Forecasting UK Inflation: An Empirical Analysis

Jared Golden
Barrie and Hibbert Ltd, Edinburgh, UK

Andrew Adams¹
Heriot-Watt University, UK

Zhuoshi Liu
Barrie and Hibbert Ltd, Edinburgh, UK

Steffen Sørensen
Moody's Corporation, New York, USA

June 2010
Discussion Paper 2010/02

Abstract
This paper considers different ways of forecasting UK RPI inflation. We show that the inflation risk premium on nominal gilts and inflation swaps vary significantly over time. The average inflation risk premia on both of these market instruments have increased considerably since 2004 and during this time the term structure of the inflation risk premium has changed from flat to upward sloping. The econometric model designed for this study is found to produce more accurate inflation forecasts than market break-even inflation rates and surveys by Consensus Economics. However, the evidence suggests that the best approach to forecasting inflation combines forward-looking and historical methods.

Keywords: forecasting inflation; break-even inflation; market expectations; inflation risk premium.

JEL Classification: C40, E43, E52.

¹ Corresponding author: Centre for Finance and Investment, School of Management and Languages, Heriot-Watt University, Edinburgh EH14 4AS. E-mail: A.T.Adams@hw.ac.uk.
Forecasting UK Inflation: An Empirical Analysis

1. Introduction

This paper analyses different ways of forecasting future UK inflation. As we include market prices, our analysis effectively is a test of the inflation expectations hypothesis and whether nominal bond yields contain an inflation risk premium.

Consensus forecasts and econometric model forecasts are not complicated by the inclusion of inflation risk premia, but consensus forecasts do not go beyond ten years and econometric models must deal with the complexity of structural breaks in making long-term forecasts. Therefore, even if break-even inflation (BEI) forecasts from gilts and inflation swaps are not as accurate as the other two types of inflation forecast, analysis of the inflation risk premium on these products, which is forward-looking, should be useful in understanding the market’s long-term expectations of inflation. This information is also useful for the Bank of England because the inflation risk premium represents the amount that the UK Government could save by issuing index-linked gilts instead of nominal gilts at the same maturities (Shen, 1998).

The simplest measure of expected inflation that can be extracted from gilts is break-even inflation (BEI), which is usually defined as the difference between the continuously compounded nominal and real zero coupon yields on otherwise identical or similar gilts. This is equivalent to the Fisher equation,

\[ \text{nominal yield} = \text{real yield} + \text{break-even inflation}, \]  

(1)

where the nominal and real yields are continuously compounded and inflation includes an inflation risk premium.

The only other financial instrument that is pinned to inflation, has sufficiently deep liquidity and enough variation in maturities to build an inflation curve is the inflation swap. An inflation swap is an agreement to exchange, at regular intervals, a fixed rate on a notional amount for a cash flow equal to the percentage change in an inflation index over a pre-specified recent period multiplied by the same notional amount. A typical swap contract is constructed so that the value at initiation is zero. Information required to build a BEI curve is provided by brokers in the form of inflation swap rates. For the purpose of comparison, these need to be made consistent with the measures of break-even inflation derived from gilts.

Assuming frictionless and arbitrage-free markets, the break-even inflation figures derived from index-linked gilts and inflation swaps should be equal. For all intents and purposes inflation swaps can also be considered free of default risk and therefore, since both financial instruments derive their value from the same retail prices index, there should theoretically be an arbitrage opportunity if the BEI figures diverge. However, there are several instances of persistent discrepancies in the BEI figures in recent years. Disparity between inflation forecasts of the investors in the two types of instruments would not cause this, because arbitrageurs would seize the opportunity to make a risk-free profit and force the BEI figures to converge almost immediately. Instead, the discrepancies are explained by practical differences between index-linked gilts and inflation swaps. The most important practical problem is incomplete markets. Hurd and Relleen (2006) explain that the severity of incomplete markets depends on the maturities of the instruments and the problem is exacerbated by trading costs, which are higher on instruments with low levels of liquidity, hence compounding the barriers to arbitrage.

For maturities where the markets for inflation derivatives and other inflation-linked instruments are incomplete, discrepancies in break-even inflation can occur due to demand and supply factors. Inflation swap markets have a reputation of excess participants wishing to receive inflation and pay fixed, such that the break-even inflation is higher than that from index-linked gilts (Hurd and Relleen, 2006). Another factor that Hurd and Relleen identify is that, when there are barriers to arbitrage, distortions in the nominal yield curve directly affect gilt break-even inflation but do not affect inflation swaps.

Quoted RPI inflation percentages generally refer to annual changes in the index level. Using these figures with monthly frequency implies a data overlap and thus autocorrelation. Calculating the month-on-month percentage changes removes the overlap but has its own problem, seasonality. Realised inflation data are subject to significant, repeatable seasonal patterns that are caused by consumption behaviour, manifested in shocks in volume traded during retail sales periods and higher energy and food consumption in the winter months (Belgrade and Benhamou, 2005).
An index-linked gilt’s nominal return is affected by the reference month due to the seasonal pattern of inflation. McGrath and Windle (2006) explain that different final reference months for two otherwise identical gilts will cause greatest differences in their break-even inflation when the gilts are close to maturity. A similar seasonality problem is experienced with swaps. It is thus necessary to adjust for the seasonal pattern in time-series models of inflation.

Cuthbertson and Nitzsche (2004) point out that the inflation risk premium is an important part of the term structure of inflation (and thus also nominal interest rates). Even though the inflation risk premium is a component of the yield on nominal securities (such as nominal gilts), it is unobservable because we need to know the implied pure inflation forecast in order to observe the risk premia. Hence, estimation techniques are required for decomposition of BEI forecasts so that we can determine the inflation risk premium that is present for a particular instrument, for a particular forecast length.

The premium is usually positive to compensate investors for the uncertainty of future inflation, but this need not be the case. Joyce, Sørensen and Weeken (2008) explain that the sign of the inflation risk premium depends on, when the economy is weak, whether people expect worse times in the near future, or rather a quick recovery. If, for example, people are expecting the economy to recover quickly then they will borrow now to smooth overall (present and future) consumption. Nominal assets are more desirable than real assets under these conditions because they provide a better hedge against the possible outcome (contrary to expectations) of an even weaker economy and lower inflation. The hedge exists because the market weakness is at least partially offset by the benefit of holding nominal versus real securities when realised inflation is lower than expected. Thus, even though people are expecting higher inflation in the future (when the economy recovers) the yield gap between nominal and real gilts will be smaller than expected inflation (because the relative increase in demand for nominal securities over real securities narrows their yield gap), reflecting a negative inflation risk premium since the inflation forecast is fixed. Figure 1 provides a graphical representation of the above argument. Similarly, opposite market expectations will cause a positive inflation risk premium. Remolona, Wickens and Gong (1998) conducted a study of market inflation expectations from the early 1980s to the late 1990s using UK gilt trading data. They found that the average inflation risk premium over the period was +100 basis points (1%).

There are several approaches to producing econometric forecasts of inflation. Affine term structure models assume that interest rates follow a random walk. Duffee (2002) devised an improved version called essentially affine term structure models that allow interest rate risk to move independently of their volatility. This has a direct application for inflation forecasting because the uncertainty regarding future nominal interest rates consists of the real forward term premium and the inflation risk premium (Durham, 2006).

Recent papers, such as Dewachter and Lyrio (2006), Christensen, Lopez and Rudebusch (2008) Liu and Spencer (2010), and Joyce, Lildholdt and Sørensen (2010), have used a no-arbitrage, essentially affine term structure framework to model inflation. With this approach, the forward looking inflation expectations and the inflation risk premia can be extracted from the nominal and/or real yield curves.

Some of these affine models have made assumptions on the relationship between the nominal yield curve and inflation. For example, Dewachter and Lyrio (2006) and Liu and Spencer (2010) have assumed that the dynamics of the long-term nominal rate is primarily driven by long-term inflation expectations. Although this assumption is consistent with standard economic theory (e.g., the Fisher equation) and is supported by some empirical evidence (Moazzami (1990) and Choi (2000)), it has been challenged by other researchers (Viren (1989) and Gupta (2002)). On the other hand, Christensen, Lopez and Rudebusch (2008) and Joyce, Lildholdt and Sørensen (2010) build models based on both nominal and real yield curves that allow for negative correlation between long-term inflation expectations and the long-term real interest rates, rather than using the...
more computationally convenient assumption of independence between the two.

Although affine models are good tools for studying inflation expectations and risk premia, they fall short on flexibility and simplicity in terms of producing inflation forecasts. This paper shows that we can use a simple econometric model to produce inflation forecasts that are superior to break-even inflation market forecasts and consensus forecasts. In addition, our econometric model enables us to decompose break-even inflation rates from bond and inflation swap market prices into expectations and risk premia.

2. Data and methodology

The data used in this paper are sourced from the Bank of England web site, Bloomberg and Consensus Economics. The data frequency is end of month, except where specified otherwise.

A summary of the sources of data is provided in Table 1.

2.1. Data transformations

The methodology adopted requires the use of continuously compounded rates of inflation. Transforming the annually compounded rates to continuous compounding is a simple task, but establishing consistency with regard to the time periods being measured is more challenging and in some cases requires judgement on the appropriate assumptions.

It is advantageous to use short-period forward rates (or even instantaneous forward rates) because these give an accurate representation of inflation at a point in time in the future. We will use forward rates with a tenor of one month. The BEI data from gilts and inflation swaps have increments to term to maturity of one year or longer. Curve fitting, or interpolation, is used to produce one-month forward rates. Producing instantaneous forward rates requires a more complicated process and in this case would not be particularly beneficial because the differences between the instantaneous and one-month forward rates will be very small.3

This section describes the transformations applied to the various data sets to produce consistent (and therefore comparable) continuously compounded, one-month forward rates.

2.1.1. Consensus forecasts

The data set from Consensus Economics is quite distinct from the others. Instead of providing ordinary spot or forward rates, Consensus Economics provides a forecast for the current calendar year at the time the forecast is made and forecasts for the ten calendar years that follow. Thus the surveys conducted and published in April measure a different term to maturity from the October surveys and each requires an appropriate assumption in order to produce spot rates (which will later be transformed into forward rates).

The surveys are usually conducted mid-month. Thus the consensus forecasts published in the April months begin with a forecast figure that refers to 3.5 months of realised inflation and 8.5 months of forecasted inflation. In order to use this data we need to assume that this annual rate represents only the average inflation expected over the 8.5 months remaining in the current calendar year.3

Surveys not conducted in April are conducted in mid-October. At this point in the calendar year there are only 2.5 months remaining. Hence, the first figure in the series of eleven forecasts (for the year of the forecast and the following ten years) represents mainly RPI inflation that has already been realised. It is appropriate to discard this figure and extend the coverage of the forecast for the following year to cover the 2.5 months from the forecast date. The data structure of the consensus forecasts is explained in Figure 2.

The spot rate curve is built by converting all rates to continuous compounding (see section 2.1.2) and adding these rates together for consecutive periods from time zero to the spot rate’s endpoint. These aggregated rates are then divided by the length of the forecast period

---

1Zero coupon, i.e., the swap rates have been transformed to correspond to fictitious swaps consisting of only one cash flow and therefore are equivalent to BEI spot rates, although differ from the gilt BEI spot rates due to the compounding.

2The consensus inflation forecast data, provided by Consensus Economics, have a different structure from the other two data sets mentioned above. The data records survey results of economists’ inflation forecasts each April and October starting October 1989. At each survey date an inflation forecast for that calendar year is recorded, as well as for the following ten calendar years. The figures are annually compounded rates. Data for April 1992 are missing for reason unknown. The missing data were smoothed over using an adjustment to the cubic spline interpolation process (see section 2.1.3) that accounted for an absence of data between October 1991 and October 1992.

3This was verified by comparing the instantaneous forward rates provided by the Bank of England with derived one-month forward rates from the spot rates. They are equal to three significant figures.

4This means that inflation forecasts made in April months will imply inflation expectations in the first year that are closer to the average realised inflation in the first 3.5 months of the calendar year than the consensus of the surveyed economists at the time.
Table 1: Sources of data.

<table>
<thead>
<tr>
<th>Data</th>
<th>Type</th>
<th>Sample</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gilt data</td>
<td>Spot rates and instantaneous forward rates (both continuously compounded)</td>
<td>January 1985 to May 2009</td>
<td>Bank of England</td>
</tr>
<tr>
<td>Inflation swap data</td>
<td>Swap rates on zero coupon inflation swaps with annual compounding</td>
<td>April 2004 to May 2009</td>
<td>Bloomberg</td>
</tr>
<tr>
<td>Consensus forecast data</td>
<td>Economists’ inflation forecasts for ten calendar years starting at year of forecast (annual compounding)</td>
<td>October 1989 to April 2009</td>
<td>Consensus Economics</td>
</tr>
<tr>
<td>Realised RPI data</td>
<td>Rolling annual changes with annual compounding</td>
<td>June 1948 to May 2009</td>
<td>UK Office for National Statistics</td>
</tr>
</tbody>
</table>

(in years) to produce continuously compounded rates of annual length.

2.1.2. Converting to continuous compounding

As mentioned above, all inflation figures used in this paper need to be continuously compounded rates. However, with the exception of the Bank of England data, all data are annually compounded. These are converted to continuously compounded rates using the equation,

$$\tilde{\pi}_{t-m,t} = \ln \left( 1 + \tilde{\pi}_{t-m,t}^{\text{ann}} \right),$$

where $\tilde{\pi}_{t-m,t}$ is the continuously compounded break-even inflation spot rate, calculated at time $t-m$ and ending at time $t$, and $\tilde{\pi}_{t-m,t}^{\text{ann}}$ is the equivalent rate with annual compounding. The symbol $\pi$ stands for inflation and the $\tilde{}$ above indicates that the figure is break-even inflation rather than realised inflation. Equation 2 is applied to the annually compounded, inflation swap spot rates, consensus forecasts and realised RPI inflation figures.

2.1.3. Cubic spline interpolation

The BEI and consensus data (which constitute all of the data except the realised RPI inflation figures) have increments in their terms to maturity that are longer than one month. The inflation swap BEI data have intervals of up to five years between consecutive lengths of the term; the gilt BEI data have six months between consecutive terms to maturity; and the consensus forecast data is for annual calendar-year forecasts. Interpolation of the spot curves is therefore necessary to produce spot curves that have term structures with one-month increments. It is preferable to interpolate the spot curves and then convert to forward curves rather than the other way round because of the inherent smoothness of the spot curves.

For each of these data sets and for any one starting date, the spot curve is clearly not linear; therefore using linear interpolation to diminish the intervals to one-year length is inappropriate. Instead, cubic spline interpolation is implemented with the so-called natural boundary conditions.5

This choice of boundary conditions is appropriate because setting the second derivative of the spline at the endpoints to zero produces the smoothest resulting curve and is consistent with the expected shape of an inflation spot curve at the extremities.

The cubic spline interpolation is conducted with the particular specifications to match each data set. The differences across the data sets are the sizes of the increments in the term structure (which in the case of the inflation swap data are not uniform) and the starting and ending points of the term structure.

2.1.4. Bootstrapping to create the forward curve

The same data sets to which cubic spline interpolation was applied in section 2.1.3 need to be transformed into forward rates. Bootstrapping enables us to deduce the implied forward rates of inflation between term dates (which have one-month intervals owing to the cubic spline interpolation) by comparing the spot rates before and after the period to which each forward rate applies and extracting the difference appropriately.

3 Implemented via a loop using the `spline` function in `Matlab`:

$$y(i,:) = \text{spline}(\text{domain}_{\text{data}}, \text{data}(i,:), \text{domain}_{\text{spline}}).$$
Equation 3 is an adapted version of the standard bootstrapping procedure. As with the cubic spline interpolation, the bootstrapping procedure is applied with particular specifications that suit each data set. The equation,

\[
(\ell)\pi_{t-\frac{\ell}{12}} = (\ell - \frac{1}{12}) \times \pi_{t-\frac{\ell-1}{12}} \times 12 \\
\text{for } \ell \in \{a, a+\frac{1}{12}, a+\frac{2}{12}, \ldots, b-\frac{1}{12}, b\}, \quad (3)
\]

is applied across each time point for \(\ell\) increasing in increments of one month from the starting point of the term structure (labelled \(a\)) to the endpoint (labelled \(b\)). This is repeated for each time point and again for each data set, selecting \(a\) and \(b\) to match the data. The values of \(a\) and \(b\) for each data set are shown in Table 2:

<table>
<thead>
<tr>
<th>Data set</th>
<th>Beginning ((a))</th>
<th>End ((b))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gilt data</td>
<td>4</td>
<td>20&lt;sup&gt;7&lt;/sup&gt;</td>
</tr>
<tr>
<td>Inflation swap data</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>Consensus forecast data</td>
<td>1</td>
<td>10</td>
</tr>
</tbody>
</table>

Subscript \(\ell\) represents the length of time between the measurement date and the end of the forecast. Equation 3 derives one-month forward rates by making use of the convenient additive property of continuously compounded rates. The part of the equation in parenthesis is multiplied by 12 to scale it from a monthly rate to the customary annual length, matching the realised RPI percentages (see section 2.1.5) and the forecasts from the model in section 2.4.

### 2.1.5. Filtering out seasonality

There is significant autocorrelation in the differences at lag 12 months. This indicates that the month-on-month changes in the level of the index are significantly affected by the calendar month in which the change is being recorded, i.e., seasonality. The tool used to remove the seasonality effect is the US Census Bureau’s X-12-ARIMA Seasonal Adjustment Program.

The same autocorrelation test is then conducted on the first differences of the seasonally adjusted RPI series. We find that the autocorrelation at lag 12 months has been removed.

Month-on-month continuously compounded inflation rates are derived from the seasonally adjusted RPI series using the following equation,

\[
\pi_{t-\frac{1}{12}} = \left[\ln(RPI^a_t) - \ln(RPI^a_{t-\frac{1}{12}})\right] \times 12, \quad (4)
\]

where \(\pi_{t-\frac{1}{12}}\) is a continuously compounded annual rate that represents the percentage change in retail prices (according to the RPI) at time \(t\) from one month prior.
The terms RPI<sub>π</sub><sup>SA</sup> and RPI<sub>π</sub><sup>SA</sup> represent the seasonally adjusted RPI levels in months \( t - \frac{1}{12} \) and \( t \) respectively.

At this point the notation is unnecessarily cumbersome because the interpolation and bootstrapping processes have converted all rates to correspond to month-on-month changes. Hereafter, notation used to represent measures of inflation (whether forward rates, realised rates or model expectations) from \( t - \frac{1}{12} \) to \( t \), such as the term on the left-hand-side of equation 4, are written without the \( t - \frac{1}{12} \) subscript. In other words, \( \pi_t \) is used as shorthand for \( \pi_{t-\frac{1}{12}} \).

### 2.2. Testing the inflation expectations hypothesis

This test investigates whether the BEI forecasts from gilts accurately predict future inflation. The regression equation,

\[
\pi_{t+\ell} = \alpha + \beta(\pi_t) + \epsilon_t
\]

is applied to the gilt data, where \( \pi_{t+\ell} \) is the RPI inflation rate (as defined above) at time \( t + \ell \); \( \pi_t \) and \( \pi_{t+\ell} \) are respectively the nominal and real forward rates and \( \pi_t \) is break-even inflation rate corresponding to the month ending at time \( t + \ell \) (consistent with \( \pi_{t+\ell} \)); \( \alpha \) and \( \beta \) are constants; and \( \epsilon_t \) is the error term. The purpose of the test is not to compare the quality of forecasts, but instead to ascertain whether there is an inflation risk premium that, being stationary or otherwise, is distorting BEI forecasts.

Equation 5 is not applied to the inflation swap data because this data set begins at April 2004 so even using only one-year BEI forecasts would not provide sufficient data points for analysis. Consensus forecasts are also not considered because the forecasts are not derived from tradable financial instruments and hence do not contain true risk premia.

Even with the gilt data alone the test is not straightforward. The twenty-year gilt BEI figures only go back to April 1992 and so the first forecasts can only be compared to realised RPI inflation in the year 2012. BEI forecasts from gilts with terms to maturity of 16 years are available from January 1985 (the start of the data set as a whole). This allows a regression for \( t \) ranging from January 2001 to May 2009 (101 data points). The shortest BEI forecast that is available from gilts dating back to 1985 is four years. Meaningful regression coefficients can thus be estimated for forecast lengths from four to sixteen years. There is a visible structural break in the RPI data in 1992. In this year the UK government abandoned the European Exchange Rate Mechanism and the Bank of England adopted a policy of inflation targeting (Shen, 1998). However, choosing to begin the hypothesis testing at 1992 for the RPI data will not have a significant impact on the results because, with even the shortest gilt BEI forecast of four years, the RPI data is only used from 1989 onwards (1985 plus four years), so the difference is just three years.

### 2.3. Measuring the inflation risk premium

The main aims of this paper are to measure the information in market prices about future inflation and to analyse the inflation risk premia (IRP) that are present in gilts and inflation swaps at different terms to maturity. This is achieved by comparing BEI forecasts with econometric model expectations as shown in the equation,

\[
\omega\tilde{\pi}_t = E_[\pi_t] + \omega IRP + \epsilon_t,
\]

where \( \omega\tilde{\pi}_t \) represents a BEI forecast for time \( t \) using market prices from \( \ell \) years prior; \( E_[\pi_t] \) is the econometric forecast of inflation at time \( t \) using the available data \( \ell \) years prior; \( \omega IRP \) is the inflation risk premium for the type of instrument (gilts or inflation swaps) at time \( t \) for forecast length \( \ell \) years; and \( \epsilon_t \) is the error term. The error term indicates that there will be some measurement error in the econometric forecasts and so the difference between the BEI figures and the econometric forecasts cannot solely be attributed to the inflation risk premium.

The advantage of this approach is that it does not rely on the realised inflation figures that are being forecast. It measures market sentiment across time and forecast length instead of the market’s accuracy in forecasting. However, it is not without fault. The inflation risk premium includes a convexity factor, which in the case of gilts, results from a difference in the convexity premia on nominal and real gilts. The convexity factor is also present in yields on inflation swaps. The approach taken in this paper does not attempt to decompose what is labelled the inflation risk premium and remove the convexity factor; neither is the measurement error taken into account. These two factors are not expected to distort significantly the inflation risk premia, and furthermore, the econometric model is tested for robustness in section 2.5.

The timespan for the application of equation 6 is limited to January 1997 to May 2009, primarily to allow the econometric model a long enough input series to produce meaningful forecasts. Forecasts made using data available at the end of the month of January 1997 use an
input series of 60 data points. There is also a structural break in the gilt BEI data in 1997 – the year in which the Bank of England gained its operation independence from the Government.

2.4. Econometric model of inflation

Theoretically an autoregressive (AR) model should fit the data well because of the strong autocorrelation in the RPI’s month-on-month changes. Regression tests, analysing residuals for autocorrelation or heteroskedasticity, were conducted and we found that it is not necessary to include Moving Average (MA) terms. The inclusion of MA terms in the econometric model is therefore not adopted.

The econometric model uses a purely autoregressive structure with specifications according to the equation,

\[ \pi_t = \alpha + \beta_1 \pi_{t-1} + \beta_2 \pi_{t-2} + \cdots + \beta_p \pi_{t-p} + \epsilon_t, \] (7)

where the number of lags, \( p \), or order of the equation, is determined using using a decision process based on the Schwarz Information Criterion (SIC).

The calculation of the model coefficients and SIC values for each monthly time point \( t \), from January 1997 to May 2009, requires an input data series, specifically RPI data. The structural break in the RPI data mentioned above should not be ignored because the econometric model would otherwise make use of inappropriate RPI data dating back to 1947. Inflation, from 1947 until inflation targeting began in 1992, was significantly higher than after 1992. The vast majority of the BEI forecast data available correspond to (one-month) periods after 1992. Therefore the econometric model uses only data from January 1992 onwards.

2.5. Comparing the inflation forecasts

It is not possible to interpret meaningful forecast accuracy measures on the inflation swap BEI forecasts because there are only 62 observations. Thus the inflation swap data set is too immature to be included in the robustness tests.

We use a number of different statistics to understand the relative ability of different methods to forecast inflation. The mean forecast error (MFE) is the simplest estimator of the forecast error. The MFE estimator is biased because the mean of the forecasts (predicted) is not equal to the mean of the realised RPI inflation figures (actual).

The mean squared error (MSE) measures the amount by which the realised quantity (RPI month-on-month inflation) differs from each type of forecast. The mean squared error consists of the variance of the mean forecast error (MFE) and the square of the MFE. Thus the mean squared error is a weighted average of the variance and bias of the MFE, where the weights are one and bias of the MFE respectively. Since inflation forecasts and realised rates are less than one (100%), the bias component of the mean squared error will have a very small weight.

The MFE and MSE measures are only meaningful when considered relative to other consistent values. This comparison is somewhat difficult to make because both the MFE and MSE values are less than one-hundredth in magnitude and moreover, the MFE values on the different forecast types can be mixed between positive and negative. Therefore we introduce scaled ranks defined by:

\[ R_x = \frac{\text{Abs}(x)}{\text{Min} [\text{Abs}(x), \text{Abs}(y), \text{Abs}(z)]}. \] (8)

where \( R_x \) is the scaled rank for forecast type \( x \); \( \text{Abs}(y) \) is the absolute value of the MSE or MFE for forecast type \( y \); and \( \text{Min}[\cdots] \) is the minimum of the entries within the brackets.

Tracking signals can add a new dimension to the analysis of forecast quality. The idea behind the tracking signal is that it can be used to identify models or forecasting techniques that are suffering from bias and need adjustment (see Alstrom and Madsen, 1996).

Tracking signals that drift upward show a bias to underestimate future inflation (while the curve continues to rise). Tracking signals that drift downward show a bias to overestimate future inflation.

3. Results

We produce four sets of inflation forecasts:

- Gilt BEI forecasts
- Inflation swap BEI forecasts
- Consensus forecasts
- Econometric model forecasts

The data and methodology provide consistent forecasts from January 1997 to May 2009, except the BEI inflation swap forecasts, which begin in April 2004, owing to the immaturity of that market. Although the measures are consistent, the boundaries to the term structures are different. Inflation swaps are the only
source of market data that provide the full one- to thirty-year term structure that is investigated.

Figure 3 shows the forecasts from all four sources on the four most recent survey dates for the consensus forecasts.9 These graphs are selected because they are fairly recent and show how significantly the forecasts can change over six-month periods.

The BEI forecasts derived from gilts and inflation swaps are similar, particularly near the boundaries of the term structure rather than in the 10- to 20-year range. Another clear observation is the change in the shape of the term structure over the 18-month period from relatively flat to upward-sloping. This is not just due to raised long-term inflation forecasts, but also a fall in short-term expectations (mainly visible in the inflation swap data).

Besides the term structure up to five years in the October 2008 graph, the consensus forecasts are always below the BEI forecasts. This outcome is expected because the consensus forecasts are not derived from market price data and so do not contain inflation risk premia. In October 2008, price levels were just beginning to fall and so the short-end of the consensus forecasts slopes downwards. By April 2009 retail prices had fallen significantly and economists were predicting a return to normal levels (which are 1 to 4% historically, with a mean of 2.55% for this sample), hence the upward-sloping short-end. However, in general the term structure of the consensus forecasts is quite flat because of the absence of risk premia.

The final group of forecasts are the econometric model expectations. The econometric model expectations from one to eleven months are included in Figure 3 to show that their term structure is not a straight line throughout — the forecasts simply converge very quickly (usually within ten to thirteen months). Like the consensus forecasts, the model expectations contain no inflation risk premia, which explains their lack of shape. It is exactly this lack of curvature that makes the econometric expectations suitable for subtraction from the BEI forecasts to produce inflation risk premia, with the term structure produced by the market participants still intact.

3.1. The inflation expectations hypothesis

Bulkley, Harris and Nawosah (2008) describe the expectations hypothesis of the term structure of interest rates as being that long-term interest rates are comprised of a weighted average of short-term forward rates, plus a constant (inflation and other) risk premium. The inflation expectations hypothesis is built upon this idea. The

---

9 The figures are adjusted for the half-month difference in timing because the Consensus Economics surveys are conducted mid-month.

Figure 3: Inflation forecasts. Inflation forecasts on the months of the four most recent consensus forecast surveys. Forecast lengths range from one month for the econometric model to thirty years for BEI forecasts from inflation swaps.
Table 3: F-tests for the inflation expectations hypothesis. The unrestricted model in the F-test is a simple linear regression model and the restricted model fixes $\alpha = 0$ and $\beta = 1$. The null hypothesis states that the realised inflation rates are accurately predicted by the unadjusted gilt BEI forecasts. $n$ represents the number of observations in the sample; $\alpha$ and $\beta$ are the regression coefficients from the unrestricted model (* indicates statistical significance at the 5% level); RSS stands for residual sum of squares, which is presented for the unrestricted and restricted models; F-statistic is the test statistic from the F-test; and p-value is the probability of obtaining an F-statistic at least as large as the one associated with the p-value in the table, if the null hypothesis is true.

<table>
<thead>
<tr>
<th>Forecast length</th>
<th>5</th>
<th>10</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n$</td>
<td>233</td>
<td>173</td>
<td>113</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>0.0104</td>
<td>0.0025</td>
<td>0.0069</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.4254</td>
<td>0.5035</td>
<td>0.4724</td>
</tr>
<tr>
<td>RSS — Unrestricted</td>
<td>0.24092</td>
<td>0.13737</td>
<td>0.11819</td>
</tr>
<tr>
<td>RSS — Restricted</td>
<td>0.31679</td>
<td>0.21533</td>
<td>0.14084</td>
</tr>
<tr>
<td>F-statistic</td>
<td>36.34</td>
<td>48.52</td>
<td>10.64</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

hypothesis states that, like interest rates, inflation can be predicted using only gilt market data.

We do not expect such a hypothesis to hold true for a typical sample period, because the inflation risk premium is likely to be significant in size and variable throughout time (non-stationary). In this section and the next we test these theories.

The inflation expectations hypothesis can be tested for forecast lengths from four to sixteen years. We select three (5, 10 and 15 years) and show the results in Table 3.

The sample sizes decrease by 60 observations from column to column, because with five-year longer forecasts in each successive column the sample period for realised RPI data must end sooner (by 5x12 months). $\alpha$ and $\beta$ are both positive throughout, as expected, because inflation is expected to be positive and the gilt BEI forecasts should be positively related to the realised inflation rates. The F-statistics and associated p-values for all three terms are statistically significant at the 1% level. This means that the gilt BEI forecasts are significantly poorer forecasters of future inflation than rates derived from affine transformations of the BEI figures. The most probable cause is an inflation risk premium of significant magnitude.

The lower level of significance in the 15-year column, including the only $\beta$ coefficient that is not significant at the 5% level, does not imply a smaller inflation risk premium on 15-year forecasts, or that long-term forecasts are superior to shorter-term ones. Instead, it is owing to the diminished power of the test as a result of the sample size being the smallest in this column.

3.2. Inflation risk premia

Inflation risk premia are calculated for gilts and inflation swaps for the same period over which their BEI forecasts are derived. Figure 4(a) shows the inflation risk premium on gilts. The term structure is visibly flat, not because it is typically so (it is seldom the case), but rather because the term structure of the inflation risk premium on gilts varies widely between upward- and downward-sloping, both concave and convex to the origin. A look ahead to Figure 6 confirms this.

The graph does not include the average of inflation risk premia on gilt BEI forecasts with lengths 20 to 25 years. This data begins only in May 2000 and the inflation risk premium increased significantly over the past decade, especially at the long-end. Therefore, including the data derived from BEI forecasts beyond 20 years would create an unnatural upward kink in Figure 4(a).

Figure 4(b) shows the inflation risk premium curve for all available inflation swap data (April 2004 to May 2009). For the sake of comparison, the graph includes the curve of the average inflation risk premium on gilt BEI forecasts over the same period. The two curves are expected to be correlated, but not necessarily as similar as shown. It would be quite normal for the two curves to differ significantly due to an aggregate effect of small differences each month than are predominantly in one direction.

Figure 4 provides a snapshot of the inflation risk premia on gilts and inflation swaps. The average (mean) inflation risk premium on gilts for the full sample period (Figure 4(a)) is 0.35%, or 35 basis points (bp). It exactly doubles to 70bp when considering only the sample period of the inflation swap data (Figure 4(b)). The average inflation risk premium on inflation swaps is inter-
estingly also 70bp when constraining its term structure to match that of the gilts; otherwise it is 63bp, where the difference is mainly owing to the short-end of the inflation swap data being the lowest part of the curve.

3.2.1. Variation in inflation risk premia

The two main ways to divide the inflation risk premium data are by forecast length and by time period. This section discusses the variation over time of the inflation risk premium at specific forecast lengths, whereas section 3.2.2 analyses the changes over time in the term structure of the inflation risk premium.

Figure 5 shows seven graphs of the inflation risk premia on gilts and inflation swaps for the sample period, January 1997 to May 2009, where each graph represents a different forecast length. Gilt data are not available for one- and thirty-year BEI forecasts, and the twenty-five-year gilt BEI forecasts begin only in May 2000.

The closeness of the gilt and inflation swap BEI forecasts is evident in each graph where both are present. Stationarity tests are conducted in their own right and not simply as a prerequisite for cointegration testing, although the tests do not include the inflation swap data because the 62 available observations for each graph are insufficient. We tested for stationarity and in each case the null hypothesis is that the series tested is non-stationary.

Only the Augmented Dickey-Fuller test statistic for the five-year gilt BEI forecasts is significant at the 5% level. The others are found to be non-stationary. The column labelled All shows the test results for the series generated from the average of the inflation risk premium curve at every point in the sample from January 1997 to May 2009. Gilt BEI data for forecast lengths 21 to 25 years are not included because the data begin mid-sample (in May 2000) and would induce artificial integration owing to the higher long-end inflation risk premia after this date.

Cointegration of the gilt data with the inflation swap data is found at the 5% significance level (results not shown) for the 10-, 15-, 20- and 25-year BEI forecasts. Testing at the other forecast lengths could not be conducted owing to either an absence of gilt data at the short-end and long-end, or the absence of an integrated series (at forecast length five years).

3.2.2. Size and term structure

The previous section analysed the variation in the time series of the BEI forecasts at particular forecast lengths. This section analyses the average level of inflation risk premia during each calendar year from 1997 through 2009 as well as the changes in the shape of the term structure of the risk premia during that time.

Table 4 shows, for each calendar year, average (mean) inflation risk premia on gilts and inflation swaps and a comparison between the two.

The results in Table 4 are associated with the graphs in Figure 6. Both the table and the figure show that the gilt BEI forecasts fell significantly from 1997 to 1998. For the next five years (to 2003) the term structure of the BEI forecasts changed to convex in 1999 and 2000, slightly concave in 2001 and 2002 and then quite flat in 2003, although the average magnitude of the inflation risk premium remained within a range only 21bp wide that included zero. After 2003 the average inflation risk premium increased steadily, along with the slope of the curve to May 2009.
The expectations hypothesis of inflation (results in Table 3) finds it likely that there is a significant inflation risk premium on gilt BEI forecasts. This result is not obvious when looking at the graphs for 1997 to 2003 in Figure 6; however, from 2004 onwards the inflation risk premia on gilts and inflation swaps are clearly significantly different from zero and positive for the vast majority of the term structure curve.

The bottom half of Table 4 shows that the correlation between gilt and inflation swap BEI forecasts is generally high. There is small negative correlation between the two curves in 2004 and 2009, but these years form the beginning and end of the data set and are incomplete. With the exception of the 2007 calendar year, the paired t-tests reject (at the 5% significance level) the null hypothesis that the gilt and inflation swap data are from the same generating process. This might come as a surprise given that they look so similar in Figure 6 and the correlation coefficients are so high. The p-values are below 1% for 2004 to 2006, because the inflation swap curves sit above the gilt curves for almost every point on the term structure during these years and the margin varies little. Hence the average difference between the two curves in each graph is large (in absolute value) and the standard deviation of the differences is very small, causing a large test statistic and the null hypothesis to be rejected. The paired t-test fails to reject the null hypothesis at the 1% level for the 2008 calendar, even though the gap between the two curves appears to be larger than in previous years, because neither curve dominates the other throughout the domain of the term structure and

Figure 5: Inflation risk premia at different forecast lengths. Risk premia on gilts and inflation swaps for the period January 1997 to May 2009. The forecast length increases in consecutive charts from one year to thirty years.
Figure 6: Inflation risk premia for calendar years 1997 to 2009. Risk premia on gilts and inflation swaps are shown for forecast lengths of one to thirty years. *The 2009 data ends May 2009 so do not represent a full calendar year. The gilt BEI forecasts of lengths 21 to 25 years begin May 2000 causing a minor kink in the graph for that calendar year.
the standard deviation of the differences is quite large relative to the mean of the differences.

3.3. Predicting inflation

We consider inflation forecasts from four sources: gilts, inflation swaps, Consensus Economics surveys and an econometric model. We have forecasts for each of these from January 1997, except the inflation swap data, which begins on April 2004. Therefore we cannot test the forecast accuracy of estimates derived from inflation swaps, since there are only 62 observations and, even if we consider only a one-year forecast, there will be 12 months of untestable data at the end of the sample. For five-year forecasts we would have only two observations to consider.

Figure 7 shows the four- to eight-year forecasts from the other three sources, along with the realised inflation curve a number of years ahead equal to the length of the forecasts in each graph. Therefore the realised RPI series end before May 2009 by the length of the forecasts. The stark fall at the end of curve in each graph represents the fall in the RPI in the final quarter of 2008 and early 2009. For graphical purposes only, the realised RPI series reflects the annual percentage change in the retail prices index over the previous year (as RPI inflation is reported in the media). This smoothes the series so that it is visually comparable to the forecasts. Calculations are undertaken using one-month forward rates that are consistent with the forecasts.

This section makes no attempt to compare the quality of forecasts made in different time periods – the sample is too small – but instead compares the forecast accuracy between gilts BEI forecasts, consensus forecasts and econometric model expectations. The comparisons begin for forecast length four years, matching the gilts BEI forecasts, and end at eight years, because this leaves 53 remaining testable observations after accounting the lag due to the forecast length. Fewer observations would make testing less statistically conclusive.

Table 5 compares the forecast errors of the three types of forecasts. Panel A shows the mean squared error (MSE) and mean forecast error of the forecasts versus realised RPI inflation. Panel B shows the corresponding scaled ranks, where one is the lowest and best rank. It can be seen clearly when looking at the scaled ranks that overall the econometric forecasts fair the best amongst the three, more so on the mean forecast error measure than on the mean squared error measure. The difference is owing to econometric forecasts being relatively constant over time and lying roughly in the centre of realised RPI inflation so that positive and negative errors tend to cancel out. Gilt BEI forecasts, on the other hand, overestimate future inflation to some extent, but overall do the best job in forecasting the trends in future inflation – the four-year and seven-year forecasts in Figure 7 provide good examples. Note that the four-year forecasts from gilts do not perform well on the measures in the table due to generally overestimating inflation rather than due to poorly predicting the trend directions. The presence of an inflation risk premium may explain this, although the cause could be an excess demand for real versus nominal gilts with four years to maturity. The consensus forecasts are not particularly useful because they are dominated by the econometric forecasts in terms of both predicting the average level and predicting the trends of future inflation.

The graphs in Figure 7 and the results in Table 5 support the proposition that a fixed adjustment to the gilt BEI forecasts, to account for the inflation risk premium, will be more accurate than the consensus and econometric forecasts. For this to be feasible in practice the adjustment must be measurable at the time of making forecasts.

### Table 4: Inflation risk premia in each calendar year: 1997 to 2009.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gilts (mean)</td>
<td>90.8</td>
<td>-13.0</td>
<td>-16.7</td>
<td>-15.2</td>
<td>1.8</td>
<td>4.2</td>
<td>2.4</td>
<td>35.6</td>
<td>38.5</td>
<td>45.9</td>
<td>68.2</td>
<td>118.2</td>
<td>150.3</td>
</tr>
<tr>
<td>Inflation swaps (mean)</td>
<td>44.6</td>
<td>51.0</td>
<td>52.5</td>
<td>66.5</td>
<td>109.0</td>
<td>118.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pearson correlation coefficient</td>
<td>-0.16</td>
<td>0.96</td>
<td>0.69</td>
<td>0.96</td>
<td>0.95</td>
<td>-0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paired t-test — test-statistic</td>
<td>-3.50</td>
<td>-7.29</td>
<td>-4.44</td>
<td>1.19</td>
<td>2.35</td>
<td>2.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p-value</td>
<td>0.002</td>
<td>0.000</td>
<td>0.000</td>
<td>0.249</td>
<td>0.029</td>
<td>0.027</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4 shows the mean inflation risk premium for each calendar year, in basis points (hundredths of a percent). The bottom panel compares the forecasts from gilts and inflation swaps, using a correlation estimate and a paired t-test. The p-value is the probability of obtaining a test statistic in the t-test that is at least as large (in absolute value) as the one associated with the p-value in the table, if the null hypothesis is true. The null hypothesis states that the two samples are derived from the same generating process.
Table 5: **Forecast errors.** Compares the quality of gilt BEI, consensus and econometric forecasts for forecast lengths four to eight years. Panel A shows the calculated error statistics, panel B presents scaled ranks for these statistics and panel C shows an inflation risk premium adjustment assuming that gilt investors paid a fixed premium that minimises forecast errors.

**Panel A: Values**

<table>
<thead>
<tr>
<th>Forecast length:</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean Squared Error of:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gilt BEI forecasts</td>
<td>1.23E-03</td>
<td>1.21E-03</td>
<td>1.28E-03</td>
<td>1.41E-03</td>
<td>1.83E-03</td>
<td>1.39E-03</td>
</tr>
<tr>
<td>Consensus forecasts</td>
<td>1.17E-03</td>
<td>1.17E-03</td>
<td>1.27E-03</td>
<td>1.47E-03</td>
<td>1.80E-03</td>
<td>1.38E-03</td>
</tr>
<tr>
<td>Econometric forecasts</td>
<td>1.15E-03</td>
<td>1.14E-03</td>
<td>1.28E-03</td>
<td>1.46E-03</td>
<td>1.78E-03</td>
<td>1.36E-03</td>
</tr>
<tr>
<td><strong>Mean Forecast Error of:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gilt BEI forecasts</td>
<td>-5.10E-03</td>
<td>-2.52E-03</td>
<td>-2.29E-03</td>
<td>-2.17E-03</td>
<td>-3.42E-03</td>
<td>-3.10E-03</td>
</tr>
<tr>
<td>Consensus forecasts</td>
<td>5.84E-04</td>
<td>3.00E-03</td>
<td>2.96E-03</td>
<td>2.58E-03</td>
<td>6.27E-04</td>
<td>1.95E-03</td>
</tr>
<tr>
<td>Econometric forecasts</td>
<td>-7.79E-04</td>
<td>1.59E-03</td>
<td>1.20E-03</td>
<td>9.14E-04</td>
<td>-9.55E-04</td>
<td>3.95E-04</td>
</tr>
</tbody>
</table>

**Panel B: Scaled ranks**

<table>
<thead>
<tr>
<th>Forecast length:</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean Squared Error of:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gilt BEI forecasts</td>
<td>1.07</td>
<td>1.05</td>
<td>1.01</td>
<td><strong>1.00</strong></td>
<td>1.03</td>
<td>1.03</td>
</tr>
<tr>
<td>Consensus forecasts</td>
<td>1.02</td>
<td>1.03</td>
<td><strong>1.00</strong></td>
<td>1.05</td>
<td>1.01</td>
<td>1.02</td>
</tr>
<tr>
<td>Econometric forecasts</td>
<td><strong>1.00</strong></td>
<td><strong>1.00</strong></td>
<td>1.01</td>
<td>1.04</td>
<td><strong>1.00</strong></td>
<td>1.01</td>
</tr>
<tr>
<td><strong>Mean Forecast Error of:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gilt BEI forecasts</td>
<td>8.72</td>
<td>1.59</td>
<td>1.90</td>
<td>2.37</td>
<td>5.46</td>
<td>4.01</td>
</tr>
<tr>
<td>Consensus forecasts</td>
<td><strong>1.00</strong></td>
<td>1.89</td>
<td>2.46</td>
<td>2.82</td>
<td><strong>1.00</strong></td>
<td>1.83</td>
</tr>
<tr>
<td>Econometric forecasts</td>
<td>1.33</td>
<td><strong>1.00</strong></td>
<td><strong>1.00</strong></td>
<td><strong>1.00</strong></td>
<td>1.52</td>
<td>1.17</td>
</tr>
</tbody>
</table>

**Panel C: Inflation risk premium (IRP) adjustment**

<table>
<thead>
<tr>
<th>Forecast length:</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant IRP adjustment on gilts to minimise errors</td>
<td>0.51%</td>
<td>0.25%</td>
<td>0.23%</td>
<td>0.25%</td>
<td>0.34%</td>
<td>0.32%</td>
</tr>
</tbody>
</table>
Inflation forecasts are graphed along with the realised RPI inflation a number of years ahead equal to the forecast length. Period of investigation: January 1997 to May 2009.

If we use the results from the inflation risk premia calculations shown in section 3.2, we will just arrive at the econometric forecasts, because the inflation risk premia are calculated as the difference between the gilt BEI and econometric forecasts. Thus, instead Panel C shows a retrospective fixed adjustment to help our understanding of the forecasts rather than directly to improve them. All of the inflation risk premium adjustments are positive and the average of 32bp is consistent with the average of 35bp shown in Figure 4(a).

Figure 8 is included to give an example of the effects of the adjustment. The revised gilt BEI curve, shown in dark blue, is 51bp below the unadjusted gilt BEI curve. The improvement or reduction in errors is coloured green and the area reflecting the increase in forecast errors is coloured orange.

The total area coloured green and the total area coloured orange appear to be approximately equal; however, the revised gilt BEI curve is an improvement because it eliminates almost all of the mean forecast error and reduces the mean squared error by more than 2%. The inflation risk premium (retrospective) adjustments cause the gilt BEI forecasts to outperform the consensus and econometric forecasts based on the measurements in Table 5.

Figure 8: Four-year gilt BEI forecasts with constant inflation risk premium adjustment. The revised gilt BEI forecast curve is 51bp below the unadjusted curve. The area reflecting forecast error reduction is coloured green and the area reflecting increments to errors is coloured orange.

Predicting inflation in four or more years time is a difficult task and the difficulty increases with the forecast length. However, we have a considerable volume of useful market data that, as shown earlier in this
section, can be used to derive better forecasts than those provided by professional economists (consensus forecasts). Econometric and BEI forecasts are not consistently more accurate than consensus forecasts and we cannot be sure when it is appropriate to use a particular data source. Therefore, we need to understand the output from the models and BEI data in order to put it to use effectively.

Figure 9 shows the tracking signals of the gilt BEI forecasts, consensus forecasts and econometric model expectations versus the realised RPI inflation data for forecast lengths four to eight years. The tracking signals are available for shorter periods as the forecast length increases, because the missing RPI inflation data is yet to be realised.

The upward drift in the consensus and econometric forecast curves indicates that realised inflation, while the curves ascend, is above the forecasts from these sources, i.e., consensus and econometric forecasts generally underestimate inflation. In the fourth quarter of 2008 inflation fell sharply and we saw negative month-on-month RPI changes in six out of seven months from October. All three sources were unable to forecast this four or more years in advance. Hence the tracking signals fall considerably during this final period in each graph of Figure 9.

The gilt BEI forecasts are shown to predict better the trends in future inflation. When comparing the tracking signals before the unusually large fall in inflation late in 2008, the consensus and econometric forecasts generally drift only upwards, whereas the gilt BEI tracking signals cross the zero line at least once in each graph before this point.

What kind of adjustment will suit each forecast type? This study does not attempt to answer this question, but the consensus and econometric forecasts are lacking the foresight of the market and the market’s expectations are clouded by a non-stationary inflation risk premium. Thus, a trade-off may prove useful.

4. Conclusions

The results show that the term structure of the inflation risk premium can take on a variety of shapes. Furthermore, the shape of the term structure changed considerably since the independence of the Bank of England.

The inflation swap BEI forecasts have an advantage over the gilt BEI forecasts because the forecasts begin at length one year and end at thirty years. Gilt BEI forecasts only run from four to twenty-five years. However, the inflation swap market is immature and we can only perform the same robustness checks on its forecasts that were applied to the gilt BEI forecasts when we have realised RPI inflation figures that reach into the second half of 2016 (for the eight-year forecast accuracy test).

It is evident that the BEI forecasts from gilts and inflation swaps are very close, and hence the inflation risk premia are also close. Where testing is possible, it is found that the two series are cointegrated so market movements in the prices of these instruments are consistent and significant supply and demand anomalies are uncommon. This makes sense because if the BEI forecast rates differ by more than a certain margin (that accounts for the different trading costs, risk discrepancies, marketability, etc.) then arbitrage trading will reduce the difference(s) to within the margin again.

The closeness between gilt and inflation swap BEI forecasts suggests that the results showing non-stationarity in the forecasts from gilts are also valid for the inflation swap BEI forecasts. Over the six calendar years of available data, the ratio of the average inflation swap inflation risk premium to the average inflation risk premium on gilts has decreased from five-fourths to about four-fifths. It is unlikely that this trend can continue much further, but it is also unclear where equilibrium between the two will lie — the inflation swap market is still young and we only have data while the inflation risk premium has been rising. In 2009, the average inflation risk premium is 150bp on gilts and 118bp on inflation swaps (almost five and three times more than in 2004, respectively). However, the term structure is now significantly upward sloping. For example, the 2009 five-year inflation risk premium on gilts is 44bp and then rises to 143bp on ten-year gilts and 190bp on fifteen-year gilts.

The findings relating to measuring inflation forecast accuracy are not only interesting but also useful for improving the quality of forecasting. Inflation in the medium and long term is difficult to forecast accurately, as indicated by the relatively poor results from all sources, but there is potential for improvement by combining the positive attributes of each type of forecast. In particular, good forecasts should have the forward-looking nature of gilt and inflation swap BEI forecasts, but still retain the ability to model likely short-term trends that are governed by our econometric model.

Overall, the econometric model produces the most accurate forecasts based on the mean squared error measure and even more convincingly when measured with the mean forecast error. This is because the econometric model describes well the short-term trends in inflation and is best (amongst the options in the paper) for pre-
Figure 9: Tracking signals for forecast lengths four to eight years. The tracking signals can be used to determine whether a particular model or source of forecast data consistently under- or over-estimates future inflation and therefore needs adjustment.

dicting the long-term average level of inflation. Moreover, only the econometric model forecasts share the mean reversion property of realised inflation rates.

However, although the forecasts from the econometric model are technically more accurate — because they fit the realised RPI inflation closest overall — it is clear that there is important information contained in the gilt BEI forecasts that should not be ignored. The tracking signals, and the mean forecast errors too, show that econometric model expectations and especially consensus forecasts of inflation underestimate inflation consistently. The market data (in the form of gilt BEI forecasts) have a significantly stronger ability to forecast movements in inflation (rate of change of the month-on-month changes in the retail prices index) in the medium and long term.

The evidence in this paper suggests that the best forecasts of inflation can be made using an approach that combines forward looking and historical methods, rather than one or the other. Including an additional affine structure (constant plus trend) to the BEI forecasts to account for the varying inflation risk premia could prove useful. However, parametrisation using an econometric model will simply produce that econometric model’s forecasts. Subtracting only a fraction of the inflation risk premium on gilt and inflation swap BEI forecasts would imply a weighted average of the BEI and econometric forecasts. This might be an effective approach although careful consideration is required to determine the right technique for calculating these weights. Owing to the non-stationarity of the inflation risk premium and the changes over time in the shape of its term structure, it is necessary to perform separate, yet consistent calculations for the weights for different forecast lengths.

References

Nominal and Real Bond Yields. Federal Reserve Bank of San Francisco, working paper 2008-34.


