



SCHOOL OF FINANCE AND ECONOMICS

UTS:BUSINESS

WORKING PAPER NO. 81

NOVEMBER, 1998

Zero-Coupon Yield Curve Estimation: A Principal Component, Polynomial Approach

Ben Hunt
Chris Terry

ISSN: 1036-7373

<http://www.business.uts.edu.au/finance/>

**Zero-Coupon Yield Curve Estimation; a Principal Component,
Polynomial Approach**

by

B F Hunt & C S Terry

Ben Hunt is Associate Professor and Head, Graduate School of the Faculty of Business, and Chris Terry is Associate Professor and Head, School of Finance & Economics, University of Technology, Sydney

Correspondence to

*Ben Hunt,
Graduate School of Business,
University of Technology, Sydney,
PO Box 123 Broadway,
N.S.W. 2007,
Australia
Fax 612.95143554
Ben.Hunt@uts.edu.au*

Abstract

Polynomial functions of the term to maturity have long been used to provide a general functional form for zero-coupon yield curves. The polynomial form has many advantages over alternative functional forms such as Laguerre, when using non-linear least squares to estimate zero-coupon yield curves with coupon bond data. Most importantly the polynomial form invariably enables convergence of the estimation process.

Unfortunately, the simple polynomial form results in estimated models of zero-coupon yield curves that approach either plus or minus infinity as the term increases. This unsatisfactory aspect of the simple polynomial is inconsistent with both theoretical considerations and observational reality.

We propose a new zero-coupon yield curve functional form consisting not of simple polynomials of term, τ , but rather constructed from polynomials of $1/(1+\tau)$. This form has the desirable property that long-term yields approach a constant value. Further, we model zero-coupon yields as a linear function of the first k principal components of p polynomials of $1/(1+\tau)$, $k < p$. Using a small number of principal components produces zero-coupon yield curves with a simple "parsimonious" structure while retaining the flexibility of larger polynomial form with its $k-1$ potential turning points. Moreover estimating with principal components exploits the principal explanatory variable's lack of colinearity.

The principal components of polynomials of $1/(1+\tau)$ model is applied to Australian coupon bond data. The results compare favourably to those obtained using the traditional polynomial of term model.

Zero-coupon Yield Curve Estimation; a Principal Component, Polynomial Approach

Introduction

An understanding of the zero-coupon treasury yield curve is essential to the pricing of an increasingly vast array of interest bearing securities and the derivatives of those securities. Zero-coupon rates for example are central to the determination of “fair” coupon bond prices, forward rates, futures prices and swap rates. Zero-coupon bond prices and yields can be inferred from a liquid market for zero-coupon bonds or treasury strips. However there are very few such markets, and thus zero-coupon yield curves are generally inferred from the market yields/prices for coupon paying bonds. This is not a trivial task.

This paper proposes a new and relatively simple method for deriving a zero-coupon yield curve from the market prices of coupon paying treasury bonds. It models zero-coupon yields, y^z , as a function of the principal components of polynomials of $1/(1+\tau)$, where τ is the term of the underlying zero-coupon instrument. This functional form not only has the flexibility, generality and tractable estimation properties of all polynomial forms, but also, and most importantly, the zero-coupon yields, implied by the model, approach some long-term limit as the term increases.

Fitting zero-coupon yield curves

A coupon bond may be priced a number of ways. The traditional procedure is to discount all of the bond's cash flows at the market determined *yield to maturity*. That is,

$$P(y^m, t) = \frac{c_1}{1 + y^m} + \frac{c_2}{(1 + y^m)^2} + \dots + \frac{c_t + F}{(1 + y^m)^t} \quad (1)$$

Where $P(y^m, t)$ is the price of a t period bond when the yield to maturity is y^m ,

c_i is the coupon payment at time t , and

F is the bond's face value.

A coupon bond is, in effect, a bundle of zero-coupon bonds with each coupon payment constituting a single zero-coupon bond. An alternative pricing method uses the constituent zero-coupon rates. A bond's cash flows are discounted with the relevant zero-coupon rate, rather than the yield to maturity, provided by the prevailing zero-coupon yield curve. Here,

$$P(y^z(\tau), t) = \frac{c_1}{1 + y^z(1)} + \frac{c_2}{(1 + y^z(2))^2} + \dots + \frac{c_t + F}{(1 + y^z(t))^t} \quad (2)$$

Where $y^z(i)$ is the zero coupon rate applicable to a term of i periods.

The zero-coupon yield curve, $y^z(\tau)$, may be estimated as a particular function of the term, τ , that minimises the sum of the squared differences between the actual market bond prices, as given in (1), and the zero-coupon bond prices as given in (2). That is, the zero-coupon yield curve is estimated by minimising,

$$SS = \sum_{i=1}^n \left(P(y^m, t_i) - P(y^z(\tau), t_i) \right)^2 \quad (3)$$

Where n is the number of coupon bonds used in the estimation process.

A consensus as to the “best” form for determining the zero coupon curve is yet to be established. One popular approach has been to use a series of splines to approximate $y^z(t)$. For example, Vasicek and Fong [1982] use exponential splines, Frishling et al [1994] use a linear spline, while Adams and Deventer [1994] a fourth order spline with the cubic term missing. Spline zero curves perfectly fit the data ensuring that the sum of squares term (equation (3)) is zero. While the perfect fit property of splines may, at first appear to be advantageous, it does contain faults. A zero-yield curve that perfectly fits the data does not allow for either data measurement error or bond pricing errors.

Data measurement error can arise from a number of sources, such as thin trading and non-synchronous trading. When there is very little trading one party may be able to impose a non-fair price upon another party. The problem of non-synchronous trading arises where the bond prices used in the estimation process are generated at significantly different points in time. Here, the bond prices represent a mixture of term structures taken from different points in time. The use of a spline function on non-fair or non-synchronous data may produce a yield curve with an erroneous idiosyncratic shape.

Spline curve models completely eliminate prediction error through the implicit estimation of very many parameters (one for each bond). In contrast to spline estimation, with its many parameters, is the “parsimonious” approach that employs a small number of parameters. In this approach to yield curve estimation the bond price data is “smoothed” by an estimation technique that minimises, but not eliminate, predicted bond price errors. This smoothing approach produces superior results when estimating zero-coupon yield curves from bond data containing errors.

Under the parsimony principle, empirical models are relatively simple, with a parsimoniously small number of parameters. Previous research has shown that while parsimonious models are relatively simple structures they nevertheless fit the observed data smoothly and well. The parsimonious approach to yield curve estimation was pioneered by Nelson and Siegal [1987]. They specified the yield curve as a four-parameter Laguerre function:

$$y^z(\tau) = \beta_0 + (\beta_1 + \beta_2\tau) e^{\beta_3\tau} \quad (4)$$

The Laguerre yield curve has many beneficial aspects. Its flexibility allows for a number of yield curve shapes such as, trenches and humps. It also produces sensible rates at the extremities of the term structure, $y(0) = \beta_0 + \beta_1$ and $y(\infty) = \beta_0$. Notwithstanding its many attractive properties, unfortunately the Laguerre form is notoriously difficult to apply. For many yield curve shapes, the non-linear least squares estimation process necessary to estimate (4) fails to converge for the Laguerre form. Bhar and Hunt [1994], Barrett et al [1995], Hunt [1995a] and Pham [1995] illuminate the problems of implementing Laguerre zero-coupon yield curves. An alternative to the Laguerre form is the polynomial yield curve.

For example, a four-parameter polynomial yield curve is specified as:

$$y^z(\tau) = \beta_0 + \beta_1\tau + \beta_2\tau^2 + \beta_3\tau^3 \quad (5)$$

Polynomial zero-coupon yield models invariably produce convergence in the estimation process. Studies that have employed polynomial forms are Langeteig and Smile [1989] Bhar and Hunt [1994], Hunt [1995a], Hunt [1995b] and Pham [1995]. This form is capable of providing all commonly observed yield curve shapes. The polynomial form is a general function that provides a good approximation to any yield curve function, as the Taylor series expansion shows. The primary failing of the polynomial form concerns the rates at the long end of the term structure.

It is generally accepted that interest rates are mean reverting and a raft of interest rate models, starting with Vasicek [1977], incorporate mean reversion. Mean reversion dictates that interest rates tend toward a fixed "mean" as the term increases. Unfortunately, however, the simple polynomials of term model are not consistent with yields tending to some long-term constant rate. On the contrary, it is a property of the polynomial form that long-term rates must diverge towards either positive or negative infinity. The instability of longer rates is a most unsatisfactory aspect of the simple polynomial term structure.

We suggest a polynomial of a simple transformation of the term, τ , that removes the problem of long-run rate instability inherent in this simple polynomial model (5) while retaining the tractable estimation property of the polynomial form.

Polynomials of $1/(1+\tau)$

Polynomials of $1/(1+\tau)$ are well behaved in that they approach a constant as τ increases.¹ We propose the following zero-coupon yield curve;

$$\begin{aligned}
 y^z(\tau) &= \beta_0 + \sum_{i=1}^p \beta_i \frac{1}{(1+\tau)^i} \\
 &= \beta_0 + \sum_{i=1}^{p-1} \beta_i w^i \\
 &= \mathbf{X}(\tau)\boldsymbol{\beta}
 \end{aligned} \tag{6}$$

Where $\boldsymbol{\beta}$ is a $(p \times 1)$ vector of coefficients $(\beta_0, \beta_1, \dots, \beta_{p-1})$,

\mathbf{X} is a $(1 \times p)$ vector of polynomials $(w^0, w^1, \dots, w^{p-1})$ and

$w = 1/(1+\tau)$.

The polynomial in $1/(1+\tau)$, like the polynomial in t , is flexible and capable of rendering all of the common yield curve shapes but unlike the polynomial in τ , is well behaved at both ends of the term structure, such that:

$$y^z(0) = \sum_{i=0}^{p-1} \beta_i$$

$$y^z(\infty) = \beta_0$$

A parsimonious modelling approach requires the yield curve to be specified with the smallest number of parameters that provide an acceptable fit to the observed bond data. Previous studies have found that models with three and four parameters are sufficient to provide a good fit to the data (see for example Nelson and Siegal [1987], Diament [1993], Hall [1989], Bhar and Hunt [1993] and Hunt [1995a, 1995b]). Following this tradition one would restrict p in equation (6) to be three or four. However, even if one, operating under the principle of parsimony, decided to estimate say, four parameters, it does not follow

that one would use the first four polynomial terms. Rather, one would employ four polynomial terms that provided the best fit. For example, Adams and van Deventer [1994] and Hunt [1995a] found that, within a four parameter model, a four order polynomial model with the cubic term omitted outperforms a straight third order polynomial model.

We suggest a new principal components approach to selecting a k parameter model from a possible p polynomial terms ($k < p$). This approach employs principal components of polynomials of $1/(1+\tau)$. Specifically, our approach uses the "best" k principal components extracted from p , polynomials of $1/(1+k)$. Equation (6) is transformed to:

$$y^z(\tau) = \mathbf{Z}(\tau)\boldsymbol{\alpha} \quad (7)$$

Where, $\boldsymbol{\alpha}$ is a $(k*1)$ vector of coefficients $(\alpha_0, \alpha_1, \dots, \alpha_k)$,
 \mathbf{Z} is a $(1*k)$ vector of functions of τ ,

$$\mathbf{Z}(\tau) = \mathbf{X}(\tau) \mathbf{A} \quad (8)$$

Where \mathbf{A} is a $k*p$ matrix transformation coefficients. Let \mathbf{Z} and \mathbf{X} be matrices composed of rows of $\mathbf{Z}(\tau)$ and $\mathbf{X}(\tau)$ for $\tau = 1$ to T . The matrix \mathbf{A} , is composed of the k characteristic roots of $\mathbf{X}'\mathbf{X}$ associated with the largest k characteristic roots of $\mathbf{X}'\mathbf{X}$. Thus, the matrix \mathbf{Z} contains k principal components of \mathbf{X} . The principal components consist of orthogonal vectors with the same number of observations as the original series that explain the maximum possible variance of the original series. If $k = p$, all of the variance of the original series is explained and $\boldsymbol{\beta} = \mathbf{A}^{-1} \boldsymbol{\alpha}$.

While the untransformed polynomials $1/(1+\tau)$ are naturally correlated, the principal components are not. A comparison of the untransformed first three polynomials of $1/(1+\tau)$, $\tau = 1$ to 10 and the first three principal components of the polynomials is presented in Table 1 and Figure 1. Table 1 shows the high degree of correlation between the first three polynomials of $1/(1+\tau)$.² This correlation is completely eliminated by combining these variables into principal components.

Figure 1 contrasts the shape of the first three polynomial variables with those of the three principal components of those variables. Note the untransformed variables do not have any minimum/maximum points, however, two of the principal component series contain turning points.

Principal component series are orthogonal to each other. In linear regression analysis orthogonality is an advantage as variables are “independent” in that they can be added to, or subtracted from, a specification without altering the value of the least squares estimated parameters on the other included variables. This property is beneficial in one’s attempt to select an additional explanatory variable from a number of candidate variables. One simply progresses through the candidate variables adding the next orthogonal variable that achieves the greatest reduction in the error sum of squares. This cannot be done with correlated variables. The reduction in error variance depends not upon the nature of the additional variable but on the combination of that variable with all of the other included variables.

Equation (7) is linear in the variables; equation (3) introduces non-linearity into the estimation process. While the desirable independence property of orthogonal variables does not strictly carry across to linear specifications, our experience indicates that mostly it does. In most cases the best combination of k explanatory variables is the best

combination of $k-1$ explanatory, plus that variable which provides the greatest reduction in error sum of squares by itself. Principal components analysis is of great assistance in identifying the ability of variable to reduce the residual error of a specification as principal components are ranked by their variance. The large variance principal components most often provide the greatest reduction in error variance.³

Example

We selected a distinct normal yield curve to demonstrate the derivation of the implicit zero-coupon yield curve. It is based on the market yields on the 7th August 1995 for a set of traded Australian Treasury bonds. The data consisted of yields and specifications for 19 coupon-paying bonds. The term of the bonds ranged from about 18 months to 12 years. The coupon rate ranged from 6.25%p.a. to 13%p.a. These yields are shown in Figure 2.

Table 2 sets out the results of sequentially adding principal components of the first eight polynomial terms of $1/(1+\tau)$. The second column in Table 2 displays how the addition of each of the first four variables provides considerable reduction in the bond price residual sum of squared residuals. The data in the third column of Table 2 show that the addition of the fourth principal component signals the end of this tendency as it not significantly different from zero.

The estimated zero-coupon yield curve, using the first three principal components of the first eight polynomial terms of $1/(1+\tau)$, is plotted in Figure 3 along with the zero-coupon yield curve estimated using the more traditional polynomial specification using the first three polynomial terms of τ .

The two zero-coupon yield curves have a similar shape over sample term. Each model has an almost identical performance in explaining bond price variation. The third order polynomial zero-coupon yield curve has a residual sum of squares of 0.284 while the three principal components zero-coupon yield curve model has a residual sum of squares of 0.281. Both figures equate to an R^2 of 99.98%. The principal components model outperforms the polynomial model in the projection of longer-term rates.

Figure 3 clearly shows the tendency for straight polynomial rates to tend towards infinity ($-\infty$ in this case). In contrast the principal components of $1/(1+\tau)$ asymptotes toward the coefficient on the constant, 10.34%p.a. in this case. If reasonable implied long-term rates are important then the principal components zero-coupon model is clearly superior to its straight polynomial equivalent.

Polynomial functions of the term to maturity have long been used to provide a general functional form for zero-coupon yield curves. The polynomial form has many advantages, over alternative functional forms such as Laguerre, when using non-linear least squares to estimate zero-coupon yield curves using coupon bond data. Most importantly the polynomial form invariably enables convergence of the estimation process.

Unfortunately, the simple polynomial form results in estimated models where the zero-coupon yield diverges either to plus or minus infinity as the term increases. This is an unsatisfactory aspect the simple polynomial model is inconsistent with both theoretical considerations and observational reality.

Summary

We used a small number of principal components of a larger number of polynomials of $1/(1+\tau)$ to produce an estimated zero-coupon yield curve with a simple “parsimonious” structure that retained the flexibility of a larger order polynomial form with its many potential turning points. Estimating the model via principal components enabled us to exploit the principal components’ lack of colinearity and also promoted convergence in the non-linear estimation process.

The yield curve constructed from polynomials of $1/(1+\tau)$ displayed the tractable estimation properties of the traditional yield curve modelled as polynomials of τ . However, unlike the traditional polynomial yield curve, the new yield curve had stable long-term yields.

The principal components of polynomials of $1/(1+\tau)$ model was applied to Australian coupon bond data. The results compare favourable to those obtained using the traditional polynomial of term model.

Endnotes

¹ The 1 in the denominator of $1/(1+\tau)$ is there to ensure the term is bounded at $\tau = 0$.

² The polynomial series have been standardised to have a mean of zero and unit variance before principal component analysis was applied. Thus the principal components have identical covariances and correlations.

³ Orthogonality of explanatory variables always eliminates any multicollinearity problems.

⁴ The data were published in the *Australian Financial Review*.

References

Adams, K.J. and van Deventer. "Fitting Yield Curves and Forward rate Curves with Maximum Smoothness", *Journal of Fixed Income*, Vol. 4, No. 1, (June 1994)

Barrett, W.B., T.F.Gosnell, Jr. and A.J.Heuson. "Yield Curve Shifts and the Selection of Immunisation Strategies", *Journal of Fixed Income*, September (1995), pp.53-64

Bhar, R. and B.F.Hunt. "Predicting the Short Term Forward Interest Rate Structure Using a Parsimonious Model", *International Review of Economics and Finance*, Vol 2., No.4. (1994)

Diament, P. "Semi-empirical Smooth Fit to the Treasury Yield Curve", *Journal of Fixed Income*, (June 1993), pp. 55-70

Frishling, W., L.Kameron and F. Stramandinoli. "On a Method of Smoothing the Forward Rate Curve from the zero Coupon Curve", paper presented at the Australasian Banking and Finance Conference, Sydney (1994)

Hunt, B.F. "Fitting Parsimonious Yield Curve Models to Australian Coupon Bond Data", School of Finance and Economics Working paper No. 51, (1995a)

Hunt, B.F. "Modelling the Yields on Australian Coupon Paying Bonds", School of Finance and Economics Working paper No. 9, (1995b)

Nelson, C. R. and A. F. Siegal. "Parsimonious Modeling of Yield Curves." *Journal of Business*, 60, No. 4 (October 1987), pp. 473-489

Pham, T. "Curve Fitting the Term Structure of Interest Rates in Australia: the Nelson-Siegel model versus Chebyshev Polynomials", Unpublished working paper, University of New South Wales, (1995)

Vasicek, O. A. "an Equilibrium Characterization of the Term Structure", *Journal of Financial Economics*, 5, (1977), pp. 177-88

Vasicek, Oldrich A., and H. Gifford Fong. "Term Structure Modeling Using Exponential Splines." *Journal of Finance*, 37, No. 2 (May 1982), pp. 339-348

Table 1: Covariance and Correlation of Polynomials and Principal Components

Covariance : Untransformed Variables

	$1/(1+t)$	$1/(1+t)^2$	$1/(1+t)^3$
$1/(1+t)$	0.0511		
$1/(1+t)^2$	0.0496	0.0520	
$1/(1+t)^3$	0.0452	0.0498	0.0493

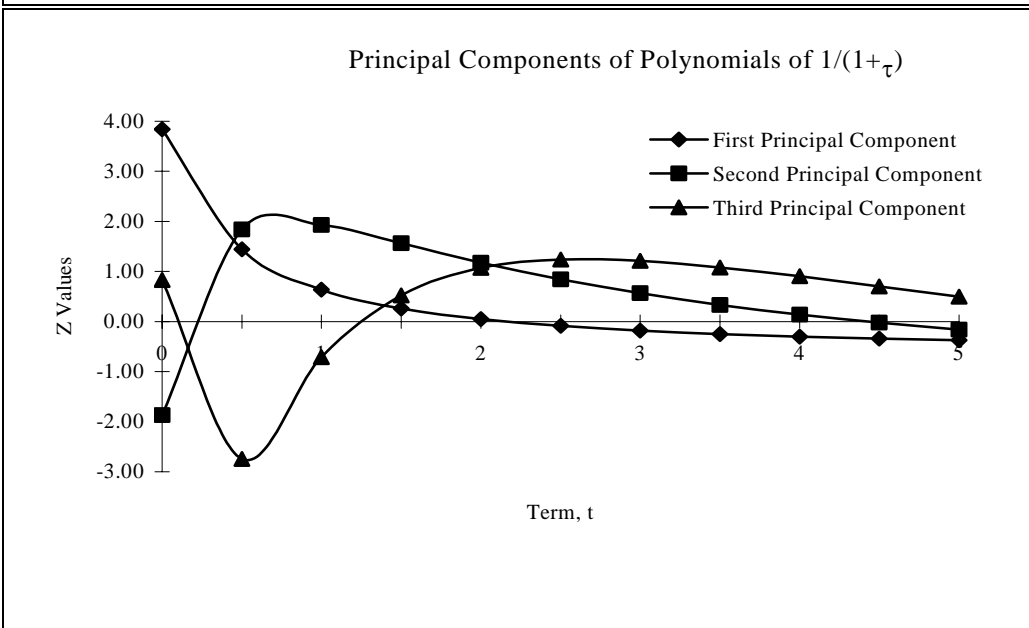
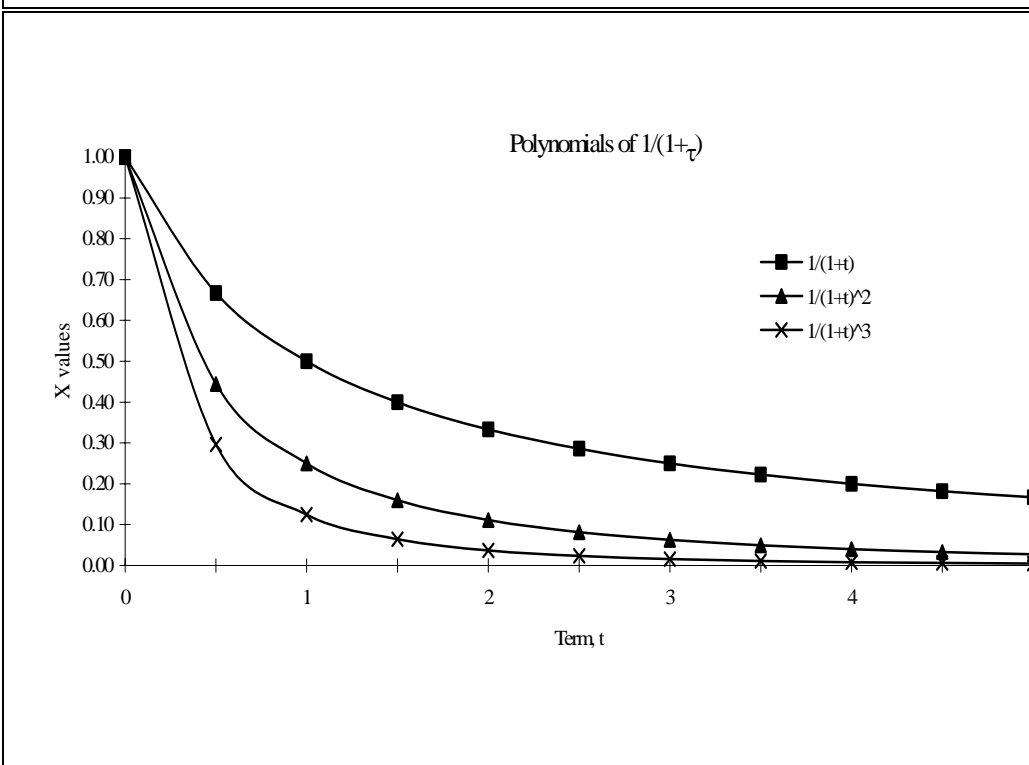
Correlation : Untransformed Variables

	$1/(1+t)$	$1/(1+t)^2$	$1/(1+t)^3$
$1/(1+t)$	1.0000		
$1/(1+t)^2$	0.9613	1.0000	
$1/(1+t)^3$	0.9001	0.9839	1.0000

Covariance and Correlation : Principal Components

	First	Second	Third
First	1.0000		
Second	0.0000	1.0000	
Third	0.0000	0.0000	1.0000

Figure 1 : Polynomials and Principal Components of $1/(1+\tau)$



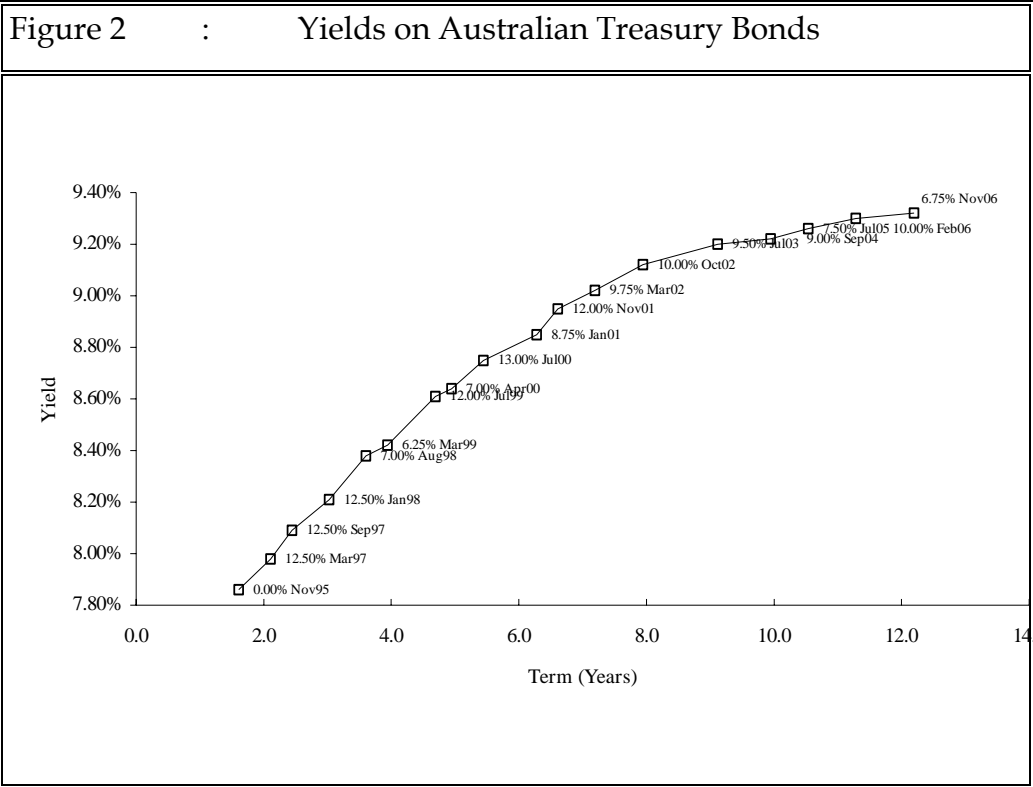
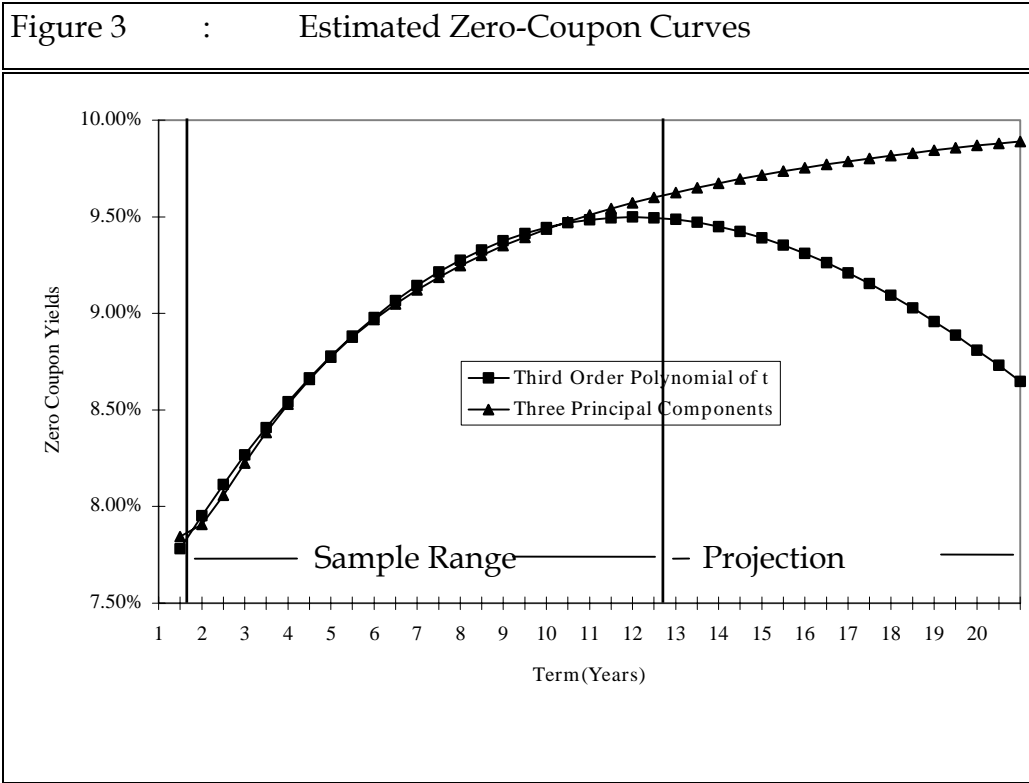


Table 2 : Zero-Curve estimation Results

Sequential Variables	Bond price Residual Sum of Squares	Probability the Additional Variable Coefficient, β_i , is Zero
Constant	9.27	
First principal component	3.36	0.0041%
Second principal component	0.76	0.0001%
Third principal component	0.28	0.0138%
Fourth principal component	0.27	48.3324%



¹ The 1 in the denominator of $1/(1+\tau)$ is there to ensure the term is bounded at $\tau = 0$.

² The polynomial series have been standardised to have a mean of zero and unit variance before principal component analysis was applied. Thus the principal components have identical covariances and correlations.