INTERPOLATING THE SOUTH AFRICAN YIELD CURVE USING PRINCIPAL-COMPONENTS ANALYSIS: A DESCRIPTIVE APPROACH

By AJ Maitland

ABSTRACT

A principal-components analysis of the South African yield curve suggests that two factors explain most of the variability in both yields and changes in yields. This result is used to select which two interest rates to model and, given a model for these rates, how to use them to reproduce the entire curve. The objective of this paper is a methodology for interpolating the South African yield curve given a restricted number of yields on that curve, while at the same time minimising the number of yields from which to estimate the remainder of the curve. The interpolated curve can then be used for the purposes of discounting nominal future cash flows. Given values for the selected yields, this methodology provides the best fit to the remainder of the curve in the sense that it minimises the expected root-mean-squared error of the residuals. The paper does not provide a model for the evolution of the yield curve.

CONTACT DETAILS

James Maitland, Stochastic Solutions cc, P.O. Box 2484, Parklands, 2121, Johannesburg; Tel: (0)11 486-1946; E-mail: stochsol@iafrica.com

1. INTRODUCTION

1.1   It is common practice in pension fund valuations for the actuary to use a single valuation rate of interest in calculating the present value of future asset and liability cash flows. The valuation rate is often assumed to be the long-term rate of interest that can be earned on future investments (Lee, 1986). Conveniently, this approach allows the use of standard actuarial commutation functions.

1.2   More recent application in asset and liability modelling (ALM) studies makes use of the simulated long-bond rate (20-year JSE-Actuaries Bond Yield) in calculating the valuation rate of interest. This provides a proxy for the market rate of interest and gives liability values closer to the market price than the traditional use of a constant discount rate. However, for nominal liabilities where the term of each cash flow is known, the long-bond rate can be a poor approximation to the term structure of interest rates and may give a present value quite different from the cost of the matching portfolio. According to the principle of no arbitrage, these two values should be identical (cf. Head *et al*, 2000: ¶5.2.1).

1.3   Ideally, the full yield curve should be available to place market values on nominal liabilities but, since it is not practical to model all maturities on the yield curve, Tilley (1992: 527) suggests the modelling of eight key yields and the use of linear interpolation to model intermediate maturities. In order to reduce the dimension of the problem further, Sherris (1995) suggests factor analysis to determine the dimension of randomness in the yield curve. In this paper, principal-components analysis (PCA) is used to determine the number of maturities, n, required to adequately describe the South African yield curve. The subset consisting of the first n principal components is then used to interpolate the entire yield curve, given a specific subset of n key points along the curve.

1.4   At this point, it is worth distinguishing between two sources of arbitrage. The first kind of arbitrage exists when two identical sets of cash flows have a different price. The situation discussed above in which liabilities are valued using a single discount rate instead of the term structure of interest rates is one example of this. The second kind exists where an immunised portfolio with a different cash-flow profile costs less than the dedicated portfolio, as discussed in Maitland (2001: ¶6.1).

1.5   In order to reduce the dimension of the model, at the expense of producing non-key yields that are not arbitrage-free, both the methodology suggested by Tilley (*op. cit.*) and that discussed in this paper give rise to arbitrage opportunities of the second kind[1]. However, as Thorlacius (2000) points out:

> For effective [ALM], accuracy and realism are important criteria and as such the simulation must provide an accurate representation of the probabilities of potential economic and market outcomes … The problem comes in trying to create robust model characteristics that reflect those observed in the real world while at the same time confining the computational demands of the model. Structures that ensure arbitrage free interest rates tend to be too simple (and thus produce unrealistic scenarios) or require a large amount of computation power.

1.6   If the purpose of the ALM is to use it for static or dynamic decision-making purposes—for example, to optimise with respect to alternative portfolio selections at future simulation dates—this statement is well justified. It is also justified for static decision-making purposes where realistic probabilities of potential future outcomes are desired. In both cases, realistic market risk premiums and investor risk preferences are relevant to the decision-making process and a simplistic arbitrage-free model would not be useful for such purposes.

1.7   For any given number of key yields, n, the methodology proposed in this paper provides the best fitting (and hence most realistic) yield curve in the sense that it minimises the expected root-mean-squared error of the residuals. The computational burden is also minimal.

1.8   It should be noted that the JSE-Actuaries Yield Curve is itself not arbitrage-free in

the sense that successive curves can give rise to arbitrage opportunities of the second kind, since cubic spline interpolation is used to construct intermediate yields (cf. Section 6). Hence, one cannot hope that a parsimonious, best fitting model will be fully arbitrage-free. Nonetheless, the JSE-Actuaries Yield Curve is still used as the basis for pricing unsecuritised, nominal cash flows and as such provides an important tool for reducing arbitrage opportunities of the first kind. Likewise, the intention of the methodology proposed in this paper is to reduce the magnitude of arbitrage opportunities of the first kind by providing a realistic estimate of the yield curve while at the same time minimising the computational burden[2].

1.9   The question whether it is appropriate to sacrifice the condition of no arbitrage in order to create a functionally simpler model that produces more realistic scenarios depends on the application. The main argument against using a term-structure model with arbitrage is that it is possible to construct investment strategies that perform unreasonably well by exploiting arbitrage opportunities of the second kind. However, if such strategies[3] are excluded and the model is simply used to reduce arbitrage opportunities of the first kind, such arbitrage opportunities should present no problem since they cannot be exploited. Since the objective is a reasonably accurate description of the par yield curve for the purposes of discounting nominal future cash flows while at the same time minimising the number of yields from which to estimate the remainder of the curve, the proposed methodology would appear acceptable[4].

1.10   Having selected the key maturities, these can then be modelled as part of a larger set of variables including other asset categories and economic variables, for the purposes of modelling the assets and liabilities of a financial institution. The dynamic model for these *n* maturities is not discussed in this paper. Use of the interpolated curve should be limited to the discounting of future cash flows and should not be extended to infer the dynamics of interpolated yields, since, as shown by Maitland (2001), these may not be arbitrage-free. Hence, the model proposed is a descriptive model rather than an equilibrium or no-arbitrage model.

1.11   If the presence of arbitrage opportunities of the second kind still gives rise to concern, the technique described by Thorlacius (2000: ¶4) can be used to remove these. This technique essentially works by adding uncertainty in the form of an independent random variable to each interpolated rate, thereby breaking the arbitrage. The standard deviation of these processes can be made small enough that the original model is not significantly disturbed. Using this technique, the fit and statistical characteristics of the original model are broadly retained while providing a model that is arbitrage-free.

2.   BACKGROUND

2.1   Before 1982 there was virtually no active secondary market in bonds. Prescribed-asset legislation forced pension and provident funds and insurance companies

to hold a certain percentage of their assets in respect of liabilities in government bonds, cash and other approved bonds. In the 1970s, insurance companies and pension funds held on average 41% of the long-term domestic marketable stock debt of the central government (compared with 47% by the Public Investment Commissioners); and 70% of local authorities' stock (Falkena *et al*, 1984: 129).

2.2   In the early 1980s an active secondary market in South African bonds began developing and has subsequently grown rapidly (McLeod, 1990). In 1986, the Johannesburg Stock Exchange (JSE) instituted a bond clearing-house and although the majority of bond trading was over the counter (OTC), a small number of trades were recorded on the JSE. Since some trades were recorded at each available maturity, and since these trades would have reflected yields traded OTC, the JSE-Actuaries Yield Curve can be considered to be a fair estimate of market yields prevailing at the time. In 1996, the bond exchange opened and the Financial Markets Control Act now requires all bond trades to be recorded by a recognised exchange.

3.   PRINCIPAL-COMPONENTS ANALYSIS

3.1   A number of empirical studies by academic researchers and practitioners conclude that the short rate is non-stationary. A partial listing of these authors includes Stock & Watson (1988), Mills (1994: 68), Ang & Moore (1994), Johansen & Juselius (1992), Juselius (1995), and Pesaran & Shin (1996). In contrast, many theoretical models of the short-term interest rate assume stationarity and include a mean reversion term (cf. e.g. Vasicek, 1977; Brennan & Schwartz, 1982; Cox, Ingersoll & Ross, 1985), although non-stationary theoretical models also exist. Wilkie (1994) and Thomson (1996) both develop empirical models for interest rates assuming stationarity on the basis of economic rather than statistical arguments. This paper does not investigate issues of stationarity and, since there is no clear consensus, a PCA of both the levels and first differences of the South African yield curve is presented.

3.2   Let x be a random d-vector with mean μ and covariance matrix Σ, and let $\mathbf{T} = (t_1, t_2, \ldots, t_d)$ be an orthogonal matrix such that $\mathbf{T}'\Sigma\mathbf{T} = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_d)$, where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_d$ are the eigenvalues of Σ. If $y = \mathbf{T}'(x - \mu)$, then $y_j = t_j'(x - \mu)$ is called the jth principal-component score of x and is the orthogonal projection of $x - \mu$ in the direction t_j (Seber, 1984: 176). Principal-components analysis explains the variance-covariance structure of the original variables through an orthogonal rotation of x such that the first principal component gives the direction of maximum variation, the second gives the next largest direction of maximum variability orthogonal to the first principal component, and so on. If Σ is positive definite, d principal components are required to reproduce the total system variability completely, but much fewer principal components may explain a reasonable proportion of the total variability and hence reduce the dimension of the model with only a small loss of information.

3.3   We define the yields for annual terms from 0 and 25 years along the JSE-Actuaries Yield Curve (with the INET (1998) codes *JAYC00*, *JAYC01*… *JAYC25*) to be our 26-dimensional random vector. If yields are stationary, the moments of the level yields exist. Table 1 provides summary statistics for key yields at annual maturities from 0 to 25 years using monthly data from January 1986 to December 1998, while Figure 1 illustrates the yield curve over this period.

TABLE 1. JSE-Actuaries Yield Curve (1986–1998)

	mean	median	standard deviation	minimum	maximum
JAYC00	14,55	14,58	3,31	8,47	22,99
JAYC01	14,38	14,59	2,72	9,24	22,31
JAYC02	14,28	14,55	2,25	10,01	21,64
JAYC03	14,37	14,72	1,94	10,36	21,06
JAYC04	14,59	14,90	1,74	10,61	20,59
JAYC05	14,84	15,21	1,59	10,91	20,24
JAYC07	15,26	15,67	1,41	11,39	19,84
JAYC10	15,58	15,80	1,33	11,75	19,72
JAYC15	15,69	15,94	1,29	11,96	19,49
JAYC20	15,65	15,77	1,25	12,02	19,34

FIGURE 1. JSE-Actuaries Yield Curve (1986–1998)
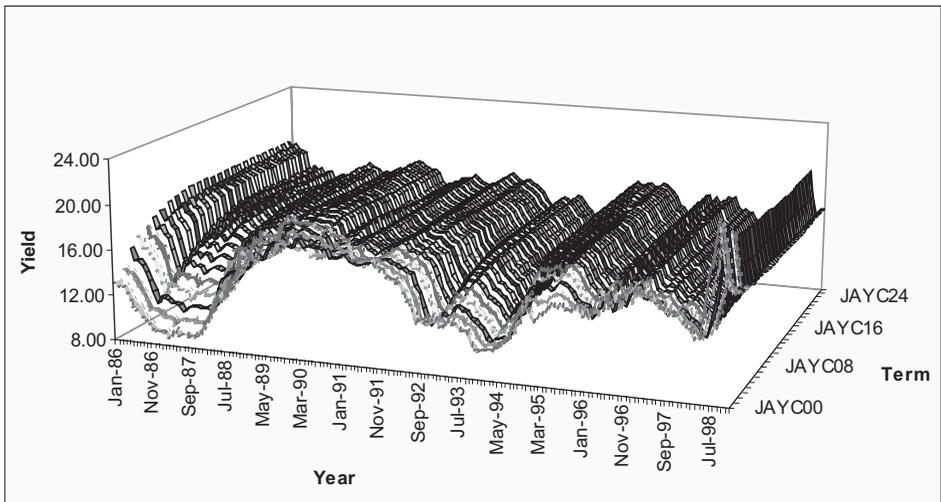
3.4   A PCA on the covariance matrix of yields reveals that the first principal component explains 77,1% of the total variability in the yield curve, the first two together explain 98,4% and the first three together explain 99,4% of the total variability in the yield curve. Figure 2 illustrates the coefficients of each of the first three principal components by term to maturity.

FIGURE 2. Coefficients for the first three principal components of yield levels
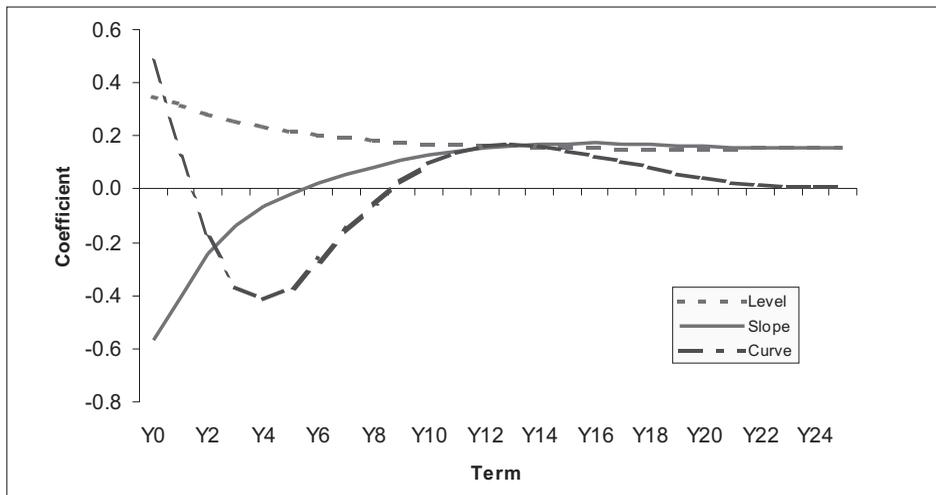
3.5   The coefficients for the first principal component are all positive, so that an increase in the score of the first principal-component results in an increase in all yields. The first principal component can therefore be regarded as a level factor. Since the coefficients are not all equal, a change in the score of the first principal component does not result in a parallel shift; instead, the short end of the curve moves more than the long end.

3.6   The coefficients for the second factor are negative at the short end and monotonically increase to a positive value at the long end. Hence, a change in the score of the second principal-component results in an opposite effect on the two ends of the yield curve, and this factor can be viewed as causing a change or twist in the slope of the yield curve. The third principal component has a negative effect on medium yields and a positive effect on short and long-term yields and hence can be interpreted as a hump factor or butterfly. Figure 3 illustrates the principal-component scores for the first three principal components from January 1986 to December 1998.

3.7   The third principal component accounts for only 1% of the total variability and the remaining 23 principal components account for about 0,5% of the total variability. Hence, two principal components appear to capture most of the variability in the yield curve. This

FIGURE 3. Principal-component scores for yields (1986–1998)
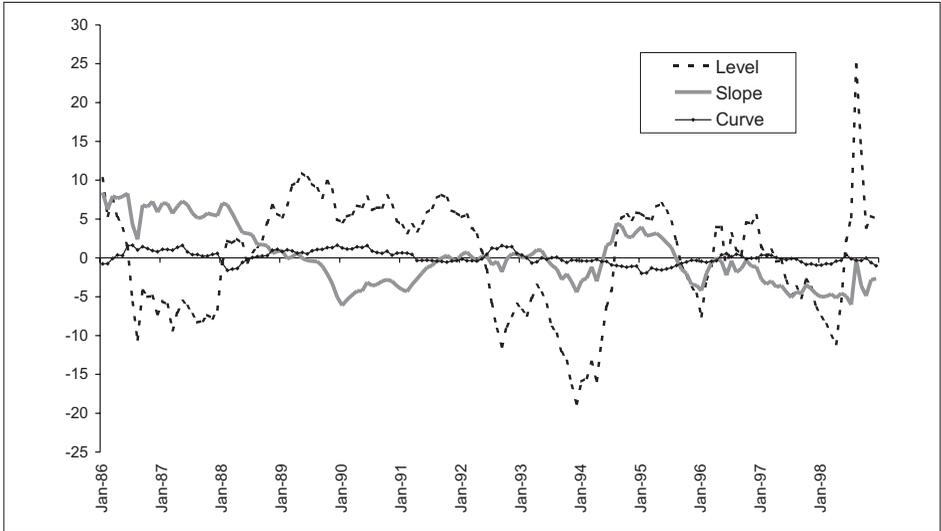
is supported by the informal scree test illustrated in Figure 6 and discussed in Section 5. Section 7 discusses how these principal components can be used to reconstruct the entire yield curve.

3.8   So far, we have considered the covariance matrix of yields. However, if yields are non-stationary, then the population moments do not exist. We now define the monthly changes in yields for annual terms from 0 and 25 years along the JSE-Actuaries Yield Curve to be our 26-dimensional random vector and again use monthly yield data from January 1986 to December 1998.

3.9   A PCA on the covariance matrix of changes in yields reveals that the first principal component alone explains 92,8% of the total variability, the first two together explain 97,3% and the first three together explain 98,4% of the total variability. Hence, two principal components again appear to capture most of the variability in yield curve changes. Figure 4 illustrates the coefficients of the first three principal components by term to maturity, while the scree test illustrated in Figure 6 supports the choice of two principal components.

3.10   The first principal component affects all maturities by similar amounts and in the same direction. It can be interpreted as a level shift factor but not as a parallel shift factor since the coefficients are unequal. Unlike the levels PCA, the short end of the curve moves less than the long end in response to the score of the first principal component. The second factor has an opposite effect on short and long yields and can be viewed as a slope change or twist factor. The third principal component has a negative effect on medium

yields and a positive effect on short- and long-term yields, and hence can be interpreted as a curvature or butterfly factor. Figure 5 illustrates the principal-component scores for the first three principal components of yield curve changes for the period January 1986 to

FIGURE 4. Coefficients for the first three principal components of yield changes
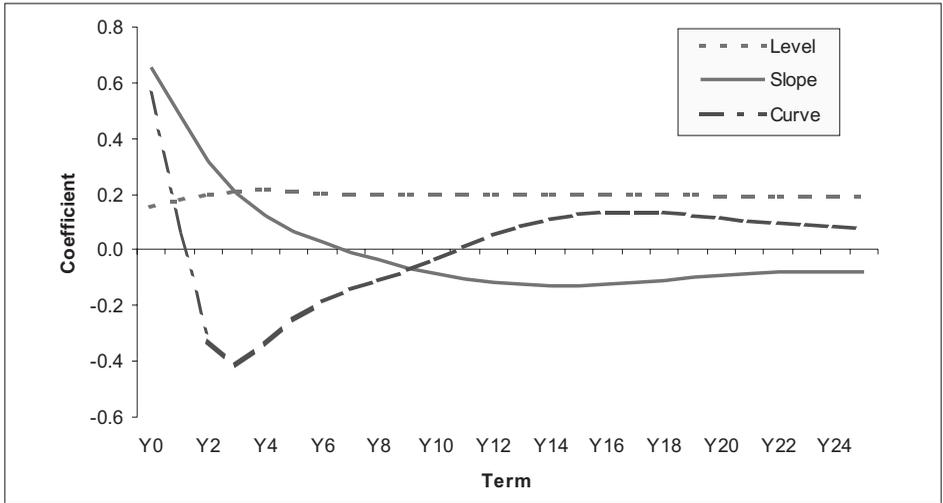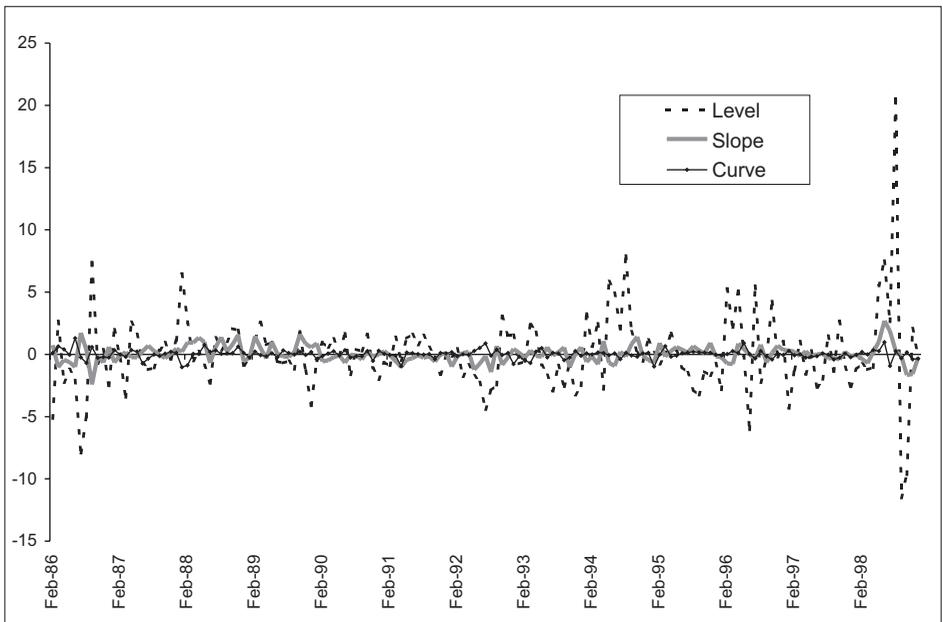

FIGURE 5. Principal-component scores for yield changes (1986–1998)

December 1998. International PCA results indicate the presence of similar components, although the proportion of variability explained by each component differs considerably from one market to another. Maitland (2001) gives a partial list of references to these results.

4. RMSE OPTIMALITY OF THE PRINCIPAL-COMPONENTS APPROXIMATION

4.1   The expected RMSE is commonly used as a simple descriptive measure of the fit of a particular model (Anderson *et al*, 1996: 59). For a maximum of *n* factors, it turns out that the first *n* principal components are optimal in the sense that they minimise the RMSE over all linear combinations of factors. A proof of this result, which, to the author's knowledge, has not been documented in the literature, is given in Appendix A.

4.2   In Section 7, a methodology is presented for interpolating the yield curve given a restricted number of yields and using the principal components. It should be noted that the principal components are optimal provided the key yields are known a priori. If the key yields are not known a priori but are projected using some stochastic model, then the RMSE may not be a valid measure of the fit of the forecast yield curve, since that may depend on the model for these key yields. It may be possible under certain conditions to separate the optimality of the projected yields from the interpolated curve whose optimality is conditional on the projected yields. However, an investigation into these conditions is beyond the scope of this paper and is left for future research.

5. THE SCREE TEST

5.1   By plotting the root-mean-squared error (RMSE) of estimated yields against the number of parameters or factors included in the model, the marginal gain in explanatory power can be visually offset against the increase in the number of parameters. This informal statistical test is known as a 'scree' test (Cattell, 1965) and Monte Carlo studies (Tucker, Koopman & Linn, 1969) have shown that it is often superior in locating major common factors when minor factors are at play. In fitting parametric curves to yield-curve data, Chaplin (1998: 344–7) supports his choice of model using this test in preference to more formal statistical tests for assessing model fit.

5.2   PCA directly locates the factors and places them in order of importance, as discussed in Section 3. The total system variance as defined by the sum of the diagonal elements of $\mathbf{\Sigma}$, $\sigma_{11} + \sigma_{22} + \ldots + \sigma_{dd}$, is equal to $\lambda_1 + \lambda_2 + \ldots + \lambda_d$. This follows since $\operatorname{tr}(\mathbf{T'\Sigma T}) = \operatorname{tr}(\mathbf{\Sigma T'T}) = \operatorname{tr}(\mathbf{\Sigma I}_d) = \operatorname{tr}(\mathbf{\Sigma})$, which follows from the properties of the trace operator and the fact that $\mathbf{T'}$ is orthogonal. As shown in equation (3), the expected mean-squared error of yields (or changes in yields) from approximating the yield curve (or changes in the yield curve) with the first *n* principal components is given by $(\lambda_{n+1} + \ldots + \lambda_d)/d$ (for the corresponding $\mathbf{\Sigma}$). Hence, the ratio $(\lambda_{n+1} + \ldots + \lambda_d)/(\lambda_1 + \ldots + \lambda_d)$ gives an equivalent scree test.

5.3   Figure 6 illustrates the scree plot for the PCA of both the yields and changes in yields. In both cases, the improvement in fit from two to three parameters is insubstantial for the purpose at hand, bearing in mind the need for parsimony in dimension of the dynamic model. For certain applications a better description of yields may be preferred, in which case more principal components can be used to approximate the yield curve.

FIGURE 6. Scree plot for yields (Y) and changes in yield (DY)
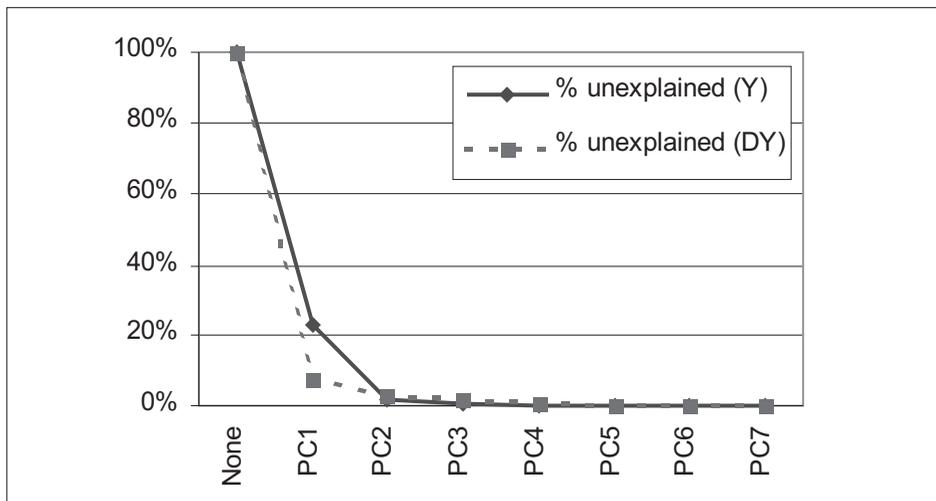
6.   A DISCUSSION OF THE PCA RESULTS

6.1   The traditional theory of immunisation as developed by Redington (1952) immunises a portfolio against parallel shifts in the yield curve. Parallel shifts imply the existence of arbitrage opportunities (cf. Boyle, 1978) and it is important to note that the first principal component does not represent an entirely parallel shift. However, for terms greater than five years, the first principal component does seem to represent a parallel shift, and for terms greater than 12 years, the second principal component also seems to represent virtually parallel shifts. Hence, the first two principal components, which represent 97,3% of the total variability, appear to indicate the regular occurrence of parallel shifts. However, this does not necessarily imply the existence of arbitrage opportunities at the long end of the curve, since, on average, 2,7% of the variability remains unexplained. Maitland (2001) shows how to identify arbitrage opportunities conditional on the absence of higher-order principal-component shifts.

6.2   Estimates of variances, covariances and correlations can be very sensitive to outliers and so we can expect principal components to have the same sensitivity. The extreme scores for the first principal component between August and October 1998

shown in Figures 3 and 5, and the corresponding large changes in the level of the yield curve evident from Figure 1, suggest the need for a PCA for sub-periods of the data. For the sub-period 1986 to 1997, the proportion of the variability explained by the first principal component of yield curve changes decreases from 92,8% (for the period 1986 to 1998) to 90,0%. It should be noted that although the extreme events occurred at points in time, the time-series properties of the scores are irrelevant for the purposes of this paper.

6.3   For level yields, the proportion of the variability explained by the first principal component reduces from 77,1% (for the period 1986 to 1998) to 76,1% for the sub-period 1986 to 1997. In both the yield and the differenced yield sub-period analyses, the principal components remain relatively unchanged, suggesting that the full-period analysis is relatively robust to the outliers from August to October 1998. A number of alternative sub-periods were considered and the results of the full period appeared to be relatively robust to the choice of sub-period.

6.4   In the above analyses, principal components are derived from the covariance matrix. If the variables in a PCA are measured on scales with widely differing ranges, it is preferable to use the correlation matrix (cf. Seber, 1984). Although the higher volatility of short rates compared with long rates results in an increased loading of the short rate on the first few factors, a PCA for both the levels yields and yield differences using the correlation matrix gives principal components and variability proportions that are similar to those obtained using the covariance matrices. Hence, the results of the PCA on the covariance matrix appear to be relatively robust to the lack of scaling. This is not too surprising given that the standard deviations of short and long yields are of the same order of magnitude.

6.5   One further point worth considering is the effect that the mathematical formulation of the JSE-Actuaries Yield Curve may have on the principal-components analysis. The curve is constructed in two steps (McLeod, 1990):
        1.   Using a form of cluster analysis, five cluster points are estimated and bonds are assigned to each cluster. The bonds in each cluster are then used to determine a weighted average term to maturity and a weighted average yield for their respective clusters. A sixth cluster with a maturity of 30 years and yield equal to the weighted average yield of the cluster with the highest weighted average yield is also determined.
        2.   Using these six cluster points, intermediate points along the curve are estimated using cubic spline interpolation.

6.6   Since the yield value of the sixth cluster is derived directly from one of the existing five cluster points, there are effectively five independent points along the curve. Hence, it is unlikely that more than five principal components would be required to reproduce most of the variability of the yield curve. The fact that two principal components capture most of the variability is a strong indication that the PCA is not constrained by the mathematical formulation of the yield curve.

7.   RECONSTRUCTING THE YIELD CURVE

7.1   Using the principal components, **T**, and the principal-component scores at time t, $\mathbf{y}_t$, of the level yields (or changes in yields), the level yields (or changes in yields) at time t, $\mathbf{x}_t$, can be reconstructed as $\mathbf{x}_t = \mathbf{T} \cdot \mathbf{y}_t + \boldsymbol{\mu}$. Since the first two principal components capture most of the variability in $\mathbf{x}$ for a PCA of both the levels and first differences,

$$\mathbf{x}_t \approx y_{1,t} \cdot \mathbf{t}_1 + y_{2,t} \cdot \mathbf{t}_2 + \boldsymbol{\mu}.$$

7.2   For users deciding which variables to include in a stochastic model, it would be possible to model $y_{1,t}$ and $y_{2,t}$. However, since the relationship between these variables and the remaining variables in the stochastic model depends on the eigenvectors $\mathbf{t}_1$ and $\mathbf{t}_2$, the resulting model may be difficult to interpret. Since a more direct and theoretically tractable relationship exists between actual yields and other stochastic variables, if any two yields (or changes in yields), $x_{a,t}$ and $x_{b,t}$, are modelled stochastically, these can be used to estimate $y_{1,t}$ and $y_{2,t}$, from which can be derived the full yield curve as explained above. More formally:

$$\text{Let} \quad \mathbf{z}_t = \begin{bmatrix} z_{a,t} \\ z_{b,t} \end{bmatrix} = \begin{bmatrix} x_{a,t} \\ x_{b,t} \end{bmatrix} - \begin{bmatrix} \mu_{a,t} \\ \mu_{b,t} \end{bmatrix}. \tag{1}$$

$$\text{Then,} \quad \mathbf{z}_t = \begin{bmatrix} t_{1,a} & t_{2,a} \\ t_{1,b} & t_{2,b} \end{bmatrix} - \begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \Im \cdot \widetilde{\mathbf{y}}_t \quad \text{say,} \tag{2}$$

$$\text{so} \quad \widetilde{\mathbf{y}}_t = \Im^{-1} \cdot \mathbf{z}_t. \tag{3}$$

$$\text{Hence } \mathbf{x}_t = \begin{bmatrix} t_1 & t_2 \end{bmatrix} \cdot \Im^{-1} \mathbf{z}_t + \boldsymbol{\mu}_t. \tag{4}$$

7.3   Monthly data for the JSE-Actuaries Yield Curve for annual terms from 0 to 25 years are available from January 1986 onwards. Before this, only yields on 3- and 20-year bonds are available as well as the Alexander Forbes Money Market Index, from which can be derived a proxy for the short rate. These three series are available from 1960 onwards. Hence, if data for the full period from 1960 to 1998 are required for modelling purposes, it is possible to model only these three points on the yield curve.

7.4   If most of the variability in the yield curve could be explained by one principal component, the correlation between yields at different terms would be close to one and the yield at any term would be sufficient to reproduce the entire yield curve. Since two principal components are required to explain most of the variability in the yield curve, we require two terms, a and b, to reproduce the entire yield curve. These two terms should be chosen so that the absolute correlation between them is as small as possible in order to minimise the error in estimating $y_{1,t}$ and $y_{2,t}$. The correlation matrix for *JAYC00*, *JAYC03* and *JAYC20* is presented in Table 2. (Level yield correlations are shown below the diagonal and differenced yield correlations above.)

TABLE 2. Yield correlations (1986–1998)

	JAYC00	JAYC03	JAYC20
JAYC00	1	0,74	0,67
JAYC03	0,87	1	0,92
JAYC20	0,44	0,75	1

7.5   The correlations between *JAYC00* and *JAYC20* in Table 2 are less than the other correlations, suggesting that *JAYC03* can be dropped from the set of model variables. Figures 2 and 4 confirm this suggestion since the greatest differences between the coefficients of the first and second principal components are at the short and long maturities. Further, for most months between January 1986 and December 1998, *JAYC03* lies between *JAYC00* and *JAYC20*. Since the difference in term between *JAYC00* and *JAYC20* is the largest, errors in forecasting *JAYC00* and *JAYC20* have a smaller effect on the forecast error for *JAYC03* than any other pair of yields might have on the remaining yield.

8.   CONCLUSION

8.1   The proposed methodology provides a way in which a yield curve can be interpolated from a restricted number of modelled yields, while at the same time minimising the number of yields from which to estimate the remainder of the curve. From a statistical perspective, the short rate and the long-bond yield should be used to reconstruct the South African yield curve, given the first and second principal components. Hence, for the purposes of reconstructing the yield curve, one need model only the short rate and the long-bond yield. If these variables are modelled as non-stationary variables, the yield curve can be reconstructed given forecast changes together with the yield curve at time zero. Otherwise, the yield curve can be reconstructed direct using equations (1) to (4). It should be noted that the optimality of the interpolated yield is conditional on the key yields being given a priori, as discussed in Section 4.

8.2   A number of other reasons exist for modelling the long-bond yield and the short rate as part of a larger set of variables, but a discussion of this is beyond the scope of this paper. However, the results in this paper give further credence to the choice of variables modelled by Thomson (1996). The methodology presented in this paper is not a framework for projecting financial and economic factors, but rather a methodology for interpolating the yield curve given these factors. It is suggested that this methodology be used to supplement the development of future stochastic investment models.

8.3   As discussed in Section 1, the interpolated curve should not be used to infer the dynamics of interpolated yields. Rather it should be used to value future cash flows in a more realistic manner. The purpose of interpolation is not to optimise with respect to

alternative terms, as described in Maitland (2001), but rather to revalue, at the end of a simulation interval, the bonds whose terms have shortened (by one interval) and to value the liabilities. For the purposes of interpolating arbitrage-free yields, readers are referred to the methodology of Heath, Jarrow & Morton (1992). Maitland (2001) provides a methodology for determining the number of principal components to include for the purposes of short-term risk analysis, based on the financial significance of additional principal components.

NOTES

1   It should be noted that the use of a single, simulated discount rate for the valuation of nominal liabilities also implies the existence of arbitrage opportunities of the second kind, since this is equivalent to assuming parallel shifts in a flat yield curve (cf. Boyle, 1978).

2   For bonds that conform to the first and second principal components, arbitrage opportunities of the first kind are completely avoided by using the methodology proposed in Section 7.

3   i.e. strategies involving the short sale of any asset or assuming liabilities are marketable.

4   NB: The proposed model is not intended for use in derivatives pricing or for constructing dynamic strategies where the existence of arbitrage opportunities of the second kind is unacceptable. For arbitrage-free derivatives pricing, the interested reader is referred to the methodology of Heath, Jarrow & Morton (1992). For constructing dynamic strategies, the interested reader is referred to Maitland (2001).

REFERENCES
Anderson N, Breedon F, Deacon M, Derry A & Murphy G (1996). *Estimating and interpreting the yield curve*. John Wiley & Sons
Ang A & Moore D (1994). The Australian yield curve. Unpublished working paper, Macquarie University, Sydney
Brennan MJ & Schwartz ES (1982). An equilibrium model of bond pricing and a test of market efficiency, *Journal of Financial and Quantitative Analysis* **17**, 75–100
Boyle PP (1978). Immunization under stochastic models of the term structure. *JIA* **105**, 177–87
Cattell RB (1965). Factor analysis: an introduction to fundamentals. *Biometrics* **21**, 190–215
Chaplin GB (1998). A review of term structure models and their applications. *BAJ* **4**, 323–49

Cox JC, Ingersoll JE & Ross SA (1985). A theory of the term structure of interest rates. *Econometrica* **53**, 385–407

Darroch JN (1965). An optimal property of principal components. *Annals of Mathematical Statitics* **36**, 1579

Falkena HB, Fourie LJ & Kok WJ (1984). *The Mechanics of the South African Financial System*. MacMillan, Johannesburg, 2nd edition

Head SJ, Adkins DR, Cairns AJG, Corvesor DO, Cule DO, Exeley CJ, Johnson IS, Spain JG & Wise AJ (2000). Pension fund valuations and market values. *BAJ* **6**, 55–143

Heath D, Jarrow R & Morton A (1992). Bond pricing and the term structure of interest rates; a new methodology. *Econometrica* **60**, 77–105

INET (1998). INET Graphics and Database Service, Intelligent Network (Pty) Ltd

Johansen S & Juselius K (1992). Testing structural hypotheses in a multivariate cointegration analysis of the PPP and UIP for UK. *Journal of Econometrics* **53**, 211–44

Juselius K (1995). Do PPP and UIP hold in the long-run. *Journal of Econometrics* **69**, 211–40

Lee EM (1986). *An introduction to pension schemes*. Institute of Actuaries, London

McLeod HD (1990). The development of a market yield curve: the South African solution. *1st AFIR International Colloquium* (2), 197–212

Maitland AJ (2001). An empirical approach to immunization in South Africa. *SAAJ* **1**, 119–38

Mills TC (1994). *The econometric modelling of financial time series*. Cambridge University Press

Pesaran MH & Shin Y (1996). Cointegration and speed of convergence to equilibrium. *Econometrica* **71**, 117–43

Redington FM (1952). Review of the principles of life office valuations. *JIA* **78**, 286–315

Seber GAF (1984). *Multivariate Observations*. Wiley

Sherris M (1995). Interest rate risk factors in the Australian bond market. *Proceedings of the 5th AFIR Colloquium*

Stock JH & Watson M (1988). Testing for common trends. *Journal of the American Statistical Association* **83**, 1097–107

Thomson RJ (1996). Stochastic investment modelling: the case of South Africa. *BAJ* **2**, 765–801

Thorlacius AE (2000). Arbitrage in asset modelling for integrated risk management. *Proceedings of the 10th AFIR Colloquium*

Tilley JA (1992). An actuarial layman's guide to building stochastic interest rate generators. *Transactions of the Society of Actuaries* **44**, 509–64

Tucker LR, Koopman RF & Linn RL (1969). Evaluation of factor analytic research procedures by means of simulated correlation matrices. *Psychometrica* **34**, 421–59

Vasicek O (1977). An equilibrium characterisation of the term structure. *Journal of Financial Economics* **5**, 177–88

APPENDIX A
OPTIMALITY OF THE FIRST n PRINCIPAL COMPONENTS

A.1 The expected RMSE is commonly used as a simple, descriptive statistical measure of the fit of a particular model (Anderson *et al*, 1996: 59). For a maximum of n factors, the first n principal components are optimal in the sense that they minimise the RMSE over all linear combinations of n factors.

A.2 PROOF

A.2.1 Since $y_t = \mathbf{T}'(\mathbf{x}_t - \boldsymbol{\mu})$, the model for $\mathbf{x}_t$ based on the first n principal components is

$$\hat{\mathbf{x}}_t = \boldsymbol{\mu} + \sum_{j=1}^{n} y_{j,t}\mathbf{t}_j ;$$ (A1)

where $\mathbf{x}_t - \hat{\mathbf{x}}_t = \boldsymbol{\varepsilon}_t = \sum_{j=n+1}^{d} y_{j,t}\mathbf{t}_j$.

A.2.2 In equation (A1), $\mathbf{x}_t$, $\boldsymbol{\mu}$, $\mathbf{t}_j$ and $\boldsymbol{\varepsilon}_t$ are all $(d \times 1)$ vectors and $y_{j,t}$ is a scalar. The mean-squared error multiplied by d (MSEd) is given by:

$$\boldsymbol{\varepsilon}_t' \boldsymbol{\varepsilon}_t = \left(\sum_{j=n+1}^{d} y_{j,t}\mathbf{t}_j\right)'\left(\sum_{j=n+1}^{d} y_{j,t}\mathbf{t}_j\right) = \sum_{j=n+1}^{d} y_{j,t}^2 \mathbf{t}_j'\mathbf{t}_j = \sum_{j=n+1}^{d} y_{j,t}^2$$ (A2)

since the eigenvectors, $\mathbf{t}_j$, are orthonormal. Hence, the expected MSEd is given by:

$$\mathrm{E}\left[\boldsymbol{\varepsilon}_t' \boldsymbol{\varepsilon}_t\right] = \mathrm{E}\left[\sum_{j=n+1}^{d} y_{j,t}^2\right] = \sum_{j=n+1}^{d} \mathrm{E}\left[y_{j,t}^2\right] = \sum_{j=n+1}^{d} \lambda_j .$$ (A3)

A.2.3 This follows since $\mathrm{V}(y_t) = \mathrm{E}[(y_t - \mathrm{E}[y_t])(y_t - \mathrm{E}[y_t])'] = \mathrm{E}[y_t y_t']$ as $\mathrm{E}[y_t] = \mathbf{T}'(\mathrm{E}[\mathbf{x}_t] - \boldsymbol{\mu}) = \mathbf{0}$ so that $\mathrm{V}(y_t) = \mathbf{T}'\mathrm{E}[(\mathbf{x}_t - \boldsymbol{\mu})(\mathbf{x}_t - \boldsymbol{\mu})']\mathbf{T} = \mathbf{T}'\Sigma\mathbf{T} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_d)$ and hence $\mathrm{E}(y_{j,t}^2) = \lambda_j$.

A.2.4 Expressing equation (A3) in terms of the trace of the residual covariance matrix gives:

$$\mathrm{E}\left[\boldsymbol{\varepsilon}_t' \boldsymbol{\varepsilon}_t\right] = tr\left\{\mathrm{E}[\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t']\right\} = tr\left\{\mathrm{E}\left[(\mathbf{x}_t^* - \hat{\mathbf{x}}_t^*)(\mathbf{x}_t^* - \hat{\mathbf{x}}_t^*)'\right]\right\};$$

$$= tr\left\{\mathrm{E}\left[(\mathbf{x}_t^* - \Xi\boldsymbol{\psi}_t)(\mathbf{x}_t^* - \Xi\boldsymbol{\psi}_t)'\right]\right\}$$ (A4)

where * denotes that the variable has been centred, Ξ is a $(d \times n)$ matrix corresponding to the first n columns of $\mathbf{T}$ and $\boldsymbol{\psi}_t$ is an n-component vector corresponding to the first n elements of y_t.

A.2.5 Next, suppose we wish to select a set of n d-component vectors (factors) $\mathbf{S} = \{s_1, s_2, \ldots s_n\}$ and corresponding $(n \times m)$ matrix of weights $\xi = \{\xi_1, \xi_2, \ldots, \xi_m\}$ for each time $t = \{1, 2, \ldots, m\}$ to approximate the true curve at each time by a linear combination of these factors. Then the total expected MSEd, τ_n, is given by:

$$\tau_n = \sum_{t=1}^{m} E[\text{MSE}d_t] = tr\left\{ E\left[\sum_{t=1}^{m} (x_t^* - \mathbf{S}\xi_t)(x_t^* - \mathbf{S}\xi_t)' \right] \right\}$$

$$= tr\left\{ E\left[(x^* - \mathbf{S}\xi)(x^* - \mathbf{S}\xi)' \right] \right\};$$

(A5)

where x^* is the $(d \times m)$ matrix corresponding to $x_1^*, x_2^*, \ldots, x_m^*$. Darroch (1965) has shown that τ_n is minimised with respect to $\mathbf{S}$ and ξ if and only if

$$\mathbf{S}\xi = \sum_{j=1}^{n} t_j t_j' x^*$$

$$\left(= \sum_{t=1}^{m} \sum_{j=1}^{n} y_{j,t} t_j \right).$$

(A6)

A.2.6 Hence, the model for x_t based on the first n principal components gives the minimum value of τ_n, which is $m(\lambda_{n+1} + \ldots + \lambda_d)$, as shown by equation (A3). Since the function $f : x \mapsto \sqrt{x / md}$ is monotonically increasing, minimising the RMSE is equivalent to minimising τ_n. This proves the result that the principal components are best in the sense that they minimise the RMSE over all linear combinations of variables.