Reserve Bank of Australia

RESEARCH DISCUSSION PAPER

A Term Structure Decomposition of the Australian Yield Curve

Richard Finlay and Mark Chambers

RDP 2008-09

A TERM STRUCTURE DECOMPOSITION OF THE AUSTRALIAN YIELD CURVE

Richard Finlay and Mark Chambers

Research Discussion Paper
2008-09

December 2008

Domestic Markets Department
Reserve Bank of Australia

The authors thank Meredith Beechey, Adam Cagliarini, Jonathan Kearns, Christopher Kent, Kristoffer Nimark and Anthony Richards for useful comments and suggestions. Responsibility for any remaining errors rests with the authors. The views expressed in this paper are those of the authors and are not necessarily those of the Reserve Bank of Australia.

Authors: finlayr or chambersm at domain rba.gov.au

Economic Publications: ecpubs@rba.gov.au

Abstract

We use data on coupon-bearing Australian Government bonds and overnight indexed swap (OIS) rates to estimate risk-free zero-coupon yield and forward curves for Australia from 1992 to 2007. These curves, and analysts' forecasts of future interest rates, are then used to fit an affine term structure model to Australian interest rates, with the aim of decomposing forward rates into expected future overnight cash rates plus term premia. The expected future short rates derived from the model are on average unbiased, fluctuating around the average of actual observed short rates. Since the adoption of inflation targeting and the entrenchment of low and stable inflation expectations, term premia appear to have declined in levels and displayed smaller fluctuations in response to economic shocks. This suggests that the market has become less uncertain about the path of future interest rates. Towards the end of the sample period, term premia have been negative, suggesting that investors may have been willing to pay a premium for Commonwealth Government securities. Due to the complexity of the model and the difficulty of calibrating it to data, the results should not be interpreted too precisely. Nevertheless, the model does provide a potentially useful decomposition of recent changes in the expected path of interest rates and term premia.

JEL Classification Numbers: C51,E43,G12
Keywords: expected future short rate, term premia, term structure decomposition, affine term structure model, zero-coupon yield

Table of Contents

A TERM STRUCTURE DECOMPOSITION OF THE AUSTRALIAN YIELD CURVE

Richard Finlay and Mark Chambers

1. Introduction

The relationship between the level of interest rates across different maturities is known as the term structure of interest rates. The term structure can be used to assess the financial markets' expectations for the future path of monetary policy. For example, the pure expectations hypothesis (which ignores the possible existence of term premia) asserts that market participants' expectations of future short-term interest rates are simply given by forward rates as observed in the market.[1]

The term structure of interest rates is often presented as a yield curve, which plots the yields to maturity of bonds with varying terms to maturity. Typically, the yield curve is presented for risk-free interest rates. In Australia, Australian Government bonds are normally used, since these are considered to have essentially zero probability of default and hence the yields do not incorporate any credit risk premia. However, the yield curve does not give a direct reading of interest rate expectations for two reasons. First, the yield to maturity of a bond is affected by the bond's coupon rate; the higher the coupon rate, the less important will be the payment at maturity as a share of the bond's total income stream and hence the yield to maturity will be affected more by short-term expectations of monetary policy relative to longer-term expectations. Second, if investors are risk-averse and the future path of interest rates is uncertain, then long-term interest rates will incorporate a term premium as compensation for investing in the face of this uncertainty.

If these two components of long-term yields can be stripped away, the resulting curve would provide a better indication of the markets' expectations of the future path of short-term interest rates, specifically the overnight interest rate used by the Reserve Bank of Australia as the instrument for monetary policy.

---

[1] By forward rate we mean an overnight interest rate which is observed in the market now but does not apply until some time in the future.

To abstract from the first of these complications, it is possible to use a set of yields on coupon bonds – that is physical government bonds – to estimate a set of yields on (hypothetical) zero-coupon bonds, which are bonds that do not make any periodic interest payments. There are a number of established methods to do this, which give broadly similar results.

The most direct method to abstract from the second complication – that is, to estimate expected future short rates separate from term premia – would be to use analysts' forecasts of future monetary policy decisions, as these give a direct reading on cash rate expectations. However, this method suffers from a number of drawbacks, chief among these being that analysts' expectations may not always be reflected in market pricing, and typically extend over only a relatively short horizon. An alternative is to specify and estimate a model of how expected future short rates and term premia evolve over time. The fact that these two elements are time-varying and are confounded in their effect on bond prices makes the choice of model crucial. The approach we employ is to combine these two methods, using data on analysts' forecasts within the model-based approach to aid separate identification of expected future short rates and term premia. Nevertheless, the central role of the assumed model (along with the computational complexities of fitting the model to data) means that it is prudent to treat the results of such a term structure model with some caution – a different model may generate different results.

Despite these caveats, the importance of the shape of the yield curve and expectations of future interest rates in understanding economic and financial market developments make the separation of yields into term premia and expectations a worthwhile exercise. To this end we employ an affine term structure model of zero-coupon yields that has been used widely in the literature and currently appears to be the best available candidate for such work.[2]

The remainder of this paper is set out as follows. Section 2 provides a brief overview of the affine term structure model, the literature on affine term structure models, and their development. Section 3 details the term structure model that we employ, while Section 4 discusses how we use estimated zero-coupon yield

2 An affine term structure model represents interest rates as being a linear combination of a small set of factors and parameters. See, for example, Duffee (2002) and Dai and Singleton (2002) for discussion of competing term structure models.

data, along with analysts' forecasts of future interest rates, as the inputs into the estimation procedure for our model. Section 5 gives the results of our estimation over two sample periods, with the output of most interest being the expected future short rates and term premia produced. Finally, Section 6 concludes. More technical detail regarding zero-coupon yield curve estimation from data on coupon-bearing Australian Government bonds, as well as the affine term structure model and its implementation, are provided in the appendices.

2. Model Overview and Related Literature

The focus of this paper is the estimation of an affine term structure model for Australian interest rates, with the aim of decomposing forward rates into expected future short rates and term premia. While mathematical details of the model are given in Section 3, a brief description of the model here provides the reader with some intuition regarding what is to follow.

We start by estimating zero-coupon yield curves from observed overnight indexed swap (OIS) and government bond data (for further details see Section 4 and Appendix A). These, along with analysts' forecasts of future interest rates, constitute the data used to estimate our term structure model.

Our term structure model describes how the cash rate might evolve. The model assumes that the cash rate can be expressed as a constant plus the sum of three latent factors, which in turn follow the continuous time equivalent of a vector auto-regressive process with normally distributed shocks. Each latent factor is assumed to have zero mean, so that according to our model, the cash rate has a constant long-run steady-state value. The cash rate moves away from this steady-state value when shocks cause the latent factors to move away from zero.

Arbitrage conditions allow us to link bond prices to the evolution of the cash rate. In a world where investors are risk-neutral, the price of a zero-coupon bond would be given by the expectation of the bond's discounted future pay-off, where discounting is with respect to the cash rate process just described. However, investors need not be risk-neutral. If they are risk-averse, they may require extra compensation for holding a bond whose value fluctuates, as opposed to cash whose value does not. This extra compensation can be considered as the term premium.

However, exactly how investors' risk preferences collectively affect term premia is not clear *a priori*. On the one hand, it is reasonable to think that investors should be compensated for holding long-term bonds over cash, since the value of long-term bonds can fluctuate and thus expose investors to the possibility of mark-to-market losses. On the other hand, for investors who have long-term fixed liabilities, a long-term bond for which the value at maturity is fixed may be *less* risky than a cash account for which the value will depend on the variable path of short-term interest rates. Term premia could therefore be positive or negative, depending on the mix of investors trading bonds.

Hence, bond prices (and therefore observed yields) depend on both expected future short rates and term premia. Of course observations of bond yields alone are not sufficient to separately identify these two components. We can get information about expected future short rates separate from term premia in two ways. First, we can obtain estimates of the latent factors which can be used to derive expected future short rates. Second, we can augment the zero-coupon yield data with analysts' forecasts of future interest rates when estimating the model – forecasts of the future cash rate are a direct reading of expected future short rates separate from term premia, and so aid in the estimation of the actual short rate process.

The latent factors are not observable, but must be estimated along with the parameters of the model. We use the Kalman filter and maximum likelihood to estimate the latent factors and parameters. The latent factors are estimated so as to provide the best fit possible between the model's implied yields and the actual observed yields. Although no economic structure is imposed on them, the latent factors tend to explain different components of the yield curve. Typically one latent factor is highly correlated with the level of the yield curve, another is correlated with the slope of the yield curve, and the third is correlated with the curvature of the yield curve.

The model of interest rates just described builds on a modelling approach that was first proposed in Duffie and Kan (1996). That work introduces the affine term structure model, an arbitrage-free multifactor model of interest rates in which the yield on any risk-free zero-coupon bond is an affine function of a set of unobserved latent factors. Duffie and Kan also provide a method to obtain the coefficients on the latent factors in the affine function and therefore to price risk-free zero-coupon bonds. The improvement of this model on the previous literature is that it is scaleable, driven by estimable factors which have arbitrary correlation, while at the same time retaining a good level of tractability.

de Jong (2000) implements this model on Treasury yield data from the United States. He estimates one-, two- and three-factor versions of the model, concluding that the one- and two-factor versions are misspecified, but that the three-factor version seems to do a good job of capturing the relevant dynamics of yields. de Jong uses a Kalman filter in estimating the models, which has the advantage that it provides tractable estimation when there are more input yields than factors. Consequently, it has become the most common technique for estimating affine term structure models.

Duffee (2002) generalises the specification of the market price of risk used by Duffie and Kan (1996) and de Jong (2000). He removes the restriction that compensation for interest rate risk must be a multiple of the variance of that risk and suggests a modification which allows it to move independently of the variance. Duffee estimates this new variant (called the '$A_0(3)$' model), the original model and a hybrid model, and demonstrates that the extra flexibility of the $A_0(3)$ model provides significant improvements to goodness-of-fit.

Dai and Singleton (2002) implement various specifications of the Duffee (2002) model on US data. They show that while regular yields fail the expectations hypothesis, the 'risk-premium adjusted' yields from the $A_0(3)$ model satisfy the expectations hypothesis. A further contribution of Dai and Singleton is that they also provide analytical formulae for the coefficients of the affine function, enabling simpler estimation than the method of Duffie and Kan (1996).

Kim and Orphanides (2005) take the $A_0(3)$ model of Duffee (2002) but incorporate survey data of analysts' forecasts of short-term interest rates as an additional input to the estimation problem. Using US data, they estimate models both with and without the forecasts and find that those models that incorporate forecasts produce a better fit. Monte Carlo trials suggest that the inclusion of forecasts helps to reduce small-sample problems arising in the estimation of highly persistent factors, especially when data sets of only limited length are available. They find that between the early 1990s and 2003, term premia in the US fell and that the fall was tied to the moderation of macroeconomic volatility seen over the period. The fall in term premia helps to explain the fall in treasury yields also observed. The model used in this paper is a variation of the Kim and Orphanides model, changed slightly to accommodate the different nature of our survey data.

Affine term structure models have also been implemented at other central banks. Kremer and Rostagno (2006) from the European Central Bank use a two-factor affine term structure model to examine the low bond yields observed in the euro area over the first half of this decade. They find a sharp reduction in estimated term premia, indicating that a reduction in risk compensation may have been driving yields lower. In addition, the term premia are found to be related to measures of liquidity, suggesting that excess liquidity may also have been playing a part in driving risk aversion down.[3]

Westaway (2006) also finds falling term premia in the United Kingdom. Given the complexities of the model and the fact that term premia are in effect residuals of the model he is, however, somewhat cautious in interpreting the results. Westaway estimates a dynamic stochastic general equilibrium (DSGE) model of a closed economy and finds that a decline in the volatility of economic shocks should lead to lower term premia, a result consistent with the term structure model. However, the DSGE model does not result in an overall fall in real yields and so cannot fully account for the low level of yields observed.

More broadly, the strategy of incorporating time-varying term premia in modelling long-run interest rates is a response to extensive empirical evidence contradicting the pure expectations hypothesis; that is, evidence that long-run interest rates are *not* simply an average of the expected path of future short-term interest rates. In particular, studies have found that long-run interest rates display both ‘excess volatility’ (fluctuating more than would be expected given the volatility of the underlying macroeconomy) and ‘excess sensitivity’ (responding to information that might be expected to only influence short-term rates).[4]

The estimation approach used in this paper belongs to the ‘pure-finance’ branch of the term structure literature, where term premia are estimated using observed yield data and perhaps some survey forecast data. This is opposed to the ‘macro-finance’ branch, typified by Rudebusch and Wu (2008), where the interaction

3 In this context, excess liquidity refers to the amount of money and liquid assets circulating in the economy.

4 See Gürkaynak, Sack and Swanson (2003) or Beechey (2004) for an overview of this literature.

between the macroeconomy and the term structure is also modelled. As noted by Kim and Wright (2005), pure-finance models, which rely on latent factors to explain the yield curve, generally have the advantage of being more robust to model misspecification, and provide a better fit to the data, than macro-finance models. Conversely, although macro-finance models generally do not fit the data as well as pure-finance models, they may be easier to interpret from an economic viewpoint given the structure that they impose.

One criticism of the term structure literature is given in Swanson (2007), who argues that different modelling techniques result in different term premia estimates, so that some degree of caution must be placed on any term premia estimate. The criticism is a reasonable one – term premia by their nature are hard to estimate since their effect on observable bond prices is confounded with expectations of future short-term rates. On the other hand, it is not entirely surprising that different modelling techniques, which make different assumptions about financial markets and the economy, should produce different results. A useful survey paper on this topic is that of Rudebusch, Sack and Swanson (2007), who review five alternative term premia estimation methodologies. They find that although different models do produce different term premia estimates, the estimates are generally not *too* different.[5]

We make some modest contributions to these term structure models; we extend the Kim and Orphanides (2005) model to accommodate a different type of forecast data (cash rate and 10-year bond yield forecasts as opposed to treasury note forecasts), and we extend the zero-coupon yield estimation method to allow the model to account for the actual cash rate prevailing at any given time. Our larger contribution is the estimation of zero-coupon bond yields, and a linear affine term structure model, for Australia.

5 Rudebusch *et al* (2007) consider term premia estimates for US data arising from five different term strucuture models, one of which is equivalent to the term structure model which we use. They find that the model which is equivalent to our model produces term premia which are very similar to those produced by two other models (correlation coefficients of 0.98 and 0.94); that the model which is equivalent to our model produces term premia which are very similar, except for a level shift, to another model (correlation coefficient of 0.96); and finally, that the last model (correlation coefficient of 0.81) produces term premia which are less similar to all other models for theoretical reasons regarding modelling assumptions.

3. The Model in Detail

In this section we outline the model we use. In what follows, scalars are lower-case and not bold, vectors are bold upper-case and lower-case, and matrices are upper-case and not bold. For mathematical convenience and to be consistent with the literature, the model is considered in continuous time; discrete time versions of such models are also possible.

Let r_t be the instantaneous short rate or cash rate and assume that

$$r_t = \rho + \mathbf{1}' \cdot \mathbf{x}_t \tag{1}$$

where $\mathbf{1} = (1,1,1)'$, $\mathbf{x}_t = (x_{1,t}, x_{2,t}, x_{3,t})'$, and

$$\mathrm{d}\mathbf{x}_t = -K\mathbf{x}_t \mathrm{d}t + \Sigma \, \mathrm{d}\mathbf{W}_t, \tag{2}$$

where K is lower triangular and Σ is diagonal, both 3×3 matrices, and $\mathbf{W}_t$ is standard multivariate Brownian motion which is analogous to a continuous time version of a random walk. Equations (1) and (2) imply that the short rate is a function of a constant ρ and three time-varying ('latent') factors, $\mathbf{x}_t = (x_{1,t}, x_{2,t}, x_{3,t})'$, with the evolution of $\mathbf{x}_t$ following a zero-mean Ornstein-Uhlenbeck process, the continuous time analogue of a vector auto-regressive process. Here the drift term $-K\mathbf{x}_t\mathrm{d}t$ is the deterministic component of the stochastic differential equation, with K controlling the speed of mean reversion, and the diffusion term $\Sigma \, \mathrm{d}\mathbf{W}_t$ is the random component, with the Brownian motion $\mathbf{W}_t$ providing random shocks to the system. As mentioned earlier, the latent factors $\mathbf{x}_t$ are not observable and need to be estimated with the parameters of the model.

Investors demand compensation for holding bonds, whose value depends on the random, and hence risky, latent factors; cash is free of this risk. The amount of compensation demanded is termed the *market price of risk*, and it is this price of risk that determines term premia (it is worth emphasising that the price of risk and term premia are not the same thing; see Section 4). We assume that the price of risk is of the form

$$\boldsymbol{\lambda}_t = \boldsymbol{\lambda}_0 + \Lambda \mathbf{x}_t \tag{3}$$

where $\boldsymbol{\lambda}_t = (\lambda_{1,t}, \lambda_{2,t}, \lambda_{3,t})'$, with $\lambda_{i,t}$ the price of risk associated with the latent factor $x_{i,t}$ at time t, $\boldsymbol{\lambda}_0 = (\lambda_{0,1}, \lambda_{0,2}, \lambda_{0,3})'$ and Λ a 3×3 matrix. This specification implies that for each i, the extra compensation demanded by investors for bearing the risk of $x_{i,t}$ is comprised of a constant $\lambda_{0,i}$ plus a linear combination of the latent factors, $(\Lambda)_{1i}x_{1,t} + (\Lambda)_{2i}x_{2,t} + (\Lambda)_{3i}x_{3,t}$.

Given this model, the arbitrage-free price of a zero-coupon bond at time t, paying 1 unit at $t+\tau$, is given by

$$P_{t,\tau} = \mathbb{E}_t^* \left[\exp\left(-\int_t^{t+\tau} r_s \mathrm{d}s\right) \right] \tag{4}$$

where the expectation is taken with respect to the risk-neutral probability distribution (also referred to as the risk-neutral measure or equivalent martingale measure).[6] The risk-neutral probability distribution adjusts the actual (real-world) probability distribution for investors' risk preferences, and so under this new distribution we can treat investors as if they were risk-neutral. This means that under the risk-neutral distribution we can price any asset by simply calculating the expected discounted present value of its future cash flows (see Appendix B for a sketch of a proof).[7]

Equations (1) and (2) describe the dynamics of the short rate under the real-world probability distribution. To obtain the dynamics of the short rate under the risk-neutral probability distribution we subtract Σ times the market price of risk, as given by Equation (3), from the drift component of $\mathbf{x}_t$, as given by Equation (2), to obtain

$$\begin{aligned} \mathrm{d}\mathbf{x}_t &= (-K\mathbf{x}_t - \Sigma\boldsymbol{\lambda}_t)\mathrm{d}t + \Sigma\, \mathrm{d}\mathbf{W}_t^* \\ &= -((K + \Sigma\Lambda)\mathbf{x}_t + \Sigma\boldsymbol{\lambda}_0)\mathrm{d}t + \Sigma\, \mathrm{d}\mathbf{W}_t^*. \end{aligned} \tag{5}$$

While $\mathbf{W}_t$ is Brownian motion under the real-world probability distribution, it is not Brownian motion under the risk-neutral distribution. However, $\mathbf{W}_t$ is related to Brownian motion under the risk-neutral probability distribution, denoted by $\mathbf{W}_t^*$, according to $\mathbf{W}_t^* = \mathbf{W}_t + \int_0^t \boldsymbol{\lambda}_s \mathrm{d}s$, or equivalently $\mathrm{d}\mathbf{W}_t^* = \mathrm{d}\mathbf{W}_t + \boldsymbol{\lambda}_t \mathrm{d}t$. In other words, $\mathbf{W}_t^*$ is derived by adjusting $\mathbf{W}_t$ for the market price of risk, given by $\boldsymbol{\lambda}_t$.[8]

---

6 See, for example, Duffie and Kan (1996).

7 For more detail on risk-neutral probability distributions see, for example, Cochrane (2001) or Steele (2001).

8 See, for example, de Jong (2000) or Dai and Singleton (2002).

Given Equations (1) and (5), Duffie and Kan (1996) show that the price of a zero-coupon bond (Equation (4)) can be simplified to

$$P_{t,\tau} = \exp\left[-\alpha_\tau - \boldsymbol{\beta}'_\tau \mathbf{x}_t\right] \tag{6}$$

where α_τ and $\boldsymbol{\beta}_\tau$ are functions of the underlying parameters ρ, K, Σ, $\boldsymbol{\lambda}_0$ and Λ (see Appendix C for details). Given that we can infer zero-coupon bond prices from government coupon bond data, we can estimate the parameters of the model by minimising the difference between zero-coupon bond prices and those prices implied by Equation (6).

Note that from Equations (1) and (2), the only parameters of the model which affect the short rate, and therefore which determine estimates of the expected future short rate, are ρ, K and Σ. On the other hand observed bond prices, as specified by Equation (6), incorporate term premia and are therefore also affected by the parameters determining the market price of risk: $\boldsymbol{\lambda}_0$ and Λ. Hence in order to separate expected future short rates from term premia we need estimates of $\boldsymbol{\lambda}_0$ and Λ as well as ρ, K and Σ. However, the matrices K and Λ only appear in the formulas for α_τ and $\boldsymbol{\beta}_\tau$ (and hence only impact on bond prices) in the form $(K+\Sigma\Lambda)$. That is, when they do appear they only appear together. This means that observed market prices in and of themselves do not identify K (which in the sense just described determines expected future short rates) separately from Λ (which likewise determines term premia).

Instead we rely on the fact that the latent factors evolve according to the real-world probability distribution as given in Equation (2), where K does appear without Λ. We also use analysts' forecasts of future interest rates, which give a clean reading on expected future short rates abstracting from term premia. As our forecast data are relatively sparse, estimates of how the latent factors $\mathbf{x}_t$ evolve play a large role in separating K from Λ, and these latent factors must in turn be estimated from the data.

4. Data and Model Implementation

Estimation of the model presented in Section 3 requires observations of zero-coupon bond yields. As zero-coupon bonds are not currently issued in Australia, we need some way to infer these yields from coupon-bearing Australian Government bonds.

We estimate zero-coupon bond prices from coupon-bearing Australian Government bond data using a modified Merrill Lynch Exponential Spline (MLES) methodology.[9] This amounts to estimating a risk-free discount function, which we take as a linear combination of hyperbolic basis functions.[10] As the estimation of the zero-coupon yield curve is not the primary focus of this paper we provide the technical details and a discussion of the issues involved in Appendix A rather than in the main text. A number of different zero-coupon estimation methodologies were considered, with the MLES method chosen due to its ease of implementation and goodness-of-fit.

To estimate the risk-free zero-coupon yield curve at the short end, we use Treasury notes when they are available and OIS rates with maturities less than or equal to 1 year when Treasury notes are not available.[11] For maturities longer than 18 months we use the yields of Australian Government bonds. Bonds with shorter maturities can become quite illiquid, and tend to suffer from pricing anomalies. We calculate zero-coupon rates at terms to maturity of 3 and 6 months, as well as for 1, 2, 4, 6, 8 and 10 years. The data are sampled at weekly intervals between July 1992 and April 2007.

We supplement these data with survey forecasts of the cash rate and the 10-year bond yield.[12] The cash rate forecast data are roughly monthly and are available from March 2000 to April 2007 for forecast horizons from 1 to 8 quarters. These forecasts are not available every month, or at all horizons when they are available; the majority come after March 2002 and are for horizons out to 1 year. The

---

[9] Our modification of the MLES procedure results in the 1-day yield being fixed at the target cash rate. See, for example, Bolder and Gusba (2002) for a discussion of competing estimation methodologies.

[10] The discount function evaluated at t gives the value today of 1 unit at time t in the future.

[11] OIS contracts are over-the-counter derivatives in which one party agrees to pay the other party a fixed interest rate in exchange for receiving the average cash rate recorded over the term of the swap. As no principal is exchanged these contracts are virtually risk-free, and so the fixed rates paid are a good approximation of the average cash rate expected to prevail over the life of the contract. Hence they can be used in place of Treasury notes to estimate the short end to the risk-free yield curve. See RBA (2002) for details of how OIS contracts operate, and Appendix A for more discussion on OIS rates.

[12] The cash rate forecast data are compiled from Bloomberg, Reuters and Consensus Economics, while the 10-year yield forecasts come from Consensus Economics.

10-year bond yield expectation data are monthly, run from December 1994 to April 2007, and are for horizons of between 3 months and 10 years. In addition to their helpfulness in identification as discussed above, survey data have been shown to counteract many small sample problems (including different parameter sets giving similar model outputs, the mean reversion of latent factors being too fast, and imprecise estimates). Although survey data give average expectations, not the marginal investor's expectation, this is unlikely to be a major problem. In fact survey data have been found to greatly improve accuracy and stability in model estimation.[13]

Since the pricing equation, Equation (6), requires knowledge of the latent factors, which are unobservable, these latent factors need to be estimated along with the parameters of the model. This is done via the Kalman filter. Using Equation (6), we can write the zero-coupon yield as implied by the term structure model at time t, for a bond maturing at time $t+\tau$, as

$$y_{t,\tau} = a_\tau + \mathbf{b}'_\tau \mathbf{x}_t \tag{7}$$

where $a_\tau = \alpha_\tau/\tau$ and $\mathbf{b}_\tau = \boldsymbol{\beta}_\tau/\tau$ are both functions of ρ, Σ, $\boldsymbol{\lambda}_0$ and $(K+\Sigma\Lambda)$. Our term structure implied zero-coupon yields should match the zero-coupon yields we have estimated using traded government bond and OIS rates, however, and so for each observation occurring at time t we can then stack the versions of Equation (7) corresponding to each maturity, τ, as follows

$$\begin{bmatrix} y_{t,0.25} \\ \vdots \\ y_{t,10} \end{bmatrix} = \begin{bmatrix} a_{0.25} \\ \vdots \\ a_{10} \end{bmatrix} + \begin{bmatrix} \mathbf{b}'_{0.25} \\ \vdots \\ \mathbf{b}'_{10} \end{bmatrix} \mathbf{x}_t + \begin{bmatrix} \eta_{t,0.25} \\ \vdots \\ \eta_{t,10} \end{bmatrix}$$

or in matrix notation

$$\mathbf{y}_t = \mathbf{a} + B\mathbf{x}_t + \boldsymbol{\eta}_t. \tag{8}$$

Here $\mathbf{y}_t$ gives the observed zero-coupon yields, and the error term $\boldsymbol{\eta}_t$ occurs because our term structure model implied yields $\mathbf{a}+B\mathbf{x}_t$ will not fit the observed

[13] See Kim and Orphanides (2005) – they compare models that use and do not use survey data, and perform Monte Carlo simulations on the effect of survey data, finding that survey data counter many of the small sample problems just discussed (note that we use surveys of cash rate expectations and bond yields, whereas they use surveys of the expected yield on US Treasury notes).

yields exactly. Note that because the a_τ and $\mathbf{b}'_\tau$ are functions of $(K+\Sigma\Lambda)$, Equation (8) on its own does not help us separate expected future short rates (determined by K) from term premia (determined by Λ).

We can use the discrete version of Equation (2), however, to write the state equation for the latent factors $\mathbf{x}_t$ as

$$\mathbf{x}_t = e^{-Kh}\mathbf{x}_{t-h} + \boldsymbol{\varepsilon}_t \tag{9}$$

where in our case $h = 7/365$ (to account for weekly sampling of the data) and $\boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \Omega_h)$ with $\Omega_h = \int_0^h e^{-Ks}\Sigma\Sigma' e^{-K's}\mathrm{d}s$.[14] In Equation (9) K appears on its own, and so with estimates of the latent factors $\mathbf{x}_t$ we can infer information about K separate from Λ.

On dates for which there are survey forecasts, Equation (7) changes slightly. Using Equations (1) and (9) we can express cash rate forecasts as

$$\tilde{y}_{t,\tau} = \rho + \mathbf{1}' \cdot e^{-K\tau}\mathbf{x}_t + \tilde{\eta}_{t,\tau} \tag{10}$$

where: τ is the length of time between t and the forecast date; $\tilde{y}_{t,\tau}$ is the cash rate forecast; and $\tilde{\eta}_{t,\tau}$ denotes the forecast error. Similarly, for bond yield forecasts we can write

$$\bar{y}_{t,\tau} = a_{10} + \mathbf{b}'_{10} \cdot e^{-K\tau}\mathbf{x}_t + \bar{\eta}_{t,\tau} \tag{11}$$

where: $\bar{y}_{t,\tau}$ is the yield forecast; and $\bar{\eta}_{t,\tau}$ denotes the forecast error.[15] Note that in both Equations (10) and (11), K appears on its own, which helps us identify K and Λ separately.

---

14 Ω_h can be evaluated as $-vec^{-1}(((K \otimes I) + (I \otimes K))^{-1} vec(e^{-Kh}\Sigma\Sigma' e^{-K'h} - \Sigma\Sigma'))$. See Kim and Orphanides (2005).

15 Here we treat a yield to maturity as a zero-coupon yield. Analysis of historically observed yield data shows that the observed yield to maturity on a 10-year bond and the estimated 10-year zero-coupon yield are in fact very close, so this should not be a problem: the difference between the two yield measures is symmetric about zero and has an average absolute size of only 4 basis points. This is less than the precision of the forecast data, which is only reported to the nearest 10 basis points, and well below the forecast error of 50 basis points per square root year that we have assumed (see Appendix C for technical specifications of the model implementation).

Incorporating the forecast data into our estimation, we can stack Equations (8), (10) and (11) to give the observation equation

$$\begin{bmatrix} y_{t,0.25} \\ \vdots \\ y_{t,10} \\ \tilde{y}_{t,\tau_1} \\ \vdots \\ \tilde{y}_{t,\tau_{n_1}} \\ \bar{y}_{t,\tau_1} \\ \vdots \\ \bar{y}_{t,\tau_{n_2}} \end{bmatrix} = \begin{bmatrix} a_{0.25} \\ \vdots \\ a_{10} \\ \rho \\ \vdots \\ \rho \\ a_{10} \\ \vdots \\ a_{10} \end{bmatrix} + \begin{bmatrix} \mathbf{b}'_{0.25} \\ \vdots \\ \mathbf{b}'_{10} \\ \mathbf{1}' \cdot e^{-\tau_1 K} \\ \vdots \\ \mathbf{1}' \cdot e^{-\tau_{n_1} K} \\ \mathbf{b}'_{10} \cdot e^{-\tau_1 K} \\ \vdots \\ \mathbf{b}'_{10} \cdot e^{-\tau_{n_2} K} \end{bmatrix} \mathbf{x}_t + \begin{bmatrix} \eta_{t,0.25} \\ \vdots \\ \eta_{t,10} \\ \tilde{\eta}_{t,\tau_1} \\ \vdots \\ \tilde{\eta}_{t,\tau_{n_1}} \\ \bar{\eta}_{t,\tau_1} \\ \vdots \\ \bar{\eta}_{t,\tau_{n_2}} \end{bmatrix}$$

or in matrix notation

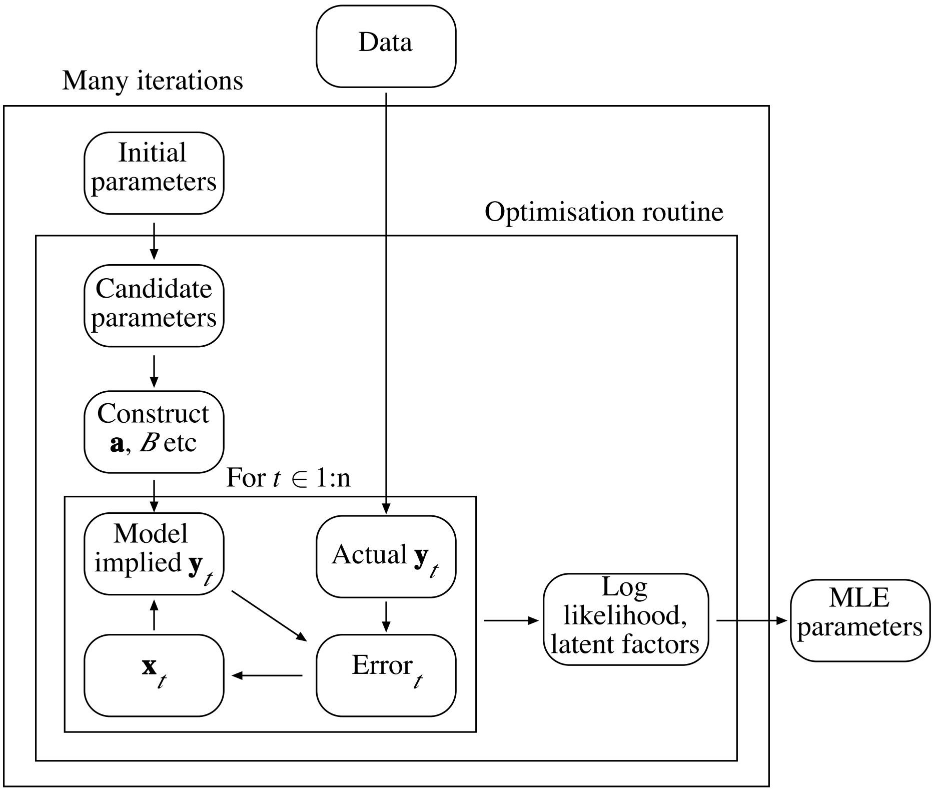
$$\tilde{\mathbf{y}}_t = \tilde{\mathbf{a}} + \tilde{B}\mathbf{x}_t + \tilde{\boldsymbol{\eta}}_t. \tag{12}$$

To summarise, Equations (12) and (9) then make up the Kalman filter observation and state equations, respectively. These can be used to compute the maximum likelihood estimate of $\mathbf{x}_t$ and the parameters ρ, Σ, $\boldsymbol{\lambda}_0$, Λ and K, using the zero-coupon yield data and survey data, which together constitute $\tilde{\mathbf{y}}_t$. See Appendix C for further details. As mentioned earlier, K enters our equations separately from Λ via the dynamics of $\mathbf{x}_t$, given by Equation (9), and via the survey forecasts as given in Equation (12).

To estimate the parameters of the model we randomly generate a vector of starting parameters, specify the starting values of the latent factors $\mathbf{x}_t$, and then use the MATLAB® `fmincon` function to search for a log-likelihood maximum. The search is based on a sequential quadratic programming routine. This is repeated 2 000 times and the set of parameters producing the highest likelihood is chosen.

This estimation procedure is displayed graphically in Figure 1. One at a time, each of the 2 000 randomly generated sets of initial parameters are fed into the optimisation routine. The routine uses the initial parameter guess to construct the parameters used by the model, such as **a** and B. Using the Kalman filter, the yield data are used to estimate the latent factors and model implied yields. The Kalman filter also produces the log-likelihood, which the optimisation routine uses to choose a new set of candidate parameters, and the procedure is then repeated. Once

Figure 1: Flow Diagram of Estimation Procedure



the optimisation routine has ended, the highest log-likelihood and the associated parameter values are stored, and the process begins again. After 2 000 iterations, the parameters that produced the highest overall log-likelihood value are chosen.

A number of alternative optimisation procedures are possible. We explored simulated annealing, as well as some other in-built MATLAB® functions, but found that the procedure described above gave the best (highest likelihood values in reasonable time) results.

Finally, from Equation (9) we have $\mathbb{E}_t[\mathbf{x}_{t+\tau}] = e^{-K\tau}\mathbf{x}_t$, so that having estimated the parameters and latent factors of the model, using Equation (1) we can calculate for time t the expected future short rate (efsr) at time $t+\tau$ as

$$\text{efsr}_{t,\tau} = \rho + \mathbf{1}' \cdot e^{-K\tau}\mathbf{x}_t.$$

Similarly, from Equation (5) where we are now considering $\mathbf{x}_t$ under the *risk-neutral* probability distribution, for $K^* = K + \Sigma\Lambda$ and $\boldsymbol{\mu}^* = K^{*-1}\Sigma\boldsymbol{\lambda}_0$ we have that $\mathbb{E}^*_t[\mathbf{x}_{t+\tau}] = e^{-K^*\tau}\mathbf{x}_t - (I - e^{-K^*\tau})\boldsymbol{\mu}^*$ (see, for example, Kim and Orphanides 2005).

Hence at time t, the model implied forward rate (fr) for time $t+\tau$ in the future is given by

$$\mathrm{fr}_{t,\tau} = \rho + \mathbf{1}' \cdot (e^{-K^*\tau}\mathbf{x}_t - (I - e^{-K^*\tau})\boldsymbol{\mu}^*).$$

But the forward rate at time t applying at time $t+\tau$ in the future consists of expectations of the cash rate at time $t+\tau$ plus the term premium. Hence, our estimate at time t of the term premium (tp) associated with borrowing or lending at time $t+\tau$ in the future is given by

$$\mathrm{tp}_{t,\tau} = \mathrm{fr}_{t,\tau} - \mathrm{efsr}_{t,\tau}.$$

5. Results

We estimate the model over two time horizons. First we use all available data so that our sample runs from July 1992 to April 2007. Then we restrict the sample to the period July 1996 to April 2007. This shorter sample covers the period when inflation expectations have been reasonably stable and consistent with the Reserve Bank's inflation target, and corresponds to a period of stable growth and low inflation. (In both cases we use the first six months of data to estimate the latent factors, but discard it when estimating model parameters.)

5.1 The Period 1993 to 2007

The primary variables of interest are the estimated expected future short rates and term premia, which are shown in Figure 2.

As expected, the forward rate and the expected future short rate tend to track each other closely at the 1-year time horizon, while at longer horizons they diverge, with more substantial positive and negative term premia emerging. Estimated term premia peak around 1994–1995, a period when inflation, inflation expectations and interest rates were all rising and economic growth was somewhat volatile. Term premia then fell steadily until around 1998, before increasing again over the next year or two. The period 1999–2000 saw a marked slowdown in global economic growth, a rise in inflation and bond yields, and a depreciation of the Australian dollar from around US 65 cents to around US 50 cents. From around 2001 to the end of the sample, term premia have been fairly steady.

Figure 2: Decomposition of the Forward Rate

Model estimated using 1993–2007 data sample

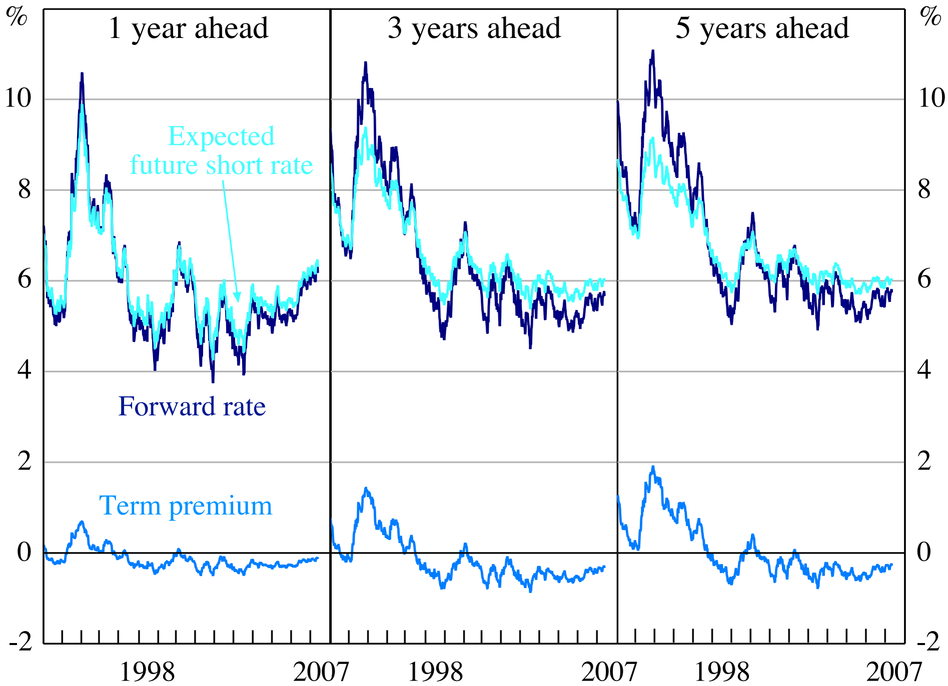


Figure 2 shows that expected future short rates are more stable than the forward rate, so that term premia tend to increase when yields are rising and decrease when yields are falling, especially for the longer-horizon samples. This result sits well with economic intuition. It seems obvious that market participants' view of the overnight cash rate a few years hence should be relatively stable – bullish macroeconomic news that would perhaps imply an increased chance of a near-term monetary tightening may raise the market's forecast of short-term interest rates, but its effect on expected interest rates years into the future is likely to be relatively small. Despite this, yields on long-term government bonds do react to such news, and to a greater extent than may seem warranted purely by changing forecasts of the real economy. As such, variations in long-term bond yields seem to be partly driven by factors other than expected future short rates, and these other factors show up as term premia in our model.

An alternate explanation raised in the literature is that this phenomenon is model driven, since the amount of variation seen in expectations of the future short rate far into the future can be significantly affected by the K matrix. In particular, large diagonal entries in the K matrix imply fast mean reversion of latent factors. This in turn means that expected future short rates will revert to the long-run value, ρ, very quickly, so that while yields may move, expected future short rates will vary little. However, this does not seem to affect our results (see discussion of the K matrix below).

Another interesting consideration is the fall in term premia, particularly pronounced between 1996 and 1998, and its subsequent stabilisation. Again, this is to some extent a function of falling yields and more stable expected future short rates. On the other hand, notwithstanding the preceding discussion, there is nothing in the model which forces term premia, as opposed to expected future short rates, to fall. It is also interesting to note that the up-ticks in term premia associated with higher inflation become more muted as time progresses. For example, underlying inflation peaked at similar levels during the periods 1994–1995, 2000–2001 and 2005–2006, yet in each case the response from term premia was very different – in the first period term premia rose by roughly 2 percentage points, in the second period by half a percentage point, and in the third period not at all (1999 also saw inflation rising, and term premia up by around 1 percentage point, but inflation peaked at a lower level than in the other three cases). The fall in the general level of term premia, and the more muted response of term premia to economic shocks, coincides with a period of relatively stable inflation and economic growth, and may reflect growing credibility of the Reserve Bank's inflation target among financial market participants. These factors undoubtedly affected both yields and term premia, making it difficult to establish causality, but it is nonetheless clear that our model-implied term premia behave as we might expect over this period.

The sustained negative term premia in the latter part of the sample is an interesting result of the model. Studies for other countries have also found negative term premia, although not to the same extent as our results. It is possible that model misspecification could be partly responsible for the size of the negative term premia, with the model placing the short rate too high and therefore term premia too low (we return to this point below). Alternatively, Australia's relatively small supply of government bonds may have resulted in yields being bid down by risk-averse and mandate-constrained investors. Indeed, over the latter part of the sample, government bond yields consistently implied lower forward rates than those seen in analysts' forecasts of the cash rate. From late 2000 to the end of the sample, bond yields have, on average, implied 2-year forward rates which are around half a percentage point below those seen in analysts' forecasts. Bond-implied forward rates were briefly above analysts' forecasts of the cash rate in 2002, when positive term premia were last seen. Over 2006 and 2007, the differential (while negative) narrowed, which also accords with our model estimates that show term premia tending slowly towards zero over the period.

In general, by allowing a flexible specification of term premia we also allow negative term premia to arise; these are in any case relatively small and have occurred over a period for which it is not inconceivable that negative term premia of the magnitude estimated indeed existed. We also note that, given the historical variability seen in the actual cash rate over the past decade or so, the level of variability seen in our estimates of the expected future cash rate seems reasonable.

Table 1: Parameter Estimates

Model estimated using 1993–2007 data sample

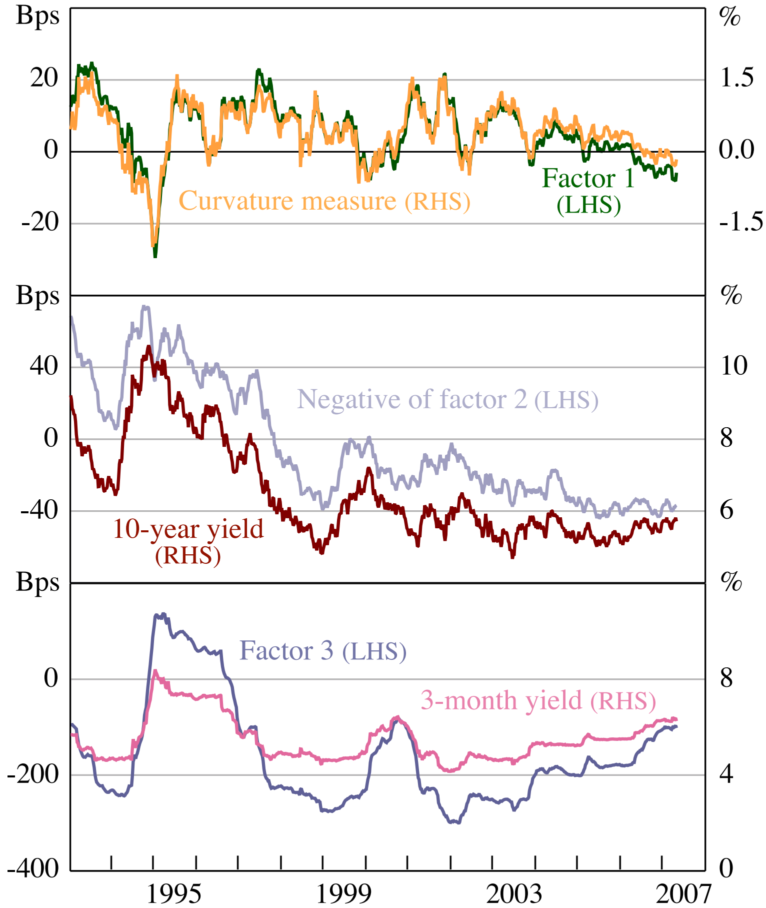
	Index number (i)					
Parameter	1		2		3	
ρ	6.97	(0.19)	–		–	
$(K)_{1i}$	1.81	(0.06)	0		0	
$(K)_{2i}$	−0.79	(0.04)	0.06	(0.00)	0	
$(K)_{3i}$	23.99	(0.50)	4.41	(0.22)	0.99	(0.02)
$(\Sigma)_{ii}$	0.15	(0.01)	0.20	(0.01)	0.42	(0.02)
$\lambda_{0,i}$	−0.11	(0.02)	0.19	(0.02)	−0.23	(0.05)
$(\Lambda)_{1i}$	−324.10	(1.03)	−43.17	(0.21)	−0.01	(0.00)
$(\Lambda)_{2i}$	17.29	(0.28)	15.75	(0.30)	3.87	(0.05)
$(\Lambda)_{3i}$	−68.11	(0.22)	82.39	(0.58)	−10.69	(0.38)

Notes: ρ and $(\Sigma)_{ii}$ are given in percentage points. Standard errors are shown in parentheses.

In Table 1 we present the estimated parameters of the model. Focusing first on the K matrix, consistent with Kim and Orphanides (2005), the smallest $(K)_{ii}$ estimated (in our case for the second latent factor) is quite small, at 0.06. In a single-factor model such a number would imply a half-life (being the time taken for a latent factor to revert half way back to its mean value of zero after experiencing a shock) of around 12.3 years (the other two half-lives would be 140 and 256 days). In term structure models such as ours, such a slow moving factor captures the longer-term characteristics of interest rates, which tend to follow the business cycle or other gradual trends, and may not mean revert for many years. Such small diagonal entries of K will improve estimation of expected future short rates at longer horizons. These estimated expectations would otherwise simply be flat and given by ρ.

Figure 3: Estimated Latent Factors

Model estimated using 1993–2007 data sample



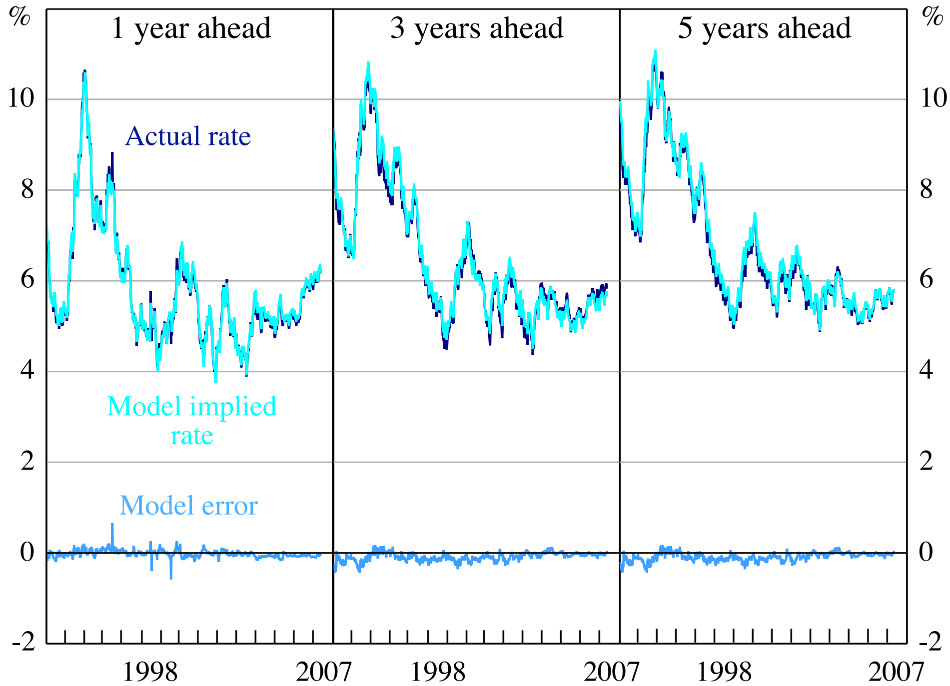
The latent factors are the processes that drive the evolution of interest rates. While they have no definitive economic interpretation, it is nevertheless interesting to examine how the model captures the yield curve with these three factors. Figure 3 displays the estimated latent factors of the model, along with various yield curve measures. The first latent factor is highly correlated with the curvature of the yield curve, where here we measure curvature as the 3-month yield plus the 10-year yield less twice the 2-year yield. The second latent factor exhibits a close relationship with long-term interest rates and consequently displays very slow mean reversion. This factor has a correlation coefficient of -0.96 with the 10-year yield. Finally, the third latent factor closely resembles short-term interest rates; it has a correlation coefficient of 0.98 with the 3-month yield.

Interpreting the term premia parameters is much more difficult as there are more of them and their effect on model outputs depends crucially on the sign and size of the $\mathbf{x}_t$ latent factors. Hence it is probably more intuitive to focus on the term premia produced (Figure 2) than on the actual numbers given in Table 1.

Finally the value of the constant ρ, which gives the short rate in steady-state and is estimated at 6.97 per cent, appears a little high. As mentioned earlier, it may be that the model is placing the long-run equilibrium short rate at a higher level than is warranted, thereby contributing to persistently negative term premia. The data sample does encompass the mid 1990s, formative years for the inflation-targeting regime and a period of relatively high cash rates which may have pushed up the estimate. Note also that our model gives the short rate at any time t as ρ plus the sum of the latent factors at t, and while the latent factors decay to zero, this happens very slowly. In fact over the sample period the value of $x_{1,t} + x_{2,t} + x_{3,t}$ has averaged -1.49 per cent. As the short rate is given by $\rho + x_{1,t} + x_{2,t} + x_{3,t}$, the equilibrium short rate over the sample could therefore be interpreted as being closer to 5½ per cent rather than to 7 per cent.

Actual and model-implied forward rates are shown in Figure 4. The fit is not perfect, but at any time the dynamic term structure model has much less flexibility in generating yields along the curve than the model used to estimate the zero-coupon yields – it must rely on the three latent factor values to generate an entire yield curve. Hence we would not expect the model to be perfect. Rather, we trust that it broadly characterises the observed actual yields, which appears to be the case – model errors are generally small and hover around zero, only departing when yields rise or fall particularly quickly.

Figure 4: Forward Rates – Actual and Model
Model estimated using 1993–2007 data sample



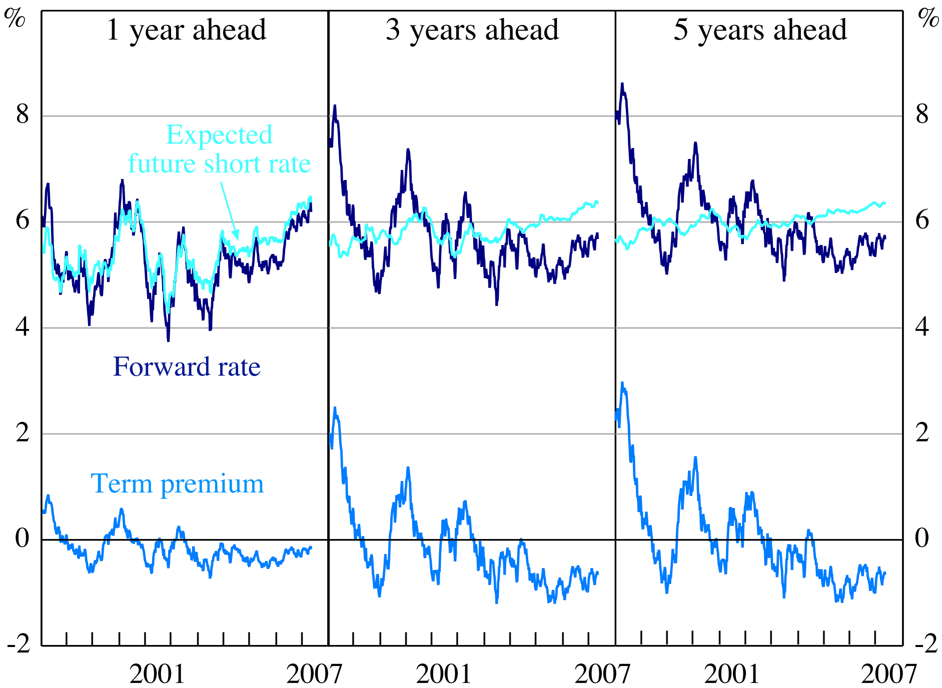
5.2 The Period 1997 to 2007

The data sample used in Section 5.1 spans the adoption by the Reserve Bank of the 2 to 3 per cent inflation target, the decline in inflation expectations, and the formal acknowledgement of Reserve Bank independence. As a result, there may be a structural break for which the model does not account. To check the robustness of the results to this possibility we estimate the model again using a restricted sample which encompasses the more stable period from 1997 to early 2007. Although the model estimates over this shorter sample are quite similar to those for the full sample, it is interesting to compare the two.

The estimated expected future short rates and term premia are given in Figure 5. The most obvious difference between Figure 5 and Figure 2 is the more stable expected future short rates and, consequently, the more variable term premia seen in the shorter sample. This is unlikely to be related to the K matrix since the shorter sample has one latent factor which is even slower to mean revert than in the case of the longer sample. The term premia parameters Λ and $\boldsymbol{\lambda}_0$ are numerically larger in the short sample model indicating, all else equal, that changes in yields will have a larger impact on term premia (and so a smaller impact on expected future short rates) than in the longer sample.

Figure 5: Decomposition of the Forward Rate

Model estimated using 1997–2007 data sample



In the shorter sample model there is an apparent upward trend in short rate expectations from around 2004 or earlier. Again it is hard to determine exactly what caused this, but it did occur during a time of rising cash rates, low but rising inflation, falling unemployment and stable growth. So for the 1-year ahead forecast at least, the trend seems plausible. The trend in the 5-year ahead forecast is smaller but still apparent. However, being of the order of less than one-quarter of a percentage point, it is probably not significant given the precision of the model .

Despite differences between the models one should keep in mind that the two are actually quite similar – the differences regard the degree of certain phenomena, not their existence. In fact, the estimates produced by the two models are generally within half a percentage point of each other; by comparison, Kim and Orphanides (2005), who also estimate models over two different sample periods, find differences in the order of around 2 percentage points.

The estimated parameters of the model over the short sample are given in Table 2. Similar to the longer sample, the smallest $(K)_{ii}$ estimated (in this case for the first latent factor) is small at 0.01. In a single-factor model such a number would imply a half-life of around 130 years (the other two half-lives would be 94 and 459 days). The extremely slow mean reversion of the first latent factor is likely due to the short sample period used, which spans a period of strong growth and low inflation, and in particular does not span a 'full' economic cycle encompassing a sizeable economic downturn.

Table 2: Parameter Estimates
Model estimated using 1997–2007 data sample

	Index number (i)					
Parameter	1		2		3	
ρ	6.78	(0.06)	–		–	
$(K)_{1i}$	0.01	(0.00)	0		0	
$(K)_{2i}$	0.55	(0.01)	2.69	(0.04)	0	
$(K)_{3i}$	2.88	(0.05)	17.38	(0.11)	0.55	(0.01)
$(\Sigma)_{ii}$	0.08	(0.00)	0.15	(0.01)	0.30	(0.01)
$\lambda_{0,i}$	1.45	(0.01)	−1.91	(0.03)	−1.06	(0.02)
$(\Lambda)_{1i}$	465.20	(0.39)	440.07	(0.26)	12.81	(0.09)
$(\Lambda)_{2i}$	−699.87	(0.44)	−763.06	(0.48)	1.16	(0.03)
$(\Lambda)_{3i}$	−171.60	(0.21)	−396.64	(0.28)	−109.58	(0.17)

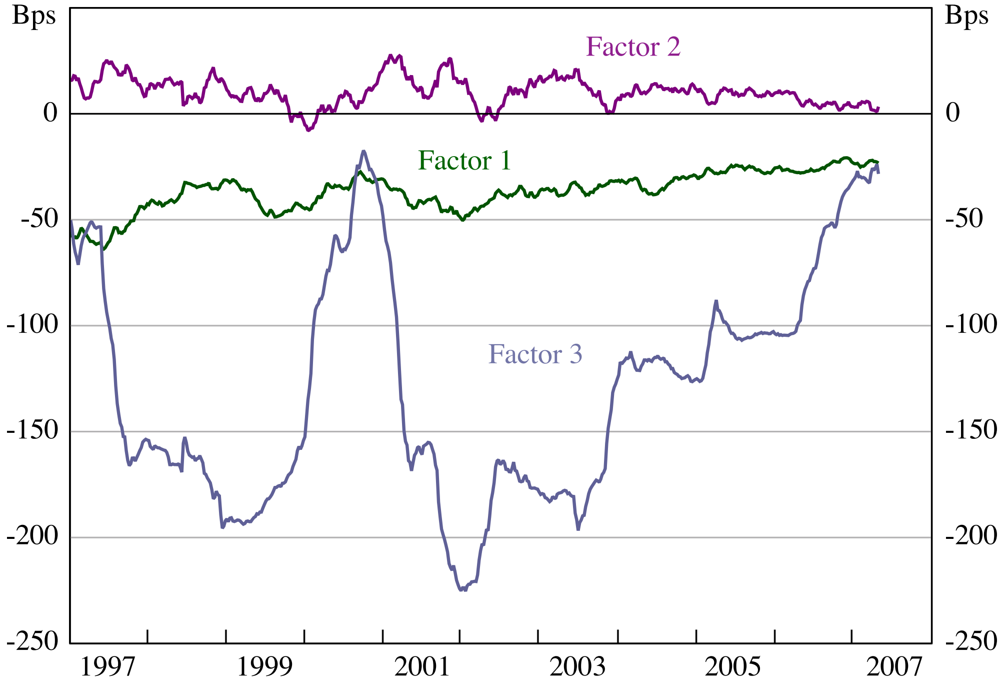
Notes: ρ and $(\Sigma)_{ii}$ are given in percentage points. Standard errors are shown in parentheses.

Figure 6 shows that the first latent factor exhibits slow mean reversion, and is in fact highly correlated with the slope of the yield curve (correlation coefficient of −0.97). In contrast, the much faster mean reverting third latent factor is highly correlated with the 3-month yield (correlation coefficient of 0.98), while the second latent factor is correlated with both the level and the curvature of yields.

It is interesting to note that for the shorter sample model, ρ is estimated at 6.78 per cent, below the 6.97 per cent estimated for the longer sample model. As touched on earlier, the credibility of the Reserve Bank's inflation target was being tested around 1994, which resulted in interest rates, and especially long-term rates, rising quite strongly. There may therefore be a structural change in interest rate dynamics around 1995–1996, associated with the Reserve Bank's success

Figure 6: Estimated Latent Factors

Model estimated using 1997-2007 data sample

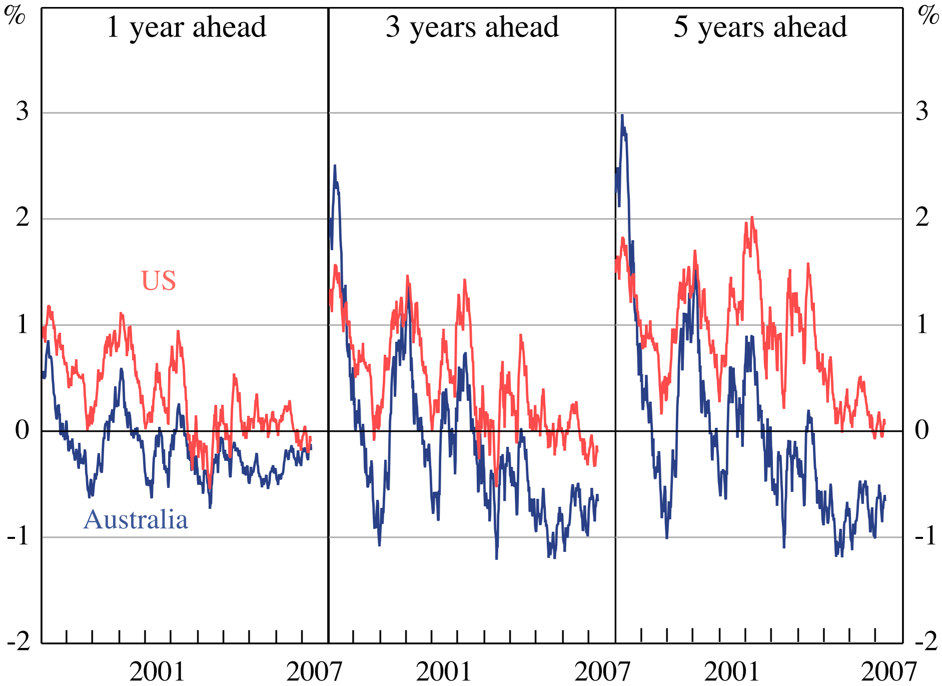


in reducing inflation and the moderation in inflation expectations, for which the model cannot account. The fact that the estimated equilibrium short rate is lower over the shorter sample is consistent with this. The lower value of ρ also manifests itself in less persistently negative term premia than those seen in the longer sample (note that similar to the longer sample model, the sum of latent factors averaged -1.56 per cent, giving an effective estimated short rate over the period of close to 5¼ per cent).

As a final point, it is interesting to compare our term premia estimates with those derived for US data. Figure 7 shows 1-, 3- and 5-year term premia as estimated by us for Australian data, as well as corresponding term premia estimated by Kim and Wright (2005), who also use the Kim and Orphanides model, for US data.[16] One can see that, excepting a level difference between the two series, the estimated Australian and US term premia track each other relatively closely – the correlation coefficient between the two series is 0.84 for the 1-year ahead term premia, 0.90 for the 3-year ahead term premia, and 0.81 for the 5-year ahead term premia. These

16 Available from http://www.federalreserve.gov/Pubs/feds/2005/200533/feds200533.xls.

Figure 7: Estimated Term Premia



results suggest that term premia may be driven by global, as opposed to country-specific, phenomena, as typified for example by the global 'search for yield' that received so much attention earlier this decade.

6. Conclusion

We have used data on coupon-bearing Australian Government bonds and OIS rates to estimate risk-free zero-coupon yield and forward curves for Australia from 1992 to 2007. These curves, and analysts' forecasts of future interest rates, were then used to fit an affine term structure model to Australian interest rates, with the aim of decomposing forward rates into expected future short rates and term premia.

The model produces plausible results, although given the complexity of the model and the difficulty of calibrating it to the data, a false level of precision should not be attributed to the results. The results show a large and sustained fall in term premia from around 1996 to 2007, when inflation credibility became more entrenched and so inflation expectations declined. This period displays relatively low inflation, stable economic growth and stable bond yields.

The model suggests that there have been small negative term premia for some periods. The finding of negative term premia has been a global phenomenon during the early to mid 2000s, as seen in the decline in long-term interest rates. This could reflect the widely discussed 'search for yield' that occurred over this period, or may be explained by an over-shooting of bond yields. In Australia's case, the relatively low supply of government bonds, which has tended to fall over the period considered, may have contributed to negative term premia as risk-averse and mandate-constrained investors bid up the price of these bonds.

Notwithstanding the above discussion, the results seem to imply that expected future short rates are relatively stable in Australia. The results also imply that, based on expectations of future monetary policy, yields on government bonds have been lower than might be expected, with term premia attached to these bonds consequently being negative, at least towards the end of our sample period.

Appendix A: Zero-coupon Yields

We estimate zero-coupon yields using the Merrill Lynch Exponential Spline (MLES) methodology adapted from Li *et al* (2001). This technique appears to be very efficient, and produces a good fit for the input data (technical details are given later in this Appendix).

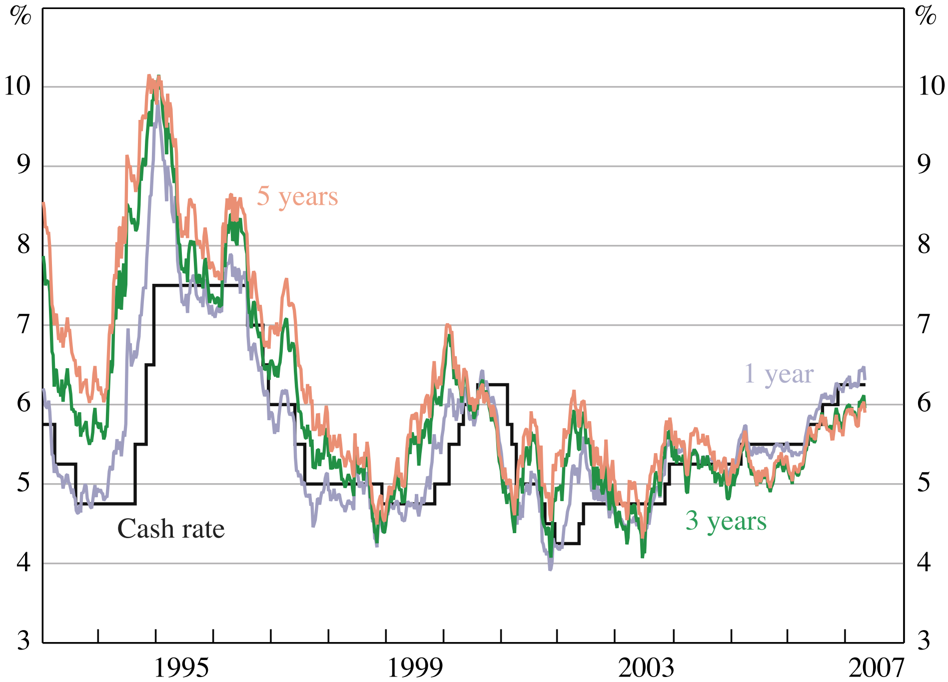
The hardest data to fit is for maturities around the 1-year mark, where from early 2001 the input data for any given day transitions from OIS rates (used as input data for maturities extending up to 1 year into the future) to bond yields (used for maturities greater than or equal to 18 months into the future). Although we regard OIS rates as the closest available substitute for risk-free Treasury note yields, they are not Treasury notes – they are swap contracts as opposed to physical bonds or notes, and they trade in a different market to physical bonds or notes, which may mean that the factors affecting OIS pricing are sometimes different from those affecting note or bond pricing. That being said, the MLES procedure still provides a good fit to the data even here – the average absolute error between the 1-year OIS yield and the MLES estimated 1-year yield is around 3½ basis points; the largest error is only 12 basis points and the error is larger than 10 basis points only three times. Taking a broader perspective, the fit of the MLES method is in fact very good, with the daily mean absolute error between fitted yields and actual yields averaging less than 2 basis points, and peaking at only 6 basis points.

Another potential and related criticism is the mixing together of OIS and bond yields to estimate a single yield curve. We believe that although Treasury note yields would be preferable for short maturity data inputs, in their absence, OIS rates are the next best, and in fact a very good, substitute. They are virtually risk-free and so can sensibly be used in the estimation of our risk-free yield curve, and they fulfil a vital function in supplying information about the short end of the yield curve that would otherwise be unavailable.

In any case, given that we are fitting a flexible term structure model, so long as the evolution of OIS yields through time is comparable to the evolution of Treasury note or bond yields, any residual risk premia inherent in OIS yields should be captured as short-dated term premia in the model. Overall, while Treasury note yields would be preferable, in their absence OIS yields provide a very good substitute.

We display estimated 1-, 3- and 5-year zero-coupon yields, as well as the interbank overnight cash rate in Figure A1.

Figure A1: Zero-coupon Yields



The technical details of the MLES methodology are as follows. We model the theoretical discount function $d(t)$ as a linear combination of hyperbolic basis functions.[17] The discount function is assumed to be of the form

$$d(t) = \sum_{k=1}^{D} \lambda_k \frac{1}{1+k\alpha t} \tag{A1}$$

where D is the number of basis functions (in our case $D = 8$), and α is an exogenous parameter, taken as 5 per cent.

Once we have estimated the λ_k coefficients we have a smooth discount function. From this it is a simple matter to compute the zero-coupon yield curve, given by

$$z(t) = -\frac{\log d(t)}{t}$$

[17] The discount function $d(t)$ gives the value today of 1 unit at time t in the future.

and the (instantaneous) forward curve given by

$$f(t) = z(t) + tz'(t).$$

Given the assumed form of the discount function, the theoretical price of bond i is given by

$$\hat{B}_i = \sum_{j=1}^{m_i} c_{ij} d(\tau_{ij})$$

where c_{ij} is the jth cash flow of bond i, occurring at time τ_{ij}, and m_i is the number of cash flows belonging to bond i. That is, the price of bond i is the sum of its discounted cash flows.

The price of a bond is the linear sum of its discounted cash flows. The discount function is assumed to be a linear sum of basis functions. This linearity allows us to write the vector of bond prices or OIS rates $\mathbf{B}$ as

$$\mathbf{B} = X\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{A2}$$

where $\mathbf{B}^T = [B_1, \cdots, B_n]$ is the vector of observed prices,[18] X is a $n \times D$ matrix with $X_{ik} = \sum_{j=1}^{m_i} c_{ij} \frac{1}{1+k\alpha\tau_{ij}}$, $\boldsymbol{\beta} = (\lambda_1, \cdots, \lambda_D)^T$ and $\boldsymbol{\varepsilon}$ a vector of errors.

For W the weight matrix,[19] if we wished to minimise the weighted squared pricing errors $\boldsymbol{\varepsilon}^T W \boldsymbol{\varepsilon}$, then the solution would be given by

$$\hat{\boldsymbol{\beta}} = (X^T W X)^{-1} X^T W \mathbf{B}. \tag{A3}$$

We wish to impose some restrictions on $d(t)$ however: $d(0)$, the discount rate at $t = 0$ should be 1, that is, one dollar today is worth one dollar. Also, the cash rate (as of today) is known and fixed, and so should be reflected in the discount function. These requirements complicate matters slightly.

From Equation (A1) it is clear that $d(0) = \sum_{k=1}^{D} \lambda_k$. Hence requiring $d(0) = 1$ is equivalent to requiring $\lambda_D = 1 - \sum_{k=1}^{D-1} \lambda_k$.

---

18 For OIS contracts the 'observed price' corresponds to the price of a discount security paying the OIS yield.

19 The weight attached to each bond is taken as its inverse duration. This has the effect of minimising fitted yield errors, as opposed to price errors.

Writing $f_k(t) = \frac{1}{1+k\alpha t}$, we can ensure that the 1-day yield is given by the overnight cash rate, r say, by requiring

$$d(\frac{1}{365}) = \sum_{k=1}^{D} \lambda_k f_k(\frac{1}{365}) = \frac{1}{1 + \frac{1}{365}\frac{r}{100}}. \tag{A4}$$

Writing $y = (1 + \frac{1}{365}\frac{r}{100})^{-1}$, from Equation (A4) these two constraints are equivalent to

$$\begin{aligned} y &= \lambda_1 f_1(\frac{1}{365}) + \cdots + \lambda_{D-1} f_{D-1}(\frac{1}{365}) + (1 - \sum_{k=1}^{D-1} \lambda_k) f_D(\frac{1}{365}) \\ &= \lambda_1 (f_1(\frac{1}{365}) - f_D(\frac{1}{365})) + \cdots + \lambda_{D-1}(f_{D-1}(\frac{1}{365}) - f_D(\frac{1}{365})) + f_D(\frac{1}{365}) \end{aligned}$$

and hence

$$\begin{aligned} \lambda_{D-1} &= \frac{y - f_D(\frac{1}{365})}{f_{D-1}(\frac{1}{365}) - f_D(\frac{1}{365})} - (\lambda_1 \frac{f_1(\frac{1}{365}) - f_D(\frac{1}{365})}{f_{D-1}(\frac{1}{365}) - f_D(\frac{1}{365})} + \cdots \\ &\qquad + \lambda_{D-2} \frac{f_{D-2}(\frac{1}{365}) - f_D(\frac{1}{365})}{f_{D-1}(\frac{1}{365}) - f_D(\frac{1}{365})}) \\ &= \lambda_{D-1}^* \end{aligned} \tag{A5}$$

We first impose the $d(0) = 1$ constraint. Writing $\mathbf{x_i}$ for the ith column of the X matrix, Equation (A2) becomes

$$\begin{aligned} \mathbf{B} &= \mathbf{x_1}\lambda_1 + \cdots + \mathbf{x_{D-1}}\lambda_{D-1} + \mathbf{x_D}\lambda_D + \boldsymbol{\varepsilon} \\ &= \mathbf{x_1}\lambda_1 + \cdots + \mathbf{x_{D-1}}\lambda_{D-1} + \mathbf{x_D}(1 - \sum_{k=1}^{D-1} \lambda_k) + \boldsymbol{\varepsilon} \end{aligned}$$

so that

$$\mathbf{B} - \mathbf{x_D} = (\mathbf{x_1} - \mathbf{x_D})\lambda_1 + \cdots + (\mathbf{x_{D-1}} - \mathbf{x_D})\lambda_{D-1} + \boldsymbol{\varepsilon}. \tag{A6}$$

Writing $\widehat{\mathbf{B}} = \mathbf{B} - \mathbf{x_D}$, $\widehat{\mathbf{x}}_i = \mathbf{x_i} - \mathbf{x_D}$, $\widehat{X} = (\widehat{\mathbf{x}}_1, \cdots, \widehat{\mathbf{x}}_{D-1})$ and $\widehat{\boldsymbol{\beta}} = (\lambda_1, \cdots, \lambda_{D-1})^T$, Equation (A6) becomes

$$\widehat{\mathbf{B}} = \widehat{X}\widehat{\boldsymbol{\beta}} + \boldsymbol{\varepsilon}. \tag{A7}$$

The estimate of $\widehat{\boldsymbol{\beta}}$ which minimises $\boldsymbol{\varepsilon}^T W \boldsymbol{\varepsilon}$ is $(\widehat{X}^T W \widehat{X})^{-1} \widehat{X}^T W \widehat{\mathbf{B}}$. Hence we have found the least squares estimate of $\boldsymbol{\beta}$ from Equation (A2), subject to the $d(0) = 1$ constraint.

If we now start from Equation (A7), replace λ_{D-1} with λ^*_{D-1} from Equation (A5) and follow the procedure above, we obtain our estimator. In this case the estimator of $\widetilde{\boldsymbol{\beta}} = (\lambda_1, \cdots, \lambda_{D-2})^T$ which solves

$$\widetilde{\mathbf{B}} = \widetilde{X} \widetilde{\boldsymbol{\beta}} + \boldsymbol{\varepsilon} \tag{A8}$$

for $\widetilde{\mathbf{B}} = \widehat{\mathbf{B}} - \frac{y - f_D(\frac{1}{365})}{f_{D-1}(\frac{1}{365}) - f_D(\frac{1}{365})} \widehat{\mathbf{x}}_{D-1}$ and $\widetilde{X} = (\widetilde{\mathbf{x}}_1, \cdots, \widetilde{\mathbf{x}}_{D-2})$ for $\widetilde{\mathbf{x}}_i = \widehat{\mathbf{x}}_i - \frac{f_i(\frac{1}{365}) - f_D(\frac{1}{365})}{f_{D-1}(\frac{1}{365}) - f_D(\frac{1}{365})} \widehat{\mathbf{x}}_{D-1}$ is given by

$$\widetilde{\boldsymbol{\beta}} = (\widetilde{X}^T W \widetilde{X})^{-1} \widetilde{X}^T W \widetilde{\mathbf{B}}. \tag{A9}$$

Hence we have solved Equation (A2) subject to both desired constraints.

Appendix B: Risk-neutral Bond Pricing

Here we examine why bonds should be priced under the risk-neutral measure. To simplify the analysis we work with a single factor model, that is

$$r_t = \rho + x_t \tag{B1}$$
$$\mathrm{d}x_t = -kx_t\mathrm{d}t + \sigma\mathrm{d}W_t \tag{B2}$$
$$\lambda_t = \lambda_0 + \lambda_1 x_t \tag{B3}$$

where all variables are scalars.

Consider the probability space $(\Omega, \mathscr{F}, \mathbb{P})$ with associated filtration $\mathscr{F}_t$ taken as the augmented filtration of $\sigma\{W_s | s \leq t\}$ (see, for example, Steele 2001). X_t is an Ito process if

$$\mathrm{d}X_t = \mu_x\mathrm{d}t + \sigma_x\mathrm{d}W_t$$

for μ_x and σ_x adapted to $\mathscr{F}_t$. Ito's lemma then states that for any function $F(x,t)$ such that F is twice differentiable in x and differentiable in t,

$$\mathrm{d}F = \left(\frac{\partial F}{\partial t} + \mu_x\frac{\partial F}{\partial x} + \frac{1}{2}\sigma_x^2\frac{\partial^2 F}{\partial x^2}\right)\mathrm{d}t + \sigma_x\frac{\partial F}{\partial x}\mathrm{d}W_t.$$

Applying Ito's lemma to Equation (B1) we trivially get

$$\begin{aligned}\mathrm{d}r_t &= -kx_t\mathrm{d}t + \sigma\mathrm{d}W_t \\ &\equiv \mu_r\mathrm{d}t + \sigma_r\mathrm{d}W_t.\end{aligned}$$

Now let $P^A(r_t,t)$ and $P^B(r_t,t)$ denote the time t price of two zero-coupon bonds with different maturity dates. Then by Ito's lemma, P^i (i = A, B) will satisfy

$$\mathrm{d}P^i = \left(\frac{\partial P^i}{\partial t} + \mu_r\frac{\partial P^i}{\partial r} + \frac{1}{2}\sigma_r^2\frac{\partial^2 P^i}{\partial r^2}\right)\mathrm{d}t + \sigma_r\frac{\partial P^i}{\partial r}\mathrm{d}W_t \tag{B4}$$
$$\equiv \mu^i\mathrm{d}t + \sigma^i\mathrm{d}W_t. \tag{B5}$$

Consider a portfolio that is long one A bond and short h B bonds. At time t this portfolio has value

$$V_t = P^A - hP^B. \tag{B6}$$

If held for dt, the portfolio's value changes by

$$\begin{aligned} \mathrm{d}V_t &= \mathrm{d}P^A - h\mathrm{d}P^B \\ &= (\mu^A - h\mu^B)\mathrm{d}t + (\sigma^A - h\sigma^B)\mathrm{d}W_t. \end{aligned} \tag{B7}$$

Hence we can make the portfolio instantaneously riskless by choosing $h = \sigma^A/\sigma^B$. In this case, the portfolio must earn the risk-free rate r_t and so

$$\mathrm{d}V_t = r_t V_t \mathrm{d}t. \tag{B8}$$

Substituting Equations (B6) and (B7) into Equation (B8) and setting $h = \sigma^A/\sigma^B$ leads to

$$\mu^A - \frac{\sigma^A}{\sigma^B}\mu^B = r_t(P^A - \frac{\sigma^A}{\sigma^B}P^B)$$

or

$$\frac{\mu^A - r_t P^A}{\sigma^A} = \frac{\mu^B - r_t P^B}{\sigma^B}.$$

Hence the ratio $(\mu - r_t P)/\sigma$ is independent of the choice of bond, and so there must exist a function λ_r such that

$$\frac{\mu - r_t P}{\sigma} = \lambda_r \tag{B9}$$

holds for any bond price P.

Now substituting μ and σ as identified by Equations (B4) and (B5) into Equation (B9) results in a Black-Scholes type partial differential equation

$$\frac{\partial P}{\partial t} = r_t P - (\mu_r - \lambda_r \sigma_r)\frac{\partial P}{\partial r} - \frac{1}{2}\sigma_r^2 \frac{\partial^2 P}{\partial r^2} \tag{B10}$$

which is solved subject to appropriate boundary conditions (bonds pay 1 unit at maturity) and the radiation condition $P \to 0$ as $r \to \infty$.

The Feynman-Kac formula then says that the solution to Equation (B10) is given by

$$P_{t,\tau} = \mathbb{E}_t^* \left[\exp(-\int_t^T r_s \mathrm{d}s) \right]$$

where r_t satisfies

$$\mathrm{d}r_t = (\mu_r - \lambda_r \sigma_r)\mathrm{d}t + \sigma_r \mathrm{d}W_t^*; \;\; \mathrm{d}W_t^* = \mathrm{d}W_t + \lambda_r \mathrm{d}t$$

and W_t^* is standard Brownian motion in the risk-neutral measure associated with $\mathbb{E}^*$.

Appendix C: Model Implementation

C.1 Formulas for a_τ and $\mathbf{b}_\tau$

From Kim and Orphanides (2005) we take the following formulas (with corrections):

$$
\begin{aligned}
a_\tau =& \frac{1}{\tau}\Big((K^*\boldsymbol{\mu}^*)'(M_{1,\tau}-\tau I)K^{*-1'}\mathbf{1} \\
&-\frac{1}{2}\mathbf{1}'K^{*-1}(M_{2,\tau}-\Sigma\Sigma' M_{1,\tau}-M_{1,\tau}'\Sigma\Sigma'+\tau\Sigma\Sigma')K^{*-1'}\mathbf{1}+\tau\rho\Big) \\
\mathbf{b}_\tau =& \frac{1}{\tau}M_{1,\tau}\mathbf{1}
\end{aligned}
$$

where

$$
\begin{aligned}
M_{1,\tau} &= -K^{*-1'}\left(e^{-K^{*'}\tau}-I\right) \\
M_{2,\tau} &= -vec^{-1}\left(((K^*\otimes I)+(I\otimes K^*))^{-1} vec\left(e^{-K^*\tau}\Sigma\Sigma' e^{-K^{*'}\tau}-\Sigma\Sigma'\right)\right)
\end{aligned}
$$

with $K^* = K+\Sigma\Lambda$, $\boldsymbol{\mu}^* = K^{*-1}\Sigma\boldsymbol{\lambda}_0$, *vec* taking a matrix to a vector column-wise, and vec^{-1} doing the opposite.

C.2 The Kalman Filter

Our implementation of the Kalman filter is based on that used by Duffee and Stanton (2004). The recursion goes from $t = 1$ forward, and is as follows:

1. Using the current value of $\mathbf{x}_t$, compute the one-step-ahead forecast of $\mathbf{x}_t$, given by $\mathbf{x}_{t+1|t} = e^{-Kh}\mathbf{x}_t$, and its variance matrix $P_{t+1|t} = e^{-Kh}P_{t|t}(e^{-Kh})' + \Omega_h$.

2. Compute the one-step-ahead forecast of $\mathbf{y}_t$, given by $\mathbf{y}_{t+1|t} = \mathbf{a} + B\mathbf{x}_{t+1|t}$, and its variance matrix $V_{t+1|t} = BP_{t+1|t}B' + R$, where R is the zero-coupon bond measurement error variance matrix.

3. Compute the forecast errors of $\mathbf{y}_{t+1|t}$, given by $\mathbf{e}_{t+1} = \mathbf{y}_{t+1} - \mathbf{y}_{t+1|t}$.

4. Update the prediction of $\mathbf{x}_{t+1}$ with $\mathbf{x}_{t+1|t+1} = \mathbf{x}_{t+1|t} + P_{t+1|t}B'V_{t+1|t}^{-1}\mathbf{e}_{t+1}$ and the variance with $P_{t+1|t+1} = P_{t+1|t} - P_{t+1|t}B'V_{t+1|t}^{-1}BP_{t+1|t}$.

For times t when we have analysts' forecasts, replace $\mathbf{y}_t$, $\mathbf{a}$, B and R by $\tilde{\mathbf{y}}_t$, $\tilde{\mathbf{a}}$, $\tilde{B}$ and $\tilde{R}$, respectively.

We then choose our parameter vector Θ to solve

$$\Theta^* = \overset{\text{argmax}}{\Theta} \sum_{t=1}^{n} \mathrm{LL}(\mathbf{e}_t, V_{t|t-1})$$

where the sample is n periods long, and the period-t approximate log-likelihood is given by

$$\mathrm{LL}(\mathbf{e}_t, V_{t|t-1}) = -\frac{1}{2}\left(\log|V_{t|t-1}| + \mathbf{e_t}' V_{t|t-1}^{-1} \mathbf{e_t}\right).$$

C.3 Implementation

To implement the model we restrict the parameters to those which result in K and K^* having positive eigenvalues. This results in $e^{-Ks} \to 0$ and $e^{-K^*s} \to 0$ as $s \to \infty$, which ensures the stability of the model (see Equation (9) and the formulas for a_τ and $\mathbf{b}_\tau$). We also require that the σ_i, being variances, are positive.

By way of parameter choices that must be made, we set $\mathbf{x}_1 = [0.005,\ 0.03,\ 0.01]'$ (with an initial standard deviation of 10 per cent) and then discard the first six months of the estimation. The standard deviation of zero-coupon yield measurement errors is set to 10 basis points, while those of the survey forecasts are set to 50 basis points per square root year.

Finally, the standard errors in Tables 1 and 2 are calculated using a random walk chain Metropolis-Hastings algorithm – for details see Geweke (1992).

References

Beechey M (2004), ‘Excess Sensitivity and Volatility of Long Interest Rates: The Role of Limited Information in Bond Markets’, Sveriges Riksbank Working Paper No 173.

Bolder DJ and S Gusba (2002), ‘Exponentials, Polynomials, and Fourier Series: More Yield Curve Modelling at the Bank of Canada’, Bank of Canada Working Paper No 02-29.

Cochrane JH (2001), *Asset Pricing*, Princeton University Press, Princeton.

Dai Q and KJ Singleton (2002), ‘Expectation Puzzles, Time-Varying Risk Premia, and Affine Models of the Term Structure’, *Journal of Financial Economics*, 63(3), pp 415–441.

de Jong F (2000), ‘Time Series and Cross-Section Information in Affine Term-Structure Models’, *Journal of Business & Economic Statistics*, 18(3), pp 300–314.

Duffee GR (2002), ‘Term Premia and Interest Rate Forecasts in Affine Models’, *Journal of Finance*, 57(1), pp 405–443.

Duffee GR and RH Stanton (2004), ‘Estimation of Dynamic Term Structure Models’, unpublished, Hass School of Business, UC Berkeley.

Duffie D and R Kan (1996), ‘A Yield-Factor Model of Interest Rates’, *Mathematical Finance*, 6(4), pp 379–406.

Geweke J (1992), ‘Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments’, in JO Berger, JM Bernardo, AP Dawid and AFM Smith (eds), *Proceedings of the Fourth Valencia International Meeting on Bayesian Statistics*, Oxford University Press, Oxford, pp 169–194.

Gürkaynak RS, BP Sack and ET Swanson (2003), ‘The Excess Sensitivity of Long-Term Interest Rates: Evidence and Implications for Macroeconomic Models’, Federal Reserve Board Finance and Economics Discussion Series No 2003-50.

Kim DH and A Orphanides (2005), ‘Term Structure Estimation with Survey Data on Interest Rate Forecasts’, Federal Reserve Board Finance and Economics Discussion Series No 2005-48.

Kim DH and JH Wright (2005), 'An Arbitrage-Free Three-Factor Term Structure Model and the Recent Behavior of Long-Term Yields and Distant-Horizon Forward Rates', Federal Reserve Board Finance and Economics Discussion Series No 2005-33.

Kremer M and M Rostagno (2006), 'The Long-Term Interest Rate Conundrum in the Euro Area: Potential Explanations and Implications for Monetary Policy', paper presented at the BIS Autumn Meeting of Central Bank Economists on 'Understanding Asset Prices: Determinants and Policy Implications', Basel, 30–31 October.

Li B, E DeWetering, G Lucas, R Brenner and A Shapiro (2001), 'Merrill Lynch Exponential Spline Model', Merrill Lynch Working Paper.

Reserve Bank of Australia (RBA) (2002), 'Overnight Indexed Swap Rates', Reserve Bank of Australia *Bulletin*, pp 22–25.

Rudebusch GD, BP Sack and ET Swanson (2007), 'Macroeconomic Implications of Changes in the Term Premium', Federal Reserve Bank of St. Louis *Review*, 89(4), pp 241–269.

Rudebusch GD and T Wu (2008), 'A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy', *Economic Journal*, 118(530), pp 906–926.

Steele JM (2001), *Stochastic Calculus and Financial Applications*, Springer, New York.

Swanson ET (2007), 'What We Do and Don't Know about the Term Premium', Federal Reserve Board of San Francisco *Economic Letter* No 2007-21.

Westaway P (2006), 'Extracting Information from Financial Markets at the Bank of England: Refining the Modelling Toolkit and Strengthening the Theoretical Underpinnings', paper presented at the BIS Autumn Meeting of Central Bank Economists on 'Understanding Asset Prices: Determinants and Policy Implications', Basel, 30–31 October.