New estimates of the UK real and nominal yield curves

By Nicola Anderson and John Sleath of the Bank's Monetary Instruments and Markets Division.

This article presents some new improved estimates of the UK yield curve, both nominal and real. It describes the rationale for changing the estimation techniques that we have previously used, in the light of our own experience and developments in the academic literature. The article also illustrates the use of data from the general collateral repo market to derive estimates of the nominal yield curve at short maturities.

Introduction

Nominal yield curves have been estimated in the Bank for more than 30 years. For the past five years, in common with many other central banks, we have used the estimation method proposed by Svensson (1994, 1995). This is a parametric method, with the entire curve described by a single set of parameters representing the long-run level of interest rates, the slope of the curve and humps in the curve. Previously we used an in-house non-parametric method described by Mastronikola (1991). And before that we used another parametric approach, with the parameters reflecting, among other things, segmentation in the market and the planning horizons of different investors.

Estimation of the real yield curve is a more recent innovation, made possible by the introduction of index-linked bonds in the United Kingdom in 1981. As these bonds are indexed only imperfectly to the price level, we have to use information from the nominal yield curve to extract the real risk-free rates of interest embodied in their prices. Until now we have been using an iterative technique developed by Deacon and Derry (1994), in which the real yield curve is described by a restricted version of Svensson's model.

As discussed by Breedon (1995), the Svensson method was preferred both to the earlier in-house method and the range of alternative options available at the time, on the basis of three key criteria. Specifically:

- the technique should aim to fit implied forward rates (rather than, for example, yields), since the final objective is to derive implied forward rates;
- it should give relatively smooth forward curves, rather than trying to fit every data point, since the aim is to supply a market expectation for monetary policy purposes, rather than a precise pricing of all bonds in the market; and
- it should allow as many economic restrictions as possible to be imposed.

For maturities of less than two years, estimates of both the real and nominal yield curves have not been thought to be reliable, and as a result have not been used by the Bank's Monetary Policy Committee, nor published in the *Inflation Report* or *Quarterly Bulletin*. This is partly because there are few gilts at the short end of the yield curve (ie with terms to maturity of two years or less), where expectations may be relatively precise and where the curve may be expected to have quite a lot of curvature. More recently, experience has led us to question whether the Svensson estimates, even at the longer maturities, are the best guide to monetary conditions in the United Kingdom.

The opportunity to shed new light on the performance of these models has arisen, partly through the relatively recent arrival of additional information from the gilt market (in the form of strips prices), and partly through the development of new techniques for estimating the yield curve. In the latter case, we find that a new model developed by Waggoner (1997) offers a number of improvements on the parametric methods currently used to estimate both the real and nominal yield curves. In addition, improvements in extracting the real yield curve from index-linked bond prices can be found using the non-iterative technique developed by Evans (1998).

The following two sections describe the problem of extracting information from the bond market and the choice of techniques currently available. We then examine some estimates of the Svensson nominal yield curve. In the light of these observations, we describe a number of criteria for comparing different methods of estimating the yield curve, and discuss how these relate to four different models. The final two sections present estimates of the yield curve using our preferred model, first extended to include general collateral (GC) repo data at the short end, and then applied to index-linked gilts.

Extracting information from the bond market

The most useful information that can be derived from the government bond market is implied forward interest rates. These are important in their own right as they reflect, albeit imperfectly, the market's expectations about the future path of interest rates. But they also provide the building-blocks for other types of information, including zero-coupon yields and the synthetic bond prices we create to derive credit spreads from corporate bonds.

Implied forward rates are the marginal rates of return that investors require in order to hold bonds of different maturities. Ideally, we would like to measure 'instantaneous' forward rates, which are related to the price of a bond as follows:

$$B(\tau) = \exp\left[-\int_0^{\tau} f(m)dm\right] \qquad \textbf{(1)}$$

where f is the forward rate, B is the price of a zero-coupon bond and τ is its maturity. Given these forward rates, it is straightforward to derive the implied forward rate of interest between any two dates in the future, and at any point in time.(1)

To measure the set of instantaneous forward rates directly from the market requires a set of observable zero-coupon bond prices across a continuum of maturities (the 'discount function'). In practice, however, as we can only observe the prices of coupon-bearing bonds,(2) the discount function is not directly observable. All we can do is to write the price of each observable bond as follows:

$$P(c,\tau) = \sum_{i=0}^{n-1} cB(\tau - i) + 100B(\tau) \qquad \textbf{(2)}$$

where τ denotes the maturity of the bond, c is the coupon payment made in each period, and n refers to the number of such payments outstanding. A more fundamental problem is that these bonds are issued across only a finite set of maturities. We therefore need a method of disentangling the discount function and 'filling in the gaps' to give a continuous curve.

Parametric versus spline-based methods

The simplest method is to define the forward rate curve, $f(m)$, as a function, $f(m,\beta)$, of a set of unknown parameters, β. This is the approach taken both by Nelson and Siegel (1987) and by Svensson (1994, 1995). In these models, the parameters are related to the long-run level of interest rates, the short rate, the slope of the yield curve and humps in the curve. Svensson's model can be regarded as an extended version of Nelson and Siegel's model, with an additional hump to help fit bond prices in the market. The precise specification of each of these models is described in the Appendix.

In each case, via equations **(1)** and **(2)**, the functional form can be used to derive a fitted value for each bond price, given the set of underlying parameters. The parameters are estimated to minimise an objective function that compares these fitted values with observations from the gilt market. A variety of objective functions are available to us; over N bonds, we choose to minimise:

$$X_P = \sum_{i=1}^{N} \left[\frac{P_i - \Pi_i(\beta)}{D_i}\right]^2 \qquad \textbf{(3)}$$

where P_i is the observed price of the ith bond, D_i is its modified duration, and $\Pi_i(\beta)$ is the fitted price. This is approximately equal to minimising the sum of squared yield residuals (although it is much quicker to calculate), and so implies roughly equal yield errors, irrespective of maturity.

Rather than specifying a single functional form to describe instantaneous forward rates, spline-based techniques fit a curve to the data that is composed of many segments, with constraints imposed to ensure that the overall curve is continuous and smooth. This is the principle advantage of spline-based techniques over parametric forms since, subject to the continuity constraints, individual segments can move almost independently of one another.

This is clearly illustrated in Charts 1a and 1b, which shows an example of a simple non-linear least squares regression to a set of arbitrary data points, using both the Svensson functional form and a cubic spline.(3) When a single data point is changed at the *long* end, the Svensson curve changes dramatically, particularly at the *short* end, whereas the spline moves only slightly to accommodate the new data, and only at the long end. Methods for fitting cubic splines to the data differ in a number of ways, including the objective function used. The effect that this has on the resulting yield curve estimates is discussed in later sections.

UK nominal interest rates estimated using Svensson

At the long end of the yield curve, the Svensson model is constrained to converge to a constant level. The rationale for this constraint is based on the assumption that forward rates reflect expectations about future short interest rates, or equivalently that the unbiased expectations hypothesis holds. Assuming that this is true, it seems implausible that agents will perceive a different path for the future short rate in 30 years time compared with, say, 25 years. So we should expect to see constant expectations and forward rates at the long end.

But how does this compare with data from the strips market? In theory, the observed strips' yields should provide a direct reading on the underlying term structure that the Svensson method is attempting to describe. Chart 2 compares the estimated yield curve with the yields on strips

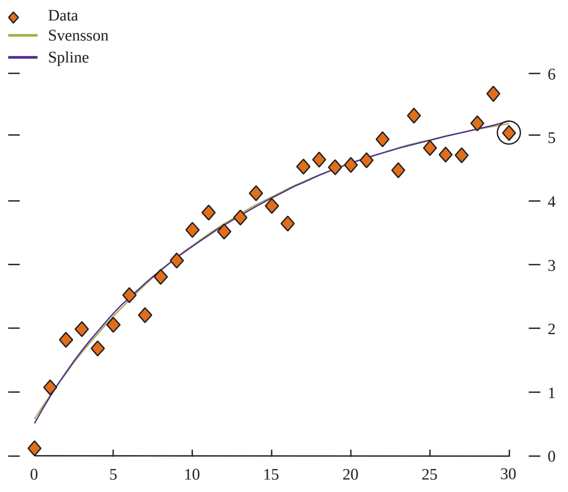
(1) The implied forward rate at time 0 between s and τ, for example, is given by $\int_s^{\tau} f(m)dm$.

(2) In fact, zero-coupon gilts have existed since the introduction of the strips market in December 1997. These separate the two components of a coupon-bearing gilt to give a principal strip with maturity equal to its redemption date and a series of coupon strips related to each payment date. The market in strips is, however, still small relative to coupon-bearing gilts. We therefore do not use strips prices to estimate the yield curve.

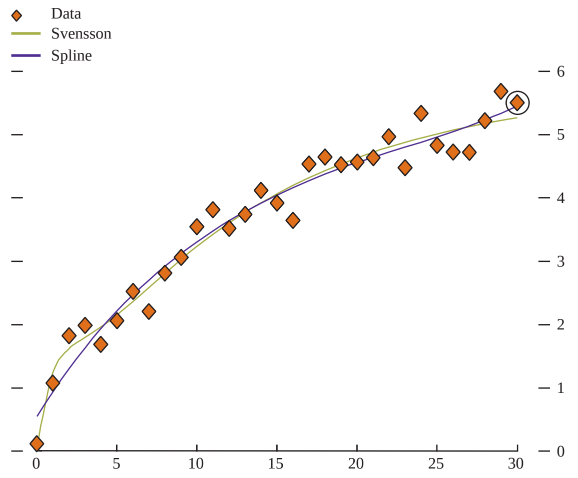
(3) The spline has been chosen to have the same number of degrees of freedom as the Svensson curve.

Chart 1 Svensson method versus cubic spline

(a) Original set of data points

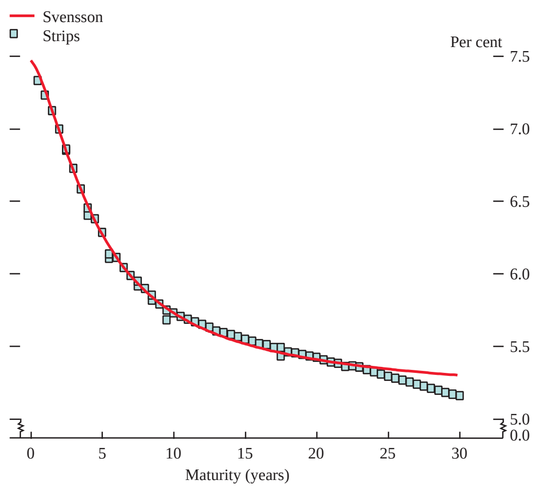


(b) Change of single data point



on a day chosen at random, 19 June 1998. The strips prices clearly display a downward-sloping term structure at the long end, compared with the constant level imposed by the Svensson yield curve. Assuming that expectations do

Chart 2
Estimated yield curve on 19 June 1998—Svensson versus strips



converge at longer maturities, this implies that there are other factors driving strips' prices (for example, risk premia and convexity terms), so that the unbiased expectations hypothesis does not hold.

Of course, the strips may be mispriced.(1) Direct evidence from the gilt market, however, suggests that a downward-sloping yield curve may be justified, at least over the maturity range that we consider. Moreover, forcing the long end of the curve to converge to a constant level can produce a significant amount of instability in the estimated yield curve. This is shown in Chart 3, where we plot the redemption yields on the 10-year benchmark bond and the longest-maturity bond (with maturity of 29 years), together with 20-year zero-coupon yield estimates derived using Svensson.

Chart 3
Time series of redemption yields at the long end

This illustrates that, as the *observed* bond yields have diverged more and more, the yield curve *estimates* have been increasingly unstable. We attribute this to the parameterised nature of the Svensson curve. Estimates at all maturities rely on a single set of parameters, of which one is the long-run level, determined largely by the yield on the longest bond. But the increasing divergence of the two redemption yield series suggests that the level of this asymptote is not well defined, at least in this maturity range. As a result, the asymptote itself is likely to be unstable, and this volatility will be transmitted into estimates of the yield curve as in Chart 1.

A comparison of techniques

In the light of this experience, we examined a number of alternative methods of yield curve estimation. In particular, we compared the performance of the Nelson and Siegel and Svensson methods with two spline-based models due to Fisher, Nychka and Zervos (1995) and Waggoner (1997). (See Appendix for details of these models.)

Our preferred model is a modification of the spline-based technique developed by Waggoner (1997), which he refers

(1) There is some concern about the reliability of strips prices in practice. This is because the market is relatively new (introduced in December 1997), and trading in strips is quite thin compared with conventional coupon gilts.

Summary of key criteria and properties of the VRP model

Criteria	Properties of the VRP model
Smoothness	
The technique should give relatively smooth forward curves, rather than trying to fit every data point, since the aim is to supply a market expectation for monetary policy purposes rather than a precise pricing of all bonds in the market.	Forward rates are estimated to maximise the fit of the model to observed bond prices while penalising curvature in the forward curve.
Flexibility	
The technique should be sufficiently flexible to capture movements in the underlying term structure. It should also be relatively less flexible at the long end than at shorter maturities, but should not necessarily asymptote within the range of maturities defined by the market.	The extent to which curvature in the forward curve is penalised—the value of the penalty parameter—depends on maturity; the shorter the maturity, the more structure is allowed in the curve. The penalty parameters are chosen to maximise the out-of-sample(1) goodness-of-fit of the model estimates.
Stability	
Estimates of the yield curve at any particular maturity should be stable, in the sense that small changes in data at one maturity (such as at the very long end) do not have a disproportionate effect on forward rates at other maturities.	Forward rates are described by a number of segments joined together. This in effect localises the influence of maturity idiosyncratic price movements to a specific portion of the curve.

(1) The term 'out-of-sample' here refers to the fit obtained for a bond excluded from the estimation. To estimate the overall out-of-sample goodness of fit we leave out each bond in turn, estimate the yield curve, and calculate the average fitting error of the omitted bonds, a procedure known as cross-validation (Davison and Hinkley (1997)).

to as the 'variable roughness penalty' (VRP) method. This model was chosen on the basis of a number of key criteria, and on its performance relative to the alternative models in a number of tests. For the sake of brevity, the results of these tests are not reported here.(1) Instead, we describe the intuition for our choice of yield curve model. The box above describes the main features of our preferred model, alongside our criteria.

In the Nelson and Siegel and Svensson methods, the yield curve estimates are guaranteed to be smooth by the parsimonious nature of the functional form—the curves are simply not flexible enough to capture the idiosyncratic price movements of every bond in the market. But this raises the question as to whether or not these methods are sufficiently flexible to capture movements in the underlying term structure. We conducted an out-of-sample test.(2) Each estimation method will produce a high in-sample goodness of fit, but this may not be indicative of the underlying term structure. The important test is whether the estimated curve can accurately price a bond that has not been used to estimate the curve.

Comparing results for the Nelson and Siegel and Svensson methods confirms Svensson's view that additional flexibility may be needed to capture variation in the underlying data. Both methods give qualitatively smooth forward curves, but the out-of-sample performance of the Nelson and Siegel method is inferior to the Svensson model.

So how do the spline-based methods compare? These techniques are specifically designed to be more flexible than the parametric forms. However, when fitting a cubic spline, we can control the smoothness of the curve by means of a roughness penalty. The objective function described in equation **(3)** is modified, so that we now minimise X_S, where:

$$X_S = X_P + \int_0^M \lambda_t(m)[f''(m)]^2\,dm \qquad \textbf{(4)}$$

$f''(m)$ is the second derivative of the fitted forward curve (and so is a measure of its curvature) and M is the maturity of the longest bond. The choice of roughness penalty, $\lambda_t(m)$, marks the main distinction between the two spline-based models we investigated. Fisher, Nychka and Zervos ('FNZ')(3) chose $\lambda_t(m)$ to be constant across all maturities, but variable from day to day.(4) In contrast, Waggoner (1997) allowed $\lambda_t(m)$ to vary across maturity, but kept it constant from day to day.

Waggoner chose a three-tiered step function for his smoothing parameter, with steps at one and ten years to maturity. This was based on the segmentation of the US market into bills, notes and bonds. The UK market cannot

(1) A forthcoming *Bank of England Working Paper* will discuss the results in full.
(2) See footnote (1) in the box above for a description of this test.
(3) Fisher, Nychka and Zervos (1995).
(4) The value of the smoothing parameter is chosen using a procedure known as generalised cross-validation. This attempts to find the optimum value based on the trade-off between goodness of fit and parsimony.

be naturally divided in the same way. We chose instead to define $\lambda(m)$ as a continuous function of only three parameters.[1] Following Waggoner, the main criterion for choosing these parameters was to maximise the out-of-sample goodness of fit averaged over our sample period.[2]

Intuitively, there are a number of reasons to suspect that the VRP method will provide us with more reliable estimates of the yield curve. First, by constraining the smoothing parameter to be maturity-invariant, FNZ assume that there is the same degree of curvature along the length of the term structure. But there are strong reasons to believe that this is not the case. In particular, investors are likely to be more informed about the precise path of interest rates at short and medium maturities (when interest rates are determined by monetary policy and business cycle conditions) than at longer maturities. Hence FNZ's curve may be too stiff at the short end (and so unable to capture the true shape of the underlying term structure) and/or too flexible at the long end (and so over-fit the data).

Comparing the goodness of fit of the two spline-based techniques supports these observations. In particular, the VRP curve outperforms the FNZ curve, which in fact does worse than both parametric forms. Intuitively, this is because it suffers from the same lack of flexibility at the short end as Nelson and Siegel. At the same time, the long end of the curve appears to be too flexible, fitting too closely to bonds included within the sample. Results for the VRP method, on the other hand, are very similar to those obtained with the Svensson model.

The main differences between the VRP and Svensson models relate to the stability criterion (see the box on previous page) and the constraints imposed at the long end. As mentioned above, the Svensson model is constrained to converge to a constant at long maturities, a property that appears to contradict evidence from the strips market. The VRP curve, on the other hand, is constrained only to be very smooth at these maturities. Chart 4 illustrates the effect that this difference has on the estimated yield curves.

Chart 4 shows that the spline-based curve is better able to capture the shape of the underlying term structure implied by strips, particularly at the long end. Note that data from the strips market were not used to derive these estimates.[3] The effect that this has on the stability of our estimates is shown in Chart 5. This compares the 20-year zero-coupon yields estimated using the new technique with those derived from the Svensson model, and shows clearly that the former are more stable.

More generally, the fact that the new model is non-parametric suggests that it is less likely to display the sort of instability highlighted above. To formalise this property and the effect that it has on the stability of the

Chart 4
Estimated yield curve on 19 June 1998—VRP model versus Svensson versus strips

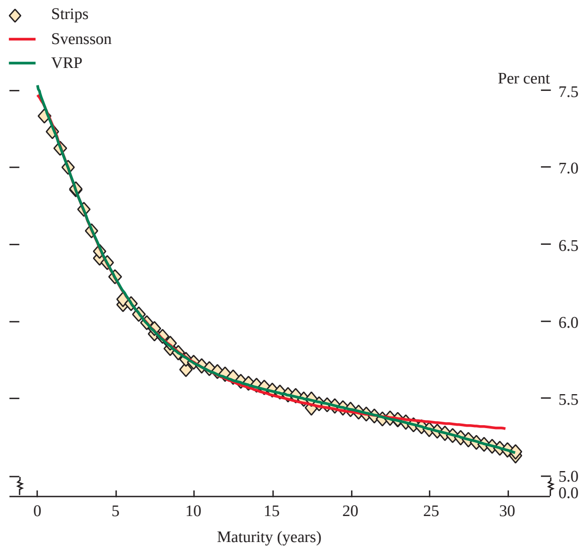
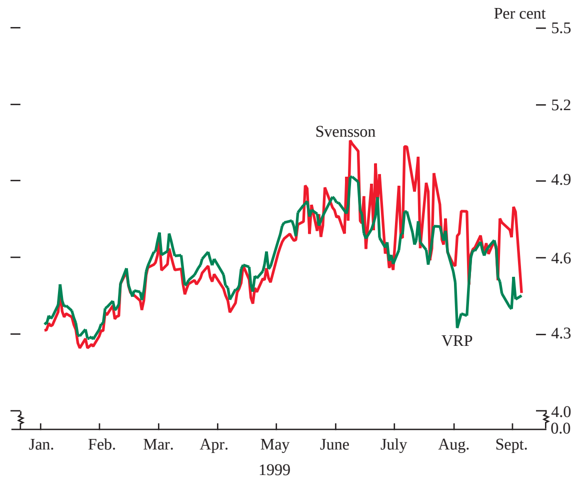


Chart 5
Comparison of 20-year yields since January 1999



resulting yield curve estimates, we conducted a stability test. All bond prices are subject to a measurement error, because of the finite minimum price change (the 'tick size'). So we require the estimated curve to be virtually unchanged if bond prices are perturbed by an amount smaller than a half of the tick size, and this forms the basis for the test. As expected, the two spline-based methods outperformed the parametric models in this test. The VRP method also outperformed the FNZ method, probably reflecting the fact that at longer maturities, the FNZ model is able to fit too closely to individual bonds.

The short end of the yield curve

At the short end of the yield curve, there are relatively few data from which estimates can be derived. An alternative approach is to introduce data from the money market. But

(1) We specify the following function: $\log \lambda(m) = L - (L - S)\exp(-m/\mu)$, where L, S and μ are the three parameters to be estimated.
(2) In practice, many combinations of these parameters gave similar out-of-sample goodness-of-fit measures. Within this set of combinations, we chose the set of parameters that corresponded to the highest level of smoothing.
(3) Market participants may use a similar yield curve to price non-trading strips from the gilts curve. If so, this reinforces our belief that the VRP curve captures the market's views.

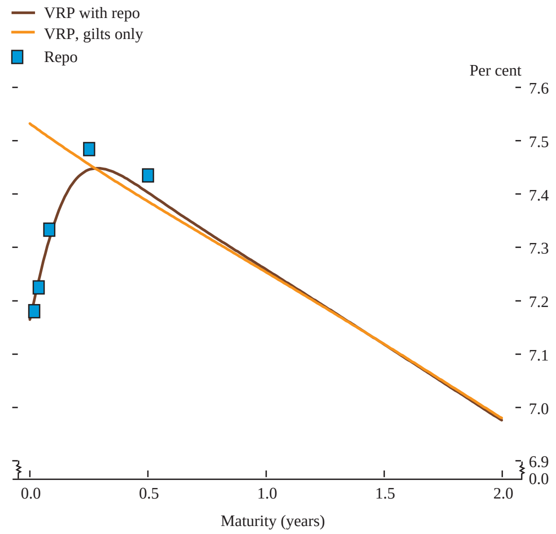
since we aim to measure the risk-free (or default-free) term structure of interest rates, the choice of data is limited. Although many short-term instruments are traded on the UK money market, their prices are not generally consistent with gilt prices, because they include a credit-risk premium. This leaves a choice of only two instruments: Treasury bills (T-bills) and GC repo rates.

T-bills are short-term zero-coupon bonds issued by the government, and so have the same risk-free characteristics as gilts. The outstanding stock of T-bills is, however, quite small, and because commercial banks use them for cash management purposes, their prices are widely accepted as being unrepresentative of the underlying fundamental rate determined by expectations.

A GC repo agreement is equivalent to a secured loan, and so the credit risk is much lower than on unsecured Libor. In addition, the repo is marked to market daily, thereby limiting the exposure of either party to large moves in the value of the collateral. The risk premium is further reduced because the collateral comprises gilts or similar instruments, for which there is virtually no chance that the issuer will default during the term of the repo. GC repo therefore provides us with the only widely traded, virtually risk-free short-term instrument.(1)

Chart 6 compares the yield curve estimates (based on the VRP method) with and without the inclusion of GC repo data, and the repo rates themselves. The difference between the two curves is striking. When the GC repo data are included, the curve exhibits a significantly different shape at the short end. At the same time, however, the two sets of estimates are virtually identical at maturities longer than one year. This is important as it indicates that, even if there is reason to doubt the reliability of the GC repo data or if these are not available, we can still have confidence in estimates at longer maturities.

Chart 6
VRP model: short-end yield estimates with and without repo rates (19 June 1998)



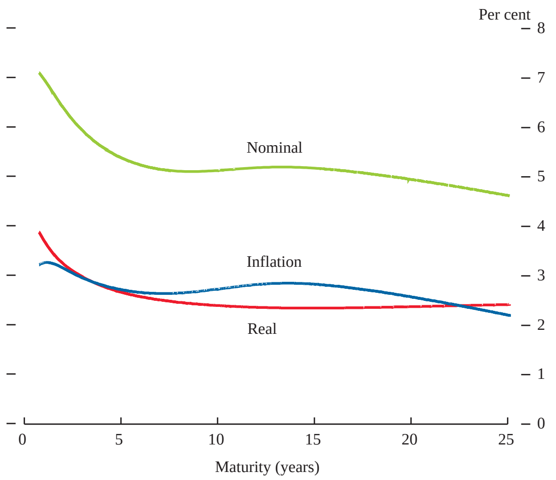
Estimation of the real term structure

The estimation of the real term structure from the prices of index-linked gilts (IGs) is considerably more complex than deriving the nominal yield curve from conventional bond prices. This is mainly because IG coupon payments are indexed to the level of RPI eight months before the cash flow occurs; for the last eight months of its life, an IG offers no inflation protection at all, and it therefore trades as a purely nominal bond. As a result, IG prices generally reflect a mixture of both the real and nominal term structures.

The approach we have been using up to now is described by Deacon and Derry (1994). By making an initial assumption about the expected future path of inflation, the real forward rate can be fitted (using a truncated Svensson curve). The difference between the real and nominal yield curves is then calculated where, assuming zero inflation-risk premia, this is determined by the market's inflation expectations. The real curve is then re-estimated using this new inflation assumption, and the process is repeated until convergence is obtained.

Evans (1998) introduced a new framework for dealing with the problems outlined above, avoiding the use of an iterative procedure. He derives a relationship between the nominal and real term structures and the term structure of (incompletely) indexed bonds,(2) allowing an interest rate curve to be fitted directly to IG prices. We have extended his work to account explicitly for the variation of the effective indexation lag for each IG's constituent cash flows, and also to deal with the delay in publication of the retail price index. A major advantage of this approach is that it is significantly more transparent than the iterative procedure.

Chart 7
Nominal, real and inflation forward curves (19 June 1998)



(1) On the other hand, we should be aware that GC repo rates can be affected by other factors. One example is gilt collateral shortages, although this effect may be diminished now that eligible collateral to be used in the Bank's operations has been extended to include many euro-denominated bonds (see *Quarterly Bulletin*, August 1999, pages 249–50).
(2) The index-linked term structure is a mathematical construct that simply allows us to price IGs using the standard discounted present value formula. It is not in itself an interesting term structure, since it is a mixture of the real and nominal curves.

Chart 7 presents preliminary estimates of the real and nominal yield curves using the VRP curve within this framework. It also shows the set of implied inflation expectations, calculated as the difference between these two curves. The real yield estimates do not differ markedly from those derived using the iterative technique. Instead, any differences in the nominal curve tend to be reflected by the set of inflation expectations. Work is still in progress to assess the relative performance of the two techniques.

Conclusion

In recent years, the Bank has used a model put forward by Svensson to estimate the UK nominal yield curve, and currently employs a similar parametric approach to derive the real yield curve. Experience of using these models has highlighted a number of problems. We have shown in this article that these problems can be resolved by using a spline-based technique. Moreover, this technique can be extended to provide estimates at the short end of the nominal yield curve, by including GC repo data. Further improvements relating to estimates of the real yield curve may by found by applying the spline technique to the theoretical framework put forward by Evans (1998).

In this article and in the November 1999 *Inflation Report*, we have presented our improved estimates of the nominal yield curve using the VRP technique. As work is still ongoing in relation to the real yield curve and inflation term structure, estimates presented continue to be based on Deacon and Derry's (1994) iterative technique. We intend to replace these with our new estimates of the real yield curve and inflation term structure in future editions of the *Quarterly Bulletin* and *Inflation Report*.

Appendix

This Appendix outlines the four methods for estimating the instantaneous forward rate curve discussed in the main text. The two parametric models were proposed by Nelson and Siegel (1987) and Svensson (1994, 1995). One of the spline-based models is the preferred choice of Fisher, Nychka and Zervos (1995), and the other is a modification (for the UK market) of the technique proposed by Waggoner (1997).

Parametric models

Nelson and Siegel proposed that the instantaneous forward rate curve could be parsimoniously modelled at all maturities by a parametric function of the form:

$$f(m,\beta) = \beta_0 + \beta_1 \exp\left(\frac{-m}{\tau_1}\right) + \beta_2\left(\frac{m}{\tau_1}\right)\exp\left(\frac{-m}{\tau_1}\right)$$

where $\beta = (\beta_0, \beta_1, \beta_2, \tau_1)'$ is the vector of parameters describing the curve, and m is the maturity at which the forward rate is evaluated. The functional form has three components: a constant term, an exponential decay term, and a 'hump-shaped' term. The curve asymptotes to a constant value of β_0 at the long end, and has a value of $(\beta_0 + \beta_1)$ at the short end.

To allow for additional flexibility in fitting the yield curve, Svensson proposed an extension to Nelson and Siegel's model, adding an extra hump term to give:

$$f(m,\beta) = \beta_0 + \beta_1 \exp\left(\frac{-m}{\tau_1}\right) + \beta_2\left(\frac{m}{\tau_1}\right)\exp\left(\frac{-m}{\tau_1}\right) + \beta_3\left(\frac{m}{\tau_2}\right)\exp\left(\frac{-m}{\tau_2}\right)$$

The curve is now described by six parameters: $\beta = (\beta_0, \beta_1, \beta_2, \beta_3, \tau_1, \tau_2)'$. Once again, the curve asymptotes to a constant value of β_0 at the long end, and has a value of $(\beta_0 + \beta_1)$ at the short end.

Smoothing cubic spline models

A generic spline is a piecewise polynomial, ie a curve constructed from individual polynomial segments joined at 'knot points', with coefficients chosen such that the curve and its first derivative are continuous at all points. The most commonly used polynomials are cubic functions, giving a cubic spline. The continuity constraints mean that any cubic spline can be written in the form:

$$S(x) = \alpha x^3 + \beta x^2 + \gamma x + \delta + \sum_{i=1}^{N-1} \eta_i \left| x - k_i \right|^3$$

for some constants, $\alpha, \beta, \gamma, \delta, \eta_i$, where k_i, $i = [0,N]$ is the set of knot points.

Though this is the simplest expression for a cubic spline, it is numerically unstable,(1) and so instead we prefer to represent our splines as a linear combination of cubic B-splines. This is a completely general transformation (any spline can be written as such a combination of B-splines of the appropriate order), which cures the numerical problems. B-splines of order n are most simply represented by the following recurrence relation:

$$B_{i,n}(x) = \frac{x - k_i}{k_{i+n-1} - k_i} B_{i,n-1}(x) + \frac{k_{i+n} - x}{k_{i+n} - k_{i+1}} B_{i+1,n-1}(x)$$

with $B_{i,1}(x) = 1$ if $k_i \le x < k_{i+1}$, and $B_{i,1}(x) = 0$ otherwise. For further details see Lancaster and Šalkauskas (1986).

With a sufficiently large number of knot points, a cubic spline can be used for interpolation. If this approach were adopted when fitting yield curves, the resulting term structures would be very different from the smooth curve that we require for monetary policy purposes. To reduce the flexibility of the spline, we can either reduce the number of knot points or impose a penalty on 'excessive' curvature (or non-smoothness). In both our spline-based models we use the latter approach, and the difference between the two methods lies in the different specifications of the penalty.

As described briefly in the main text, Fisher, Nychka and Zervos (1995) specify a roughness penalty that is constant across maturities, but which varies from day to day. So the objective function can be written:

$$X_{FNZ} = X_P + \lambda_t \int_0^M \left[f''(m) \right]^2 dm$$

where X_P is the duration-weighted sum of squared price residuals, and $f''(m)$ is the second derivative of the forward curve, and so a measure of its curvature. The constant λ_t is chosen for each day. If a large value is used, the curve is very smooth, and the effective number of parameters is reduced. Alternatively, a small value results in a very flexible curve, increasing the (in-sample) goodness of fit. The 'generalised cross-validation' technique is used to derive the optimum value of λ_t based on this trade-off between parsimony and goodness of fit, using the estimated curve and the observed bond prices.

Waggoner's VRP method (and our modification) uses a roughness penalty that is constant from day to day, but depends on maturity. The objective function can be written:

(1) Computers have only finite accuracy, and the calculation of a spline using this expression typically involves subtracting very large, similar numbers, resulting in (potentially) large errors.

$$X_{VRP} = X_P + \int_0^M \lambda(m)[f''(m)]^2 \, dm$$

In this case, we need to specify a functional form for the smoothing function. We use:

$\log \lambda(m) = L - (L - S) \exp(-m / \mu)$

where, L, S and μ are parameters to be estimated. The smoothing parameters were chosen to maximise the out-of-sample goodness of fit, with a preference for higher smoothing when (as was found to be the case in practice) several combinations of the parameters gave similar out-of-sample goodness of fit measures.

References

Breedon, F (1995), 'Bond prices and market expectations of inflation', *Bank of England Quarterly Bulletin*, Vol 35(2), pages 160–65.

Davison, A C and Hinkley, D V (1997), 'Bootstrap methods and their applications', Cambridge University Press, Cambridge.

Deacon, M and Derry, A (1994), 'Deriving estimates of inflation expectations from the prices of UK government bonds', *Bank of England Working Paper*, No 23.

Evans, M D D (1998), 'Real rates, expected inflation, and inflation risk premia', *Journal of Finance*, Vol 53, page 187.

Fisher, M, Nychka, D and Zervos, D (1995), 'Fitting the term structure of interest rates with smoothing splines', *Finance and Economic Discussion Series*, 95-1, Federal Reserve Board.

Lancaster, P and Šalkauskas, K (1986), 'Curve and surface fitting: an introduction', Academic Press, London.

Mastronikola, K (1991), 'Yield curves for gilt edged stocks: a new model', *Bank of England Discussion Paper (Technical Series)*, No 49.

Nelson, C R and Siegel, A F (1987), 'Parsimonious modelling of yield curves', *Journal of Business*, Vol 60, page 473.

Svensson, L (1994), 'Estimating and interpreting forward interest rates: Sweden 1992–94', *IMF Working Paper*, No 114.

Svensson, L (1995), 'Estimating forward interest rates with the extended Nelson & Siegel method', *Sveriges Riksbank Quarterly Review*, 1995:3, page 13.

Waggoner, D (1997), 'Spline methods for extracting interest rate curves from coupon bond prices', *Working Paper*, No 97-10, Federal Reserve Bank of Atlanta.