. 4, pages 336–50, available at www.bankofengland.co.uk/publications/Documents/quarterlybulletin/2013/qb130405.pdf.

Harrison, R, Nikolov, K, Quinn, M, Ramsay, G, Scott, A and Thomas, R (2005), The Bank of England Quarterly Model, available at http://webarchive.nationalarchives.gov.uk/20100114080129/http://www.bankofengland.co.uk/publications/other/beqm/beqmfull.pdf

Harvey, D, Leybourne, S and Newbold, P (1997), 'Testing the equality of prediction mean squared errors', *International Journal of forecasting*, Vol. 13, No. 2, pages 281–91.

Holly, S and Weale, M (eds) (2000), *Econometric Modelling: Techniques and Applications*, Cambridge University Press, Cambridge.

Independent Evaluation Office of the International Monetary Fund (2014), *IMF Forecasts: Process, Quality, and Country Perspectives*, Evaluation Report, Independent Evaluation Office of the International Monetary Fund, Washington.

Kapetanios, G, Labhard, V and Price, S (2008), 'Forecast combination and the Bank of England's suite of statistical forecasting models', *Economic Modelling*, Vol. 25, No. 4, pages 772–92.

King, M (2010), 'Uncertainty in macroeconomic policy making — art or science', available at www.bankofengland.co.uk/archive/Documents/historicpubs/speeches/2010/speech432.pdf

Lansing, K J and Pyle, B (2015), 'Persistent overoptimism about economic growth', *Federal Reserve Bank of San Francisco Economic Letter 2015-03*.

Labbé, F and Pepper, H (2009), 'Assessing recent external forecasts', *Reserve Bank of New Zealand Bulletin*, Vol. 72, No. 4, pages 19–25.

McKeown, J and Paterson, L (2014), 'Enhancing the transparency of the Bank of England's Inflation Report', available at www.voxeu.org/article/enhancing-transparency-bank-england-s-inflation-report.

Mincer, J A and Zarnowitz, V (1969), 'The evaluation of economic forecasts', in Mincer, J A (ed), *Economic Forecasts and Expectations: Analysis of Forecasting Behaviour and Performance*, National Bureau of Economic Research, New York.

Mitchell, J (2008), 'Density forecast revisions and forecast efficiency', National Institute of Economic and Social Research.

Murphy, E and Senior, S (2013), 'Changes at the Bank of England', *Bank of England Quarterly Bulletin*, Vol. 53, No.1, pages 20–28, available at www.bankofengland.co.uk/publications/Documents/quarterlybulletin/2013/qb130102.pdf.

Müller, U K (2014), 'HAC corrections for strongly autocorrelated time series', *Journal of Business & Economic Statistics*, Vol. 32, No. 3, pages 311–22.

Nelson, C R (1972), 'The prediction performance of the FRB-MIT-PENN model of the U.S. Economy', *American Economic Review*, Vol. 62, No. 5, pages 902–17.

Newey, W K and West, K D (1987), 'A Simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix', *Econometrica*, Vol. 55, No. 3, pages 703–08.

Ng, S and Wright, J H (2013), 'Facts and challenges from the Great Recession for forecasting and macroeconomic modeling', *Journal of Economic Literature*, Vol. 51, No. 4, pages 1,120–54.

Nickell, S and Quintini, G (2001), 'The recent performance of the UK labour market', speech given to the Economics Section of the British Association for the Advancement of Science, Glasgow.

Nordhaus, W D (1987), 'Forecasting efficiency: concepts and applications', *The Review of Economics and Statistics*, Vol. 69, No. 4, pages 667–74.

Pain, N, Lewis, C, Dang, T, Jin, Y and Richardson, P (2014), 'OECD forecasts during and after the financial crisis: a post mortem', *OECD Economics Department Working Paper 1107*, OECD Publishing.

Patton, A J and Timmermann, A (2007), 'Properties of optimal forecasts under asymmetric loss and nonlinearity', *Journal of Econometrics*, Vol. 140, pages 884–918.

Pesaran, M H and Timmermann, A (1992), 'A simple nonparametric test of predictive performance', *Journal of Business & Economic Statistics*, Vol. 10, No. 4, pages 461–65.

Romer, C D and Romer, D H (2000), 'Federal Reserve information and the behavior of interest rates', *American Economic Review*, Vol. 90, No. 3, pages 429–57.

Rossi, B and Sekhposyan, T (2015), 'Forecast rationality tests in the presence of instabilities, with applications to Federal Reserve and survey forecasts', *Journal of Applied Econometrics*.

Stockton, D (2012), 'Review of the Monetary Policy Committee's Forecasting Capability', available at www.bankofengland.co.uk/publications/Documents/news/2012/cr3stockton.pdf.

Stratford, K (2013), 'Nowcasting world GDP and trade using global indicators, *Bank of England Quarterly Bulletin*, Vol. 53, No. 3, pages 233–43, available at www.bankofengland.co.uk/publications/Documents/quarterlybulletin/2013/qb130305.pdf.

Timmermann, A (2006), 'An evaluation of the World Economic Outlook Forecasts', *International Monetary Fund Working Paper WP/06/59*.

Wallis, K F (2004), 'An assessment of Bank of England and National Institute inflation forecast uncertainties', *National Institute Economic Review*, Vol. 189, pages 64–71.

Warsh, K (2014), 'Transparency and the Bank of England's Monetary Policy Committee', available at www.bankofengland.co.uk/publications/Documents/news/2014/warsh.pdf.

Weale, M (2011), 'Uncertain uncertainty', available at www.bankofengland.co.uk/archive/Documents/historicpubs/speeches/2011/speech486.pdf.