Long term extrapolation and hedging of the South African yield curve

by
Michael Patrick Thomas

Submitted in partial fulfillment of the requirements for the degree
Magister Scientiae
in the Faculty of Natural & Agricultural Sciences
University of Pretoria
Pretoria

April 2008

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

DECLARATION

I, the undersigned, hereby declare that the dissertation submitted herewith for the degree Magister Scientiae to the University of Pretoria contains my own, independent work and has not been submitted for any degree at any other university.

Name: Michael Patrick Thomas
Date: April 2008

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Title: Long term extrapolation and hedging of the South African yield curve
Name: Michael Patrick Thomas
Supervisor: Prof E Maré
Department: Mathematics and Applied Mathematics
Degree: Magister Scientiae

Summary

The South African fixed interest rate market has historically had very little liquidity beyond 15 - 20 years. Most financial institutions are currently prepared to quote and trade interest rate risk up to a maximum term of 30 years. Any trades beyond 30 years usually attract very onerous spreads and raise relevant questions regarding an appropriate level of mid - rates. However, there are many South African entities whose business involves taking on exposure to interest rates beyond 30 years, such as life insurance companies and pension funds. These entities have historically used very traditional approaches to hedging their interest rate exposures across the whole term structure and have typically done little to gain any further protection.

We can generalise the problems faced by any entity exposed to long term interest rate risk in South Africa:

1. The inadequacy of traditional matching methods (i.e. immunisation and bucketing) to cope with the long term interest rate risks.
2. The non-observability of interest rate data beyond the maximum term in the yield curve. Associated with this is the inability to adequately quantify interest rate risk.
3. The lack of liquidity in long term interest rate markets. Associated with this is the inability to adequately hedge long term interest rate risk.

We examine various traditional approaches to matching / hedging interest rate risk using information available at observable / tradable terms on the nominal yield curve. We then look at the reasons why these approaches are not suitable for hedging long term interest rate risk.

Some modern methods to forecast and hedge long term interest rate risks are then examined and the possibility of their use in managing long term interest risk is explored. On the back of these investigations, we propose a number of possible yield curve extrapolation procedures and methodology for performing calibrations.

Using some general theoretical hedging results, we perform a case study which analyses the performance of various theoretical hedges over a historical period from October 2001 to March 2007. The results indicate that extrapolation and hedging of the yield curve is able to significantly reduce Value-At-Risk of long term interest rate exposures.

A second case study is then performed which analyses performance of the various theoretical hedges using out-of-sample simulated yield curve data.

We find that there appears to be a significant benefit to the use of yield curve extrapolation techniques, specifically when used in conjunction with a hedging strategy. In some cases we find that the more simple extrapolation techniques actually increase risk (significantly) when used in conjunction with hedging. However, for some of the more advanced techniques, risk can be significantly reduced.

For an entity looking to deal with long term interest rate risk, we find that the choice of extrapolation technique and hedging strategy go hand-in-hand. For this reason the cost of hedging and reduction in risk are strongly correlated. The results we obtain suggest that it is necessary to weigh the benefits against the cost of hedging. Further, this cost seems to increase with increasing reduction in risk. The research and results presented here are related to those in the paper **Long Term Forecasting and Hedging of the South African Yield Curve** presented by Thomas and Maré at the 2007 Actuarial Convention in South Africa.

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

ACKNOWLEDGEMENTS

I would like to thank the following people who have assisted me in producing this dissertation:

Professor Eben Maré for his valuable insights and contributions throughout the production of this dissertation.

Andrew Smith, a founder of the Smith-Wilson model, for sharing his unique perspectives and his encouragement to challenge conventional thinking.

Dmitri Gott, a good friend, whose eccentricities know no bounds. Without his encouragement and high spirits (in both senses), I would never have started this research.

Ilze Janse van Rensburg, for an endless supply of patience...

CONTENTS

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

LIST OF TABLES

LIST OF FIGURES

NOTATION

The following notation will apply generally throughout the dissertation:

$f_t(\tau)$: denotes the forward rate at time t for term $\tau - 1$ to τ , where a unit of time is equal to 1 year.

$z_t(\tau)$: denotes the zero coupon yield at time t for a term of τ.

$s_t(\tau)$: denotes the at-the-money annual swap rate at time t for a term of τ.

$P_t(\tau)$: denotes the price of a zero coupon bond at time t of term τ.

Note that in certain areas notation may depart slightly from that above, particularly in Chapter 2 where results from other papers / researchers are explored. However, any alternative notations will be described fully as they are used.

TERMINOLOGY

Guaranteed Annuity Option
In a life insurance context, this is an option for a policyholder to convert the maturity proceeds of their retirement policy into a life annuity at guaranteed terms. This effectively represents an option on interest rates and mortality at a future date.

Defined Benefit Pension
A pension where the ultimate retirement benefit is expressed in terms of an employee's salary at (or prior to) retirement. Because the ultimate benefit is in the form of an annuity, these bear a large interest rate risk.

Pension in Accumulation Period
A pension where an employee has not yet reached their retirement date and hence the ultimate retirement benefit is unknown.

Intertemporal Consistency
Suppose that we have a family of forward rate curves, denoted by χ. Suppose also that we have an interest rate model M which represents behaviour of the financial markets. Bjork and Christensen (1999) define the concept of **consistency** as follows: the pair (M, χ) are consistent if all forward curves which may be produced by M are contained within the family χ.

Yield Curve
The term structure of interest rates, specifically an expression of the interest rates that are applicable by term outstanding. These can be expressed in a number of different ways, including zero coupon rates, par rates, etc.

Immunisation
Refers to the process of protecting oneself against interest rate risk by matching the duration of one's assets to liabilities.

1. INTRODUCTION

For many years life insurance companies have been selling annuity related products to their policyholders. Company pension funds have similarly been undertaking liabilities to pay pensions in the form of a defined benefit on retirement. As a result, both types of entity have exposed themselves to extremely long term interest rate risk. Unfortunately, the term of such risks often extends well beyond the longest point on the tradable yield curve. This creates serious problems for entities looking to hedge their interest rate risks.

In South Africa, various entities have tried to follow an immunisation approach to hedge their interest rate risk. Other entities have attempted to split their liability into buckets (i.e. grouping by term) and immunise each bucket separately. Few entities have opted for a respectable derivative based strategy to hedge such exposures.

While it is clear that traditional immunisation is only partially effective as it offers little protection against non-parallel shifts in the yield curve, many entities have opted for a bucketing approach. Generally such an approach leaves an insurer / pension fund with less risk as it immunises groups of liabilities across different terms.

However, this is only true for a very limited scope of risks such as annuities and pensions in payment. More advanced risks such as guaranteed annuity options and pensions in their accumulation period cannot be adequately managed through the use of immunisation and bucketing. The key reason for this is that the interest rate risks associated with guaranteed annuity options and pensions in their accumulation period are often contingent and contain optionality. It can therefore be very difficult to apply a traditional approach to match liabilities when they are not certain. In such cases a well engineered derivative strategy could provide a good management tool to these risks.

Such a strategy could involve the use of a variety of financial instruments including swaps, caps / floors, swaptions, etc. Unfortunately these instruments are usually only available, with reasonable liquidity, out to a maximum term. In South Africa the maximum term is 30 years. In more liquid foreign markets, such as the UK and US, this term may be up to 50 years. This limitation provides a significant barrier to the development of adequate derivative based risk management strategies.

A further problem arises from non-observability beyond the maximum term of the yield curve. This makes it exceedingly difficult to quantify the extent of an entity′s interest rate risk beyond this term.

In light of the above discussion it can be seen that three primary problems have emerged:

1. The inadequacy of traditional matching methods (i.e. immunisation and bucketing) to cope with the long term interest rate risks on life insurers' and pension funds' balance sheet.

2. The non-observability of interest rate data beyond the maximum term in the yield curve. Associated with this is the inability to adequately quantify interest rate risk.

3. The lack of liquidity in long term interest rate markets. Associated with this is the inability to adequately hedge interest rate risk.

Therefore, as a result of the above problems, this dissertation will aim to achieve the following tfour objectives:

1. Research the extent of work that others have performed related to forecasting and hedging long term interest rates.

2. Explore various methods of quantifying long term interest rate risks. This is intended to focus specifically on the yield curve. An extension of this research could possibly focus on performing a similar study of implied interest rate volatilities.

3. Explore the implementation of alternative hedging strategies for long term interest rate risks. This dissertation aims to focus on relatively simple interest rate risks in order to clearly establish a theoretical framework.

4. Compare and contrast the efficiency and adequacy of the proposed strategies with traditional strategies.

In order to achieve these objectives, the dissertation will be structured as follows:

In the first part of Chapter 2 we perform a literature study of the traditional methods used to hedge interest rate risks. We describe the concepts of immunisation and duration bucketing, along with duration vectors and the M-square measure. The application of these concepts to interest rate risk management is then considered, and we describe why these approaches are not adequate for managing long term risks beyond the maximum tradable term.

In the second part of Chapter 2 we perform a literature study of some modern methods to forecast and hedge long term interest rate risks. We discuss the emergence of the concept of stochastic duration, including the role of principal component analyses in interest rate risk management. Some of the functional form approaches to interest rate modelling are also discussed, along with relevant associated research. Further, we highlight a lesser-known interest rate model, the Smith-Wilson model.

In Chapter 3 we propose various yield curve extrapolation procedures based on the discussions in the literature study. Calibration of these procedures is also discussed.

In Chapter 4 we derive generic results relating to the proposed extrapolation procedures in Chapter 3. These results will be used to derive theoretical hedge portfolios for the various extrapolations.

In Chapter 5 we perform our first case study. The generic results of Chapter 4 are used to derive theoretical hedges over a historical period from October 2001 to March 2007. We track the weekly performance of the various extrapolation procedures when used to forecast and hedge a theoretical 50 year zero coupon bond. The results of the exercise are used to draw conclusions regarding performance of the various approaches. In addition, we perform an extension of the case study by applying the same exercise to a 35 year zero coupon bond.

In Chapter 6 we perform our second case study. This is done by performing a similar exercise to Chapter 5, except we perform the analysis on out-of-sample / simulated yield curve data, rather than historical data.

2. LITERATURE REVIEW: TRADITIONAL AND MODERN METHODS OF MANAGING INTEREST RATE RISK

2.1 *Traditional Methods for Managing Interest Rate Risk*

The two most common forms of interest rate protection tools are Immunisation and Duration Bucketing.

Immunisation relies on the concept of Macaulay duration as introduced by Macaulay (1938). The weaknesses of this concept have been well publicised, and it has been shown that the measure only works well when the shifts in the yield curve are parallel and small. Hence it does not cope well with more complicated movements such as twists or inversions. However, many entities still base their liability matching (hedging) strategies upon this measure.

Duration bucketing is another technique that is commonly adopted by various entities for liability matching. It involves dividing a profile of liability cashflows into buckets by term. This approach will also be discussed along with its performance and weaknesses.

Another method that will be discussed is the M-squared method as covered by Fong and Vasicek (1984).

References include: Agca(2002), Bierwag (1977), Bierwag (1978), Bierwag (1983), Boyle (1980), Ingersoll (1978), Redington (1952)

2.1.1 *Immunisation*

Immunisation is a technique whose development has been accredited to Redington (1952), although the concept of duration was originally developed by Macaulay (1938). Hicks (1946) also developed a similar concept yet all three authors seemed to reach their conclusions independently.

Redington identified the concepts of Duration (or mean term) and Convexity that can apply to any set of known future cash flows. Each concept can be described as follows:

Given that we are currently at time t_0; suppose we have a set of future cash flows described by $(C_{t_1}, C_{t_2}, \ldots, C_{t_n})$, where $t_0 < t_1 < t_2 < \ldots < t_n$. Suppose that the term structure of interest rates is a flat i% per annum. Then

$$Duration_C = D_C = \frac{\sum_{s=0}^{n} C_{t_s} \times t_s \times (1+i)^{-t_s}}{\sum_{s=0}^{n} C_{t_s} \times (1+i)^{-t_s}}, \tag{2.1}$$

$$Convexity_C = C_C = \frac{\sum_{s=1}^{n} C_{t_s} \times t_s^2 \times (1+i)^{-t_s}}{\sum_{s=0}^{n} C_{t_s} \times (1+i)^{-t_s}}. \tag{2.2}$$

Now suppose we have a set of liability cash flows $(L_{t_1}, L_{t_2}, \ldots, L_{t_n})$. Then Redington shows that by choosing an asset cash flow profile $(A_{t_1}, A_{t_2}, \ldots, A_{t_n})$ such that $D_A = D_L$, and $C_A > C_L$, then for any small change in the level of i%, the balance sheet position given by the change in (A - L) will always generate a surplus. Intuitively this approach equates to matching the duration of asset and liability cash flows, while keeping the dispersion of asset cash flows (around the duration) greater than that of the liability cash flows.

The most prominent limitation of this approach to interest rate hedging is that it will only provide protection against small, parallel shifts in the term structure of interest rates. Other key weaknesses include the assumption of a flat term structure, as well as the implicit assumption that arbitrage profits can be made through the maintenance of Reddington's second condition.

Much work has subsequently been done to generalise Redington's approach and overcome inherent weaknesses in the concept of immunisation. The following section provides a brief summary of this work:

Generalisation of the Immunisation Approach Fisher and Weil (1971) performed a generalisation of Redington's approach by assuming a non-flat term structure of interest rates. Under this approach a new duration measure was defined where each cashflow was discounted by its spot yield on the term structure. The measure was almost identical to the Redington measure except each cash flow is discounted with respect to its associated spot rate. Therefore, suppose that the spot rate for an outstanding term of t time units is i_t, where $t > 0$. Then the measure can be expressed as follows:

$$D_C^F = \frac{\sum_{s=0}^{n} C_{t_s} \times t_s \times (1+i_{t_s})^{-t_s}}{\sum_{s=0}^{n} C_{t_s} \times (1+i_{t_s})^{-t_s}}. \tag{2.3}$$

Even though duration matching under this measure accounted for a non-flat term structure of interest rates, it still required unexpected shifts in the term structure to be additive.

Bierwag (1977) then extended the approach to allow for multiplicative unexpected shifts in the term structure. He derived a duration measure D^B which is implicitly defined as follows:

$$r(D_C^B) = \frac{\sum_{j=1}^{n} C_{t_j}.(\frac{t_j}{q}).i_{t_j}.(1+i_{t_j})^{-t_j}}{\sum_{j=1}^{n} C_{t_j}.(1+i_{t_j})^{-t_j}}. \tag{2.4}$$

In the above equation q is defined as the "planning period." Bierwag also examined the

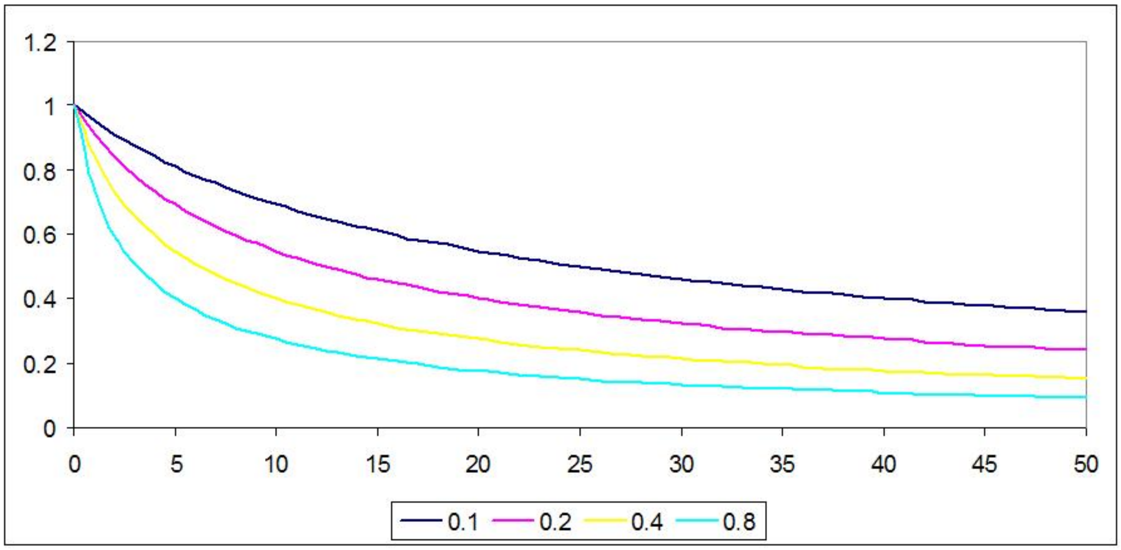


Fig. 2.1: Graph of Khang Δs_t for varying α

meaning of duration where the interest rate process is both multiplicative and additive.

Various researchers have tried to improve the duration measure to provide protection against non-parallel shifts in the term structure of interest rates. Khang (1979) extended the approach to allow for the case when short term interest rates fluctuate more than long term interest rates. According to this approach, Khang hypothesised that each spot rate (s_t) in the term structure would change subject to the following general formula:

$$s(t)^+ = s(t) + \frac{\lambda}{\alpha t} ln(1 + \alpha t). \tag{2.5}$$

In the above formula, the α parameter quantifies the ratio of changes in short term interest rates to changes in long term rates, and must be chosen subject to the underlying interest rate process. The Δs_t function is shown in Figure 2.1 for varying levels of α. In reality α this would need to be estimated from empirical evidence. Khang then specified that the measure of duration for such a process would be D^K such that:

$$ln(1 + D_C^K) = \frac{\sum_{j=1}^{n} C_{t_j}.ln(1 + \alpha t_j).(1 + i_{t_j})^{-t_j}}{\sum_{j=1}^{n} C_{t_j}.(1 + i_{t_j})^{-t_j}}. \tag{2.6}$$

Bierwag, Kaufman and Toevs (1983) researched the use of duration matching to immunise multiple liability cash flows. They also reconciled the concept of immunisation back to general equilibrium theory. Their research concluded that Redington's convexity condition is not necessary to achieve immunisation. This is because no-arbitrage opportunities exist under general equilibrium.

The research produced by Khang, Bierwag and many others provided a natural evolution of the duration matching framework to many of the more advanced interest rate hedging theories which will be covered later in this dissertation.

2.1.2 Duration Bucketing

Duration Bucketing is an approach used in practice that is based on the same principle as immunisation. The key difference between the two approaches is that Duration Bucketing attempts to "group" liability cashflows by term into a number of "liability buckets" then immunise each bucket separately. Therefore the approach is a simple extension of the immunisation approach and it is intended to provide some protection against non-parallel term structure shifts.

In practice the liability cash flows will need to be analysed and a time-horizon will need to be identified for the group. This horizon will be defined as (t_{min}, t_{max}), where t_{min} is the term until payment of the earliest liability cash flow and t_{max} is the term until payment of the latest liability cash flow. These terms will usually be changed as follows:

- t_{min} may be rounded to the nearest earlier month or year.
- t_{max} may be rounded to the nearest later month or year.

This horizon is then divided up into a number of sub-intervals (t_0, t_1), (t_1, t_2), ... , (t_{n-1}, t_n), such that $t_0 <= t_{min}$ and $t_n >= t_{max}$. In theory these sub-intervals should be chosen such that interest rate risk is relatively homogenous across each sub-interval. Hence it is intended that there should be high correlation in the size and direction of unexpected changes in all spot rates within each sub-interval. The number and size of the sub-intervals would be determined taking into account the following factors:

1. Historical estimates of the correlation structure underlying the term structure of interest rates.
2. Forecast changes in the future correlation structure underlying the term structure of interest rates.
3. Projected volatility of the term structure.
4. Risk appetite of the entity using the approach. Ultimately a lower risk appetite would usually lead to a greater number of smaller sub-intervals.

Once the liability cash flows have been grouped into their respective buckets, the standard immunisation technique is applied to each bucket. Note that in the limit this approach reduces to matching each liability cash flow separately.

This approach is able to provide increased protection against both parallel and non-parallel shifts in the term structure. However, the approach has a number of weaknesses:

1. It can be expensive to administer as it may require more frequent trading of more instruments than the standard immunisation approach.

2. Within each sub-interval the approach offers little protection against non-parallel shifts in that section of term structure, even though there is likely to be increased protection across all intervals.

3. There will often exist relatively strong correlations between different sub-intervals. By trying to hedge each interval separately no account is being taken for this cross-interval correlation, which suggests that this matching strategy is not optimal as it does not take account of all information in the term structure.

4. This approach describes a principle for hedging, however it still remains for the user to pick the specific assets to use for hedging. For example, if a user was to select their hedge for each sub-interval from a universe of two bonds (i.e. one very long bond and one very short bond), then there is no guarantee that this approach will result in better performance than the standard immunisation approach.

5. If this approach is used for hedging a dynamic set of liability cashflows over a given period of time, then the duration of each sub-group of assets will not be a smooth function of time. This will be caused by assets creeping across sub-intervals as their maturity dates shorten.

One way of ensuring that appropriate assets are chosen per sub-interval is to impose the requirement that the term of any asset used to immunise a sub-interval should fall within the sub-interval. However, this accentuates the problem of discontinuous asset-group durations per sub-interval, and often such assets may not always be easily available when the sub-intervals are defined narrowly.

Interestingly, the concept of Duration Bucketing is closely related to the more advanced *Key Rate Duration* originally described by Ho (1992) and recently explored by Poitras (2005). The key similarity between the two approaches is that they both try to segment sections of the term structure and implicitly assume that the interest rate risk within each segment can be described by a single factor.

2.1.3 M-Squared

The M-Squared (or M^2) measure, as defined by Fong and Vasicek (1984), was derived as a tool to assist in selecting the best duration matching portfolio from a set of potential portfolios. The key result derived by Fong and Vasicek was as follows:

Suppose that:

- We have a set of future cash flows described by $(C_{t_1}, C_{t_2}, , C_{t_n})$, where $t_0 < t_1 < t_2 < \ldots < t_n$,
- We are trying to hedge these cashflows with respect to a specific time horizon (H),
- The present value of these cashflows at t_0 is described as $P_C = \sum_{s=0}^{n} C_{t_s} \times (1 + i_{t_s})^{-t_s}$,
- Forward rates change instantaneously from f_t to $f_t^n = f_t + \Delta f_t$, where Δf_t may

be an arbitrary function of term(t).

Define $P_C^H = P_C \times (1 + i_{t_H})^H$. Then the following theorem (Fong and Vasicek (1984)) holds:

Theorem 2.1.1: Let K be an arbitrary constant. If $\frac{\delta i(t)}{\delta t} \leqslant K$ for all $t > 0$, then

$$\frac{\Delta P_C^H}{P_C^H} \geqslant -\frac{1}{2} \times K \times M^2, \tag{2.7}$$

where

$$M^2 = \frac{\sum_{j=1}^{n} (t_j - H)^2 \times C_{t_j} \times (1 + i_{t_j})^{-t_j}}{P_C}. \tag{2.8}$$

There are a number of things worth noting from this result:

1. Equation (2.7) provides a lower bound on the change in the risk-neutral expected future value of the portfolio for any given time horizon. It relies on the assumption that any changes in the term structure forward rates will be a smooth function of term. However, they hypothesised that beyond this there is no reliance on any assumptions regarding the nature or dimensionality of the interest rate process.

2. This lower bound is made up as the product of two terms: $-\frac{1}{2}K$, and M^2. The first term depends only on the interest rate process while the second term depends only on the initial term structure and the structure of the cashflow profile.

3. K represents the upper bound on the change in the slope of the term structure forward rates, with respect to term. As such it provides a measure of the extent to which the yield curve can twist.

4. Once K has been specified, it immediately follows that M^2 is a direct measure of the sensitivity of the structure to interest rate movements. Therefore M^2 logically follows as a tool which can be used for risk measurement.

5. Even though the investor may not have control over the K factor, the investor does have control over the structure of the cash flow profile and hence has direct control over the quantity M^2. Hence a risk-averse investor could build their structure with the aim of minimising M^2 as this would ensure reduced sensitivity of the structure to interest rate movements.

6. It is interesting to see that while duration is the weighted average of time to payments on a structure, M^2 is a similarly weighted variance of the time to payment around the horizon, where the weighting factors are the present value of each payment.

The development of the M^2 measure has provided risk managers with an additional tool for selecting an optimal hedging strategy for a given liability profile. Typically, the choice of a hedge would be an optimisation problem expressed as follows:

From the horizon of available assets, select an asset profile A such that:

1. $D_A = D_L$,

2. M^2 *is minimised, where* $H = D_L$.

An alternative formulation could simply be to minimise M^2 where $H = D_L$.

Example Suppose that a portfolio manager has an obligation to pay R100m in 1.5 years. In order to hedge this obligation he has available to him, 3 zero coupon bonds of varying terms, namely 0.5 years, 1 year and 2 years. There are no restrictions on the amount he can invest in each bond. Interest rates are initially a level 5% p.a., though movements in term structure are not necessarily parallel.

Denote $P^C = \sum_{i=1}^{3} C_{t_i}(1+i_{t_i})^{-t_i}$. Under the M^2 framework, the hedge portfolio is determined from the following three requirements:

$$1.\ P^C = 100\,000\,000(1+i_{t_i})^{t_i},$$

$$2.\ \sum_{i=1}^{3} t_i \frac{C_{t_i}(1+i_{t_i})^{-t_i}}{P^C} = 1,$$

$$3.\ \textit{Minimise}\ M^2.$$

If we assume that $C_{t_i} \geqslant 0$, then we find that the solution to this problem is given by the case where $C_{t_1} = 0$:

Parameter	*Solution*
$C_{0.5}$	0
C_1	48.795m
C_2	51.235m
M^2	0.25

Tab. 2.1: M^2 Example - Solution for $C_{t_i} \geqslant 0$

The Duration Puzzle Further research of the M^2 approach has been performed by a number of contributors. Work done by Ingersoll (1983), Bierwag et al. (1993) examined what has become known as "the duration puzzle." This is based on the argument that M^2-minimising portfolios (without a maturity matching bond) perform worse than portfolios containing a maturity-matching bond, and has been supported by empirical evidence. Bierwag et al. (1993) went further to show that minimising M^2 is not independent of the underlying stochastic process, as had been assumed in previous research.

2.1.4 Duration Vector

The Duration Vector approach was documented by Chambers, Carleton and McEnally (1988). This approach is rooted in traditional immunisation theory as described by Redington. This approach calculates an infinite vector of partial derivatives, where the k-th element represents the k-th partial derivative of the value of the liability stream with respect to the interest rate, divided by the liability value. In order to achieve a matched position this approach elects a dimension (k) to which matching must occur. It then goes about to select assets (bonds) which replicate (in aggregate) the elements of the infinite vector to the k-th order, such that the total value of assets (A) equals the total value of liabilities(L).

Therefore traditional duration matching is simply a special case of the duration vector approach where k = 1, i.e. traditional duration matching simply involves choosing assets such that $\frac{\partial A}{\partial i}/A = \frac{\partial L}{\partial i}/L$, where A = L.

Under the assumption that the term structure of continuously compounded interest rates can be expressed as a polynomial, Chambers and Carleton (1981) demonstrate that the finite and non-instantaneous return of a default free bond can be expressed as a dot product of a duration vector and a shift vector. More specifically, let P_t denote the price of a zero coupon bond with maturity t, then they show that:

$$\frac{P_{t_{s+1}}}{P_{t_s}} = k_{t_s;t_{s+1}} + \sum_{w=1}^{\infty} D_{t_s}(w).q_{t_s;t_{s+1}}(w), \tag{2.9}$$

where

- $D_{t_s}(w) = \sum_{j=1}^{n} \frac{C(t_j).B_{t_s}(t_j)}{P_{t_s}}(t_j - 1)^w$,
- $C(\bullet)$ represents the series of cashflows on the respective bond,
- $B_{t_s}(T)$ represents the price at time t_s of a bond maturing at time T,
- $k_{t_s;t_{s+1}}$ is the return on a zero coupon bond from time t_s to maturity at time t_{s+1},
- q(w) is a random variable containing information regarding the term structure shift from time t_s to t_{s+1}.

Note that this equation assumes no cashflows occur in the period $(t_s; t_{s+1})$. Note also that the measures $D_{t_s}(1)$ and $D_{t_s}(2)$ are closely related to the traditional convexity and duration measures. The key difference is that term minus 1 is used in the calculation.

Therefore, measures of $D(\bullet)$ for $w \geqslant 1$ describe the one-period return component arising as a result of level shifts in the term structure. Measures for $w \geqslant 2$ describe the one-period return component arising from changes in the slope of the term structure, while measures for $w \geqslant 2$ describe the one-period return component arising from changes in the curvature of the term structure.

In terms of hedging, this theory requires that a hedging portfolio should be set up which has equivalent $D(\bullet)$ measures to the liability portfolio. In theory a hedging portfolio would identically replicate the movements in value of the liability portfolio if all of its duration measures exactly matched to those of the liability portfolio at all times. This is because the return on the liability portfolio, as described by the above equation, can be written as a function of the dot product of the duration vector and a general function of the change in term structure. Provided that the term structure could be expressed at all times as a polynomial function of time, the hedge would work regardless of the underlying interest rate process.

One practical problem with using this approach for hedging is that it requires the duration measures to be matched up to an infinite order. However, Chambers, Carleton and McEnally (1988) perform empirical testing on this model and find that the equation holds relatively closely over shorter ranges of summation. They test the effectiveness of the hedge when matching $D(w)$ up to the order of n, for n = 1,2...9. A 4 year period of quarterly returns is used. It is found that matching to a higher order improves effectiveness of the hedge. The results of their analysis indicate that matching beyond the 5th order begins to add marginally little benefit. However, they point out that this result should be applied to a fairly simple liability. Hedging more complicated interest rate derivatives with this approach may require duration matching out to higher orders.

2.1.5 Other Uses of Duration

Various papers have investigated the use of duration outside of immunisation and risk management. Durand(1974) examined the potential integration of duration with profitability analyses in capital budgeting. Blocher and Stickney (1978) examined the effect of changes in the firm's required rate of return on the duration of its projects and suggested project selection rules appropriate for the project manager's risk tolerance. Bierwag and Khang (1978) showed that a duration-derived immunisation strategy is optimal where an investor's preferences are adequately described by Fishburn's (1977) measure of risk. Tito and Wagner (1977) suggested evaluating pension fund managers on the basis of portfolio return for a given duration. Keintz and Stickney (1977) considered various problems with using duration concepts to immunise pension fund interest rate risks.

2.1.6 Traditional Approaches for Managing Life Insurance / Pension Fund Risks

The development of the duration and immunisation concepts were extremely important steps forward in level of thinking regarding interest rate risk. Their simplicity has made them easily practicable and popular in bond portfolio management strategies. However, it has been necessary to adapt the approaches to overcome various inherent weaknesses.

Adaptations of the duration concept have included the Duration Bucketing and M-square approaches as described above. However, a key reliance that still underpins these approaches is the ultimate certainty of (liability) cash flows being managed. As soon as we move into the context of most life insurance companies and pension funds, fixed-interest liabilities are no longer necessarily "fixed," i.e. they may be contingent. An example of such a risk is a guaranteed annuity option (GAO) which provides a policyholder with the option to convert the maturity proceeds on their savings policy into an annuity at a guaranteed rate. Ignoring demographic risk, this is effectively an option on interest rates which strikes when interest rates drop below a specified level. It would be difficult to effectively hedge these risks using the traditional techniques described as they rely on a high level of certainty in the liability being hedged. Wilkie et al. (2003) provide a more comprehensive definition of guaranteed annuity options. Pelsser (2003) describes the nature of the interest rate risk inherent in guaranteed annuity options.

A further problem which undermines these traditional approaches described relates to liquidity and observability of interest rates. These traditional approaches implicitly make the assumption that interest rates are observable and tradable at all relevant terms. In markets such as those in South Africa, where there is little or no liquidity in fixed interest instruments beyond 30 years, it is sometimes necessary to have a means of forecasting interest rates at terms where these are not directly observable. This is particularly true in the case of life insurers and pension funds who often have interest rate exposures well beyond the tradable term of 30 years.

On the basis of the above evidence we can see that it is necessary to consider some more modern approaches to hedging interest rate risk.

2.2 Modern Methods of Managing Interest Rate Risk

This section covers various methods for hedging interest rate risk. These methods have been researched / developed relatively recently in comparison to the traditional duration type methods discussed in the previous section. In addition, these models represent a step forward in the level of thinking around interest rate risk, particularly because they are all explicitly based on models of yield curve behaviour.

The approaches that will be discussed include the Stochastic Duration, Functional Form, the Smith-Wilson and the Principal Component approaches.

2.2.1 Stochastic Duration

This approach was highlighted by Boyle (1980), however the approach had previously been suggested by Ingersoll, Skelton and Weil (1978). In principle, we begin with the assumption that it is possible to identify the factors driving changes in the yield curve. We then derive the sensitivity of our hypothetical interest rate risks to each of the driving factors. In order to hedge the risks, it then becomes necessary to find a hedging portfolio which generates equal and opposite sensitivities to those of our interest rate risk.

At the time that Boyle's paper was written, duration had become a popular tool in interest rate risk management. Therefore, the key outcome of this paper was not to derive an innovative, new interest rate risk measure, but rather to illustrate that duration is an inappropriate measure of risk.

Non-infinitesimal uniform shifts in the Yield Curve Ingersoll, Skelton and Weil (1978) refute much of the work performed by previous authors regarding the meaningfulness of duration for risk measurement. They prove that duration is meaningful for non-flat term structures only when changes in the yield curve are infinitesimal and of uniform proportional magnitude. Therefore additive, uniform, non-infinitesimal shifts in a non-flat term structure, assumed by many previous authors, cannot occur in a competitive equilibrium. This invalidated duration as an adequate measure of risk for all such models of the term structure. Further, they go on to develop a measure of risk which is consistent with competitive equilibrium in the case of a non-flat term structure. This is done by assuming interest rate dynamics to follow the process:

$$dr = \mu(r)dt + \sigma(r)d\phi, \tag{2.10}$$

where:

- r is the short rate process,
- ϕ is a continuous time Poisson process with parameter λ,
- σ represents the size of the random shock to which the spot rate is exposed, given that such a shock occurs.

For the simple case where $\sigma(t)$ is a constant $\sigma > 0$, the appropriate measure of risk is shown to be:

$$\frac{\sum_{s=0}^{n} C_{t_s} \times (1 - e^{-\sigma t}) \times (1 + i_{t_s})^{-t_s}}{\sum_{s=0}^{n} C_{t_s} \times (1 + i_{t_s})^{-t_s}}. \tag{2.11}$$

The primary differences to the traditional duration measure are:

1. Risk increases with maturity of the cash flow at less than a linear rate. This is because longer intervals of time will be exposed to more shocks, but the average number of shocks per period will be more tightly distributed around the expected amount. Shorter intervals would have greater volatility in the average number of shocks per period.

2. Where traditional duration is measured in units of time, this measure does not have an easily identifiable dimension.

The second difference can be resolved by defining the stochastic duration to be as follows:

$$D^S = -\frac{1}{\sigma} ln[1 - \frac{\sum_{s=0}^{n} C_{t_s} \times (1 - e^{-\sigma t}) \times (1 + i_{t_s})^{-t_s}}{\sum_{s=0}^{n} C_{t_s} \times (1 + i_{t_s})^{-t_s}}]. \tag{2.12}$$

In the case where $\mu = 0$, $\lambda(r) = \lambda r$, and σ constant, then the appropriate measure of duration is:

$$D^{S1} = \frac{\sum_{s=0}^{n} C_{t_s} \times A(t_s) \times (1 + i_{t_s})^{-t_s}}{\sum_{s=0}^{n} C_{t_s} \times (1 + i_{t_s})^{-t_s}} - 1, \tag{2.13}$$

where:

$$A(T) = \frac{1 + \lambda}{\lambda + e^{\sigma(1+\lambda)T}}. \tag{2.14}$$

Continuous time, stochastic yield curve movements Continuous time, stochastic modelling of the term structure had been researched by Vasicek (1977), Cox, Ingersoll and Ross (1977, 1985), and Brennan and Schwartz (1977). Boyle (1980) reconciled the ideas developed by these thinkers back to the immunisation framework. As had been the case with many previous researchers, his work pointed out a series of weaknesses with the use of duration as a risk management tool.

Going back to the original thinking behind Macaulay's duration measure, we can say that the definition of duration could be expressed as follows:

$$D = -\frac{\frac{\partial A}{\partial r}}{A},$$

where:

- A represents the present value of a series of nominal future cash flows,
- r is the short rate,
- The term structure is assumed flat such that i(t) = r, and r is the only risk factor.

Hence, provided that r is the only source of volatility in the term structure, immunisation could be achieved for a set of liability cash flows by having:

$$\frac{\partial A}{\partial r} = \frac{\partial L}{\partial r}, \tag{2.15}$$

$$A = L, \tag{2.16}$$

where

- A represents the value of asset cash flows,
- L represents the value of liability cash flows

Viewed in this light, it seems natural to relax the assumption of a flat term structure and reformulate the definition of duration from first principles. Boyle's research adopted this exact approach and compared the results of the two approaches where the term structure followed Vasicek and Cox-Ingersoll-Ross processes.

Vasicek Process Under the Vasicek model of interest rates, the short rate process is assumed to be:

$$dr = \alpha(\gamma - r)dt + \rho dW_t, \tag{2.17}$$

The term structure, defined in terms of zero coupon bond prices, is decribed by:

$$P(t, s, r) = e^{F(\alpha,T)(D-r)-TD-\frac{\rho^2}{4\alpha}(F(\alpha,T))^2}, \tag{2.18}$$

where

- T = (s - t),
- $F(\alpha, T) = \frac{1}{\alpha}(1 - e^{-\alpha T})$,
- $D = \gamma - \frac{1}{2}\frac{\rho^2}{\alpha^2}$.

Boyle then uses the pricing function above to calculate:

$$\frac{\partial P}{\partial r} = -PF. \tag{2.19}$$

Hence for a set of cash flows C, we have:

$$-\frac{\frac{\partial P_C}{\partial r}}{P_C} = \frac{\sum_{s=0}^{n} C_{t_s} \times F(\alpha, t_s - t_0) \times (1 + i_{t_s})^{-t_s}}{\sum_{s=0}^{n} C_{t_s} \times (1 + i_{t_s})^{-t_s}}. \tag{2.20}$$

CIR Process Under the CIR model of interest rates, the short rate process is assumed to be:

$$dr = \kappa(\mu - r)dt + \sqrt{\sigma^2 r} dW_t. \tag{2.21}$$

The term structure, defined in terms of zero coupon bond prices, is decribed by:

$$P(t, s, r) = A(T)e^{-r.B(T)}. \tag{2.22}$$

where

- T = (s - t),
- $A(T) = [\frac{2\lambda e^{(\kappa-\lambda)\frac{T}{2}}}{(\lambda+\kappa)(1-e^{-\lambda T}+2\lambda e^{-\lambda T})}]^{\frac{2\kappa\mu}{\sigma^2}}$,
- $B(T) = [\frac{2(1-e^{-\lambda T})}{(\lambda+\kappa)(1-e^{-\lambda T})+2\lambda e^{-\lambda T}}]$,
- $\lambda^2 = \kappa^2 + 2\sigma^2$.

Boyle then uses the pricing function above to calculate:

$$\frac{\partial P}{\partial r} = -PB. \tag{2.23}$$

Hence for a set of cash flows C, we have:

$$-\frac{\frac{\partial P_C}{\partial r}}{P_C} = \frac{\sum_{s=0}^{n} C_{t_s} \times B(t_s - t_0) \times (1 + i_{t_s})^{-t_s}}{\sum_{s=0}^{n} C_{t_s} \times (1 + i_{t_s})^{-t_s}}. \tag{2.24}$$

Boyle then proceeded to derive hedges for various interest rate risks based on these risk measures and compared his results with those derived using the traditional duration measure. His results indicated that the potentially accurate hedging portfolio could not be derived using traditional immunisation, as this tended to underweight the importance of having long dated cash flows in the hedging portfolio. The key weakness in his research was that he had only used a one factor model for describing the interest rate process. However, he had successfully taken the concept of immunisation back to its fundamentals and expressed the problem as one of hedging interest rate exposures based on their driving risk factors (or principal components). This was subtly different from much of the work in the preceding 40 years, which had either tried to generalise the concept of duration in terms of non-parallel term structure shifts, or had tried to impose additional ad-hoc requirements into the original immunisation framework. In this respect Boyle's paper was one of the first of its kind that became well recognised, and helped to change the thinking around interest rate hedging. In research conducted since Boyle's paper, the trend in thinking around interest rate risk has moved toward a risk factor (or principal component) type approach.

Stochastic Duration Extended into Multiple Factors Various papers have subsequently been produced which expand the stochastic duration theory. Generally speaking, any paper which has succeeded in deriving an analytical (or even numerical) expression for the term structure, based on an initial assumption about the process governing the term structure, has added to the stochastic duration theory. This is because stochastic duration is simply the sensitivity of price changes with respect to each of the inputs in the term structure process. Longstaff and Schwartz (1992) produced a 2-factor equilibrium model of the term structure and showed that the level and volatility of the short rate could be expressed as a function of these factors, where the function is invertible:

They start by assuming that realised returns (Q) on the market portfolio are governed by the SDE:

$$\frac{dQ}{Q} = (\mu X + \theta Y)dt + \sigma\sqrt{Y}dZ_1, \tag{2.25}$$

where X and Y are state variables governed by the following equations:

$$dX = (a - bX)dt + c\sqrt{X}dZ_2, \tag{2.26}$$

$$dY = (d - eY)dt + f\sqrt{Y}dZ_3. \tag{2.27}$$

Wealth is described by the following equation:

$$dW = W\frac{dQ}{Q} - Cdt,$$

where C represents consumption.

It is then shown, subject to assumptions regarding utility, that the price of a contingent claim (F) satisfies the following partial differential equation:

$$\frac{x}{2}F_{xx} + \frac{y}{2}F_{yy} + (\gamma - \delta x)F_x + (\eta - \xi y - (-\frac{J_{WW}}{J_W})COV(W,Y))F_y - rF = F_\tau, \tag{2.28}$$

where $x = \frac{X}{c^2}$, $y = \frac{Y}{f^2}$, $\gamma = \frac{a}{c^2}$, $\delta = b$, $\eta = \frac{d}{f^2}$, and $\xi = e$.

It is then shown that risk free rate (r) can be written as:

$$r = \alpha x + \beta y, \tag{2.29}$$

where $\alpha = \mu c^2$, $\beta = (\theta - \sigma^2)f^2$.

In order to obtain an invertible relationship from a set of unobservable variables (x and y) to a set of observable variables; the volatility of the short rate is defined as:

$$v = E[(r - E[r])^2]. \tag{2.30}$$

This translates to the following equation:

$$v = \alpha^2 x + \beta^2 y. \tag{2.31}$$

Provided that $\alpha \neq \beta$, then the transformation of (x,y) to (r,v) is invertible such that:

$$x = \frac{\beta r - v}{\alpha(\beta - \alpha)}, \tag{2.32}$$

$$y = \frac{v - \alpha r}{\beta(\beta - \alpha)}. \tag{2.33}$$

The dynamics of r and v can then be written as:

$$dr = (\alpha\gamma + \beta\eta - \frac{\beta\delta - \alpha\xi}{\beta - \alpha}r - \frac{\xi - \delta}{\beta - \alpha}v)dt + \alpha\sqrt{\frac{\beta r - v}{\alpha(\beta - \alpha)}}dZ_2 + \beta\sqrt{\frac{v - \alpha r}{\beta(\beta - \alpha)}}dZ_3, \tag{2.34}$$

$$dv = (\alpha^2\gamma + \beta^2\eta - \frac{\alpha\beta(\delta - \xi)}{\beta - \alpha}r - \frac{\beta\xi - \alpha\delta}{\beta - \alpha}v)dt + \alpha^2\sqrt{\frac{\beta r - v}{\alpha(\beta - \alpha)}}dZ_2 + \beta^2\sqrt{\frac{v - \alpha r}{\beta(\beta - \alpha)}}dZ_3. \tag{2.35}$$

Solving this set of equations yields a term structure which is described by the following analytical expression for the price of a riskless zero coupon bond of term τ:

$$F(r, v, \tau) = A^{2\gamma}(\tau)B^{2\eta}(\tau)e^{\kappa\tau + C(\tau)r + D(\tau)v}, \tag{2.36}$$

where

$$A(\tau) = \frac{2\phi}{(\delta + \phi)(e^{\phi\tau} - 1) + 2\phi},$$
$$B(\tau) = \frac{2\psi}{(\nu + \psi)(e^{\psi\tau} - 1) + 2\psi},$$
$$C(\tau) = \frac{\alpha\phi(e^{\psi\tau} - 1)B(\tau) - \beta\psi(e^{\phi\tau} - 1)A(\tau)}{\phi\psi(\beta - \alpha)},$$
$$D(\tau) = \frac{\psi(e^{\phi\tau} - 1)A(\tau) - \phi(e^{\psi\tau} - 1)B(\tau)}{\phi\psi(\beta - \alpha)},$$

and

$$\nu = \xi + \lambda,$$
$$\phi = \sqrt{2\alpha + \delta^2},$$
$$\psi = \sqrt{2\beta + \nu^2},$$
$$\kappa = \gamma(\delta + \phi) + \eta(\nu + \psi).$$

Therefore:

$$\frac{\partial F}{\partial r} = CF,$$
$$\frac{\partial F}{\partial v} = DF.$$

In the one factor case the stochastic duration of a set of cashflows was taken as $-\frac{\frac{\partial P}{\partial r}}{P}$. However, in the two factor case the stochastic duration is represented by a vector where the elements are a measure of change in value with respect to each factor, i.e. $(-\frac{\frac{\partial P}{\partial r}}{P}, -\frac{\frac{\partial P}{\partial v}}{P})$.

Therefore, for a set of cashflows C, the stochastic duration can be written as:

$$(\frac{\sum_{s=0}^{n} C_{t_s} \times -C(t_s - t_0) \times (1 + i_{t_s})^{-t_s}}{\sum_{s=0}^{n} C_{t_s} \times (1 + i_{t_s})^{-t_s}}; \tag{2.37}$$

$$\frac{\sum_{s=0}^{n} C_{t_s} \times -D(t_s - t_0) \times (1 + i_{t_s})^{-t_s}}{\sum_{s=0}^{n} C_{t_s} \times (1 + i_{t_s})^{-t_s}}). \tag{2.38}$$

2.2.2 Principal Component Approach

Principal Component Analyses (Phoa (2000)) of the term structure of interest rates have become increasingly popular in the last decade or so. Various studies have been conducted across various economies to assess whether movements in individual yields can be explained by systematic changes in the term structure. The key objective is to investigate whether these systematic shifts tend to repeat themselves and assess to what extent they can explain all term structure movements.

The process of performing a principal component analysis is described by Phoa (2000). Suppose we have a matrix A, then $\underline{v}$ is an eigenvector of A with eigenvalue λ, if $A\underline{v} = \lambda\underline{v}$. In all cases the eigenvalue is unique, while the vector is a scalar multiple.

For A to be a correlation matrix, it is necessary that A is both symmetric and positive definite. Thus for any non-zero vector $\underline{w}$, we must have $\underline{w}A\underline{w} > 0$. By construction, each eigenvector of a matrix is orthogonal to all of the others. Each eigenvalue represents the relative weighting that measures the influence of that specific eigenvector.

The process for performing a principal component analysis is as follows:

1. Estimate the correlation matrix across the term structure of interest rates. This would typically be done using appropriate historical data. Note that we refer to a general correlation structure, therefore we could estimate the correlation structure of the absolute spot, forward or swap rates. Alternatively, we could estimate the correlation structure of the change in the spot, forward or swap rates.

2. Calculate the eigenvalues and eigenvectors of the correlation structure.

3. The eigenvectors can be interpreted as the fundamental yield curve shifts. The corresponding eigenvalues indicate the extent to which each of the shifts affects yield curve behaviour.

4. Remove any spurious eigenvectors and eigenvalues from the results to obtain the estimated principal components of the term structure. If there are any spurious eigenvectors, they will typically have very small eigenvalues relative to the larger eigenvalues in the results. This process of eliminating spurious eigenvectors is often subjective.

Assuming that the process governing yield curve movements is time-homogenous, the results of this analysis provide a reasonable amount of information regarding the process. Specifically, to the extent that the process can be defined as a set of independent factors which continuously affect the curve to a greater / lesser extent, then the eigenvectors will provide estimates of these factors and the eigenvalues will provide estimates of the relative strength of each factor.

It is worth mentioning that the derived principal components will only be estimates of the true underlying components. The above process may have eliminated eigenvectors which should not have been eliminated, or there may be estimation error in the levels of the derived principal components. Indeed, there may even have been fundamental economic shifts in the underlying estimation data which could confound the results of the analysis. Fortunately,

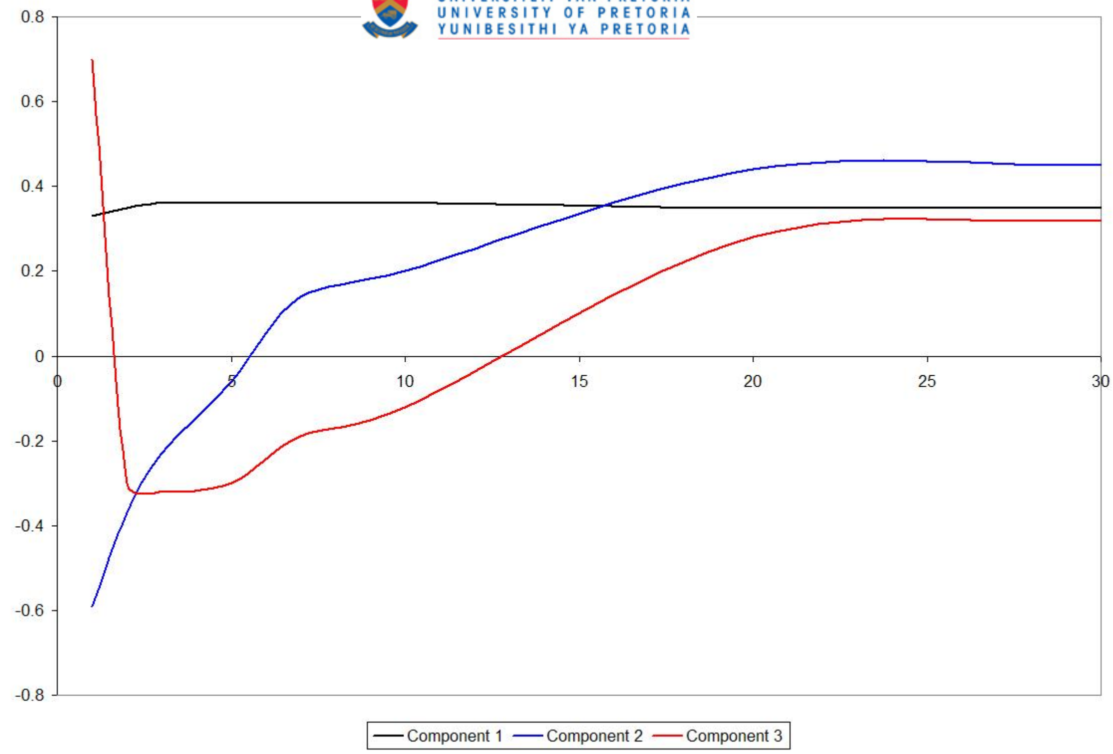


Fig. 2.2: Phoa(2000) principal component functions on US treasury term structure

most principal component analyses in liquid markets have been found to yield similar results which have been found consistent with macroeconomic theory. Therefore, any results which are substantially different would need to be justified on the basis of economic reasoning and more extensive historical testing.

Phoa (2000) performs a principal component analysis on actual US treasury bond data from 1993 to 1998. Table (2.2) below shows the results of his anaysis, while Figure (2.2) shows the principal components. The analysis concludes that the dominant shift in the spot yield curve is a parallel shift which explains over 90% of variability. The second most important shift is a slope shift in which short yields fall and long yields rise (or vice versa). The third most important shift is a curvature shift, in which short and long yields rise while medium yields fall (or vice versa). The remaining eigenvectors are deemed to be insignificant.

$Weight$	$1Year$	$2Year$	$3Year$	$5Year$	$7Year$	$10Year$	$20Year$	$30Year$
0.3%	0.00	0.05	-0.20	0.31	-0.63	0.50	0.32	-0.35
0.3%	0.00	-0.08	0.49	-0.69	0.06	0.27	0.30	-0.34
0.2%	0.01	-0.05	-0.10	0.25	0.30	-0.52	0.59	-0.48
0.4%	-0.05	-0.37	0.65	0.27	-0.45	-0.34	0.08	0.22
0.6%	0.21	-0.71	0.03	0.28	0.35	0.34	-0.27	-0.26
1.1%	0.70	-0.30	-0.32	-0.30	-0.19	-0.12	0.28	0.32
5.5%	-0.59	-0.37	-0.23	-0.06	0.14	0.20	0.44	0.45
91.7%	0.33	0.35	0.36	0.36	0.36	0.36	0.35	0.35

Tab. 2.2: Phoa(2000) principal component analysis of US treasury term structure

On the basis of a principal component analysis, it is often very tempting to conclude that by hedging the significant components of the curve one can generate a very high level of protection against interest rate movements. This can be true, especially during periods when the market is not showing extreme movements. However, as Phoa (2000) describes, this approach has a number of weaknesses:

1. When a principal component analysis results in a relatively large number of rejected components, it is likely that the rejected components (in aggregate) contribute significantly to the correlation structure. Therefore by excluding the less significant components one could be inferring a significantly different correlation structure.

2. A principal component analysis looks at the yield curve holistically and assumes that each factor has consistent relevance across the whole curve. However, this is not necessarily true because principal component analyses on sub-intervals of the term structure may yield different results to an analysis on the whole term structure. Therefore, if used for risk management there is a risk of idiosyncratic yield curve shifts occurring which would not have been picked up in the principal component analysis.

3. The approach assumes homogeneity in the process affecting the curve over time, which is not necessarily true. A change in this process may invalidate one's risk hedging process.

4. It is possible that a significant risk factor has been ignored by the approach, particularly one which acts rarely and has an extreme effect. By performing the analysis over a finite period one cannot be sure that all risk factors have been observed and captured in the analysis.

James Maitland (2002) performed a principal component analysis on South African government bond data from January 1986 to December 1998. The results indicated that approximately 92.8% of variability is explained by the first (level) component, 97.3% is explained by the first two (level and slope) components, while 98.4% is explained by the first three components (level, slope and bow). These results are similar to those obtained from Phoa (2000) above.

Niffiker, Hewins and Flavell (2000) perform a principal component analysis on the swap curves of 10 major currencies. Their results indicated that the first two factors (parallel and twist) explained between 97.1% and 98.6% of variation in the swap curves across the respective currencies. They then carried the analysis forward to propose a VAR calculation framework based on synthetic (empirically derived) factors driving yield curve behaviour.

2.2.3 Functional Form Approach

One of the earlier functional form papers was put forward by Cooper (1977), in which he summarised much of the previous work that had been done regarding functional form models of the term structure. In this paper Cooper assessed four previously suggested functional forms for the term structure and concluded that the three factor spot rate form of $R_t = e^{A+B\times t+C\times ln(t)}$ gave marginally better performance than the alternatives.

Perhaps the most significant work that has been done recently regarding functional forms was the establishment of the Nelson-Siegel framework by Nelson and Siegel (1987). This framework hypothesises that the term structure of interest rates is made up of three components: a level component, a slope component, and a bow component. Any potential systematic movement in the yield curve is stipulated as being a fuction of these three components. This approach has gained much support from economic circles due to its intuitive appeal. However, research performed by Bjork and Christensen (1999) and Filipovic (1999) and (2000) concludes that the Nelson-Siegel approach is not theoretically consistent as a function of time, i.e. it implies arbitrage opportunities. Subsequent work perfomed by Diebold and Li (2005) and Krippner (2006) has elevated the original Nelson-Siegel framework to a level that is intertemporally consistent while still remaining intuitively appealing.

Nelson-Siegel Approach Nelson and Siegel (1987) proposed the following model for forward rates at a given point in time (t):

$$f_t(\tau) = \beta_{1,t} + \beta_{2,t}e^{-\frac{\tau}{\lambda}} + \beta_{3,t}(\frac{\tau}{\lambda})e^{-\frac{\tau}{\lambda}},$$

where

- τ represents the term of the forward rate,
- $\beta_{1,t}$, $\beta_{2,t}$, $\beta_{3,t}$ represent time dependent stochastic variables,
- λ is a shape parameter.

This leads to the following specification for the spot curve at time (t):

$$s_t(\tau) = \beta_{1,t}h_{1,t}(\tau) + \beta_{2,t}h_{2,t}(\tau) + \beta_{3,t}h_{3,t}(\tau),$$

where

$$h_{1,t}(\tau) = 1,$$
$$h_{2,t}(\tau) = \frac{1 - e^{-\frac{\tau}{\lambda}}}{\frac{\tau}{\lambda}},$$
$$h_{3,t}(\tau) = \frac{1 - e^{-\frac{\tau}{\lambda}}}{\frac{\tau}{\lambda}} - e^{-\frac{\tau}{\lambda}}.$$

Therefore, the spot curve can be seen as a linear combination of three component functions with different shapes: a flat curve, a sloped curve, and a humped curve. These are depicted in Figure (2.3) for $\lambda = 0.179328$.

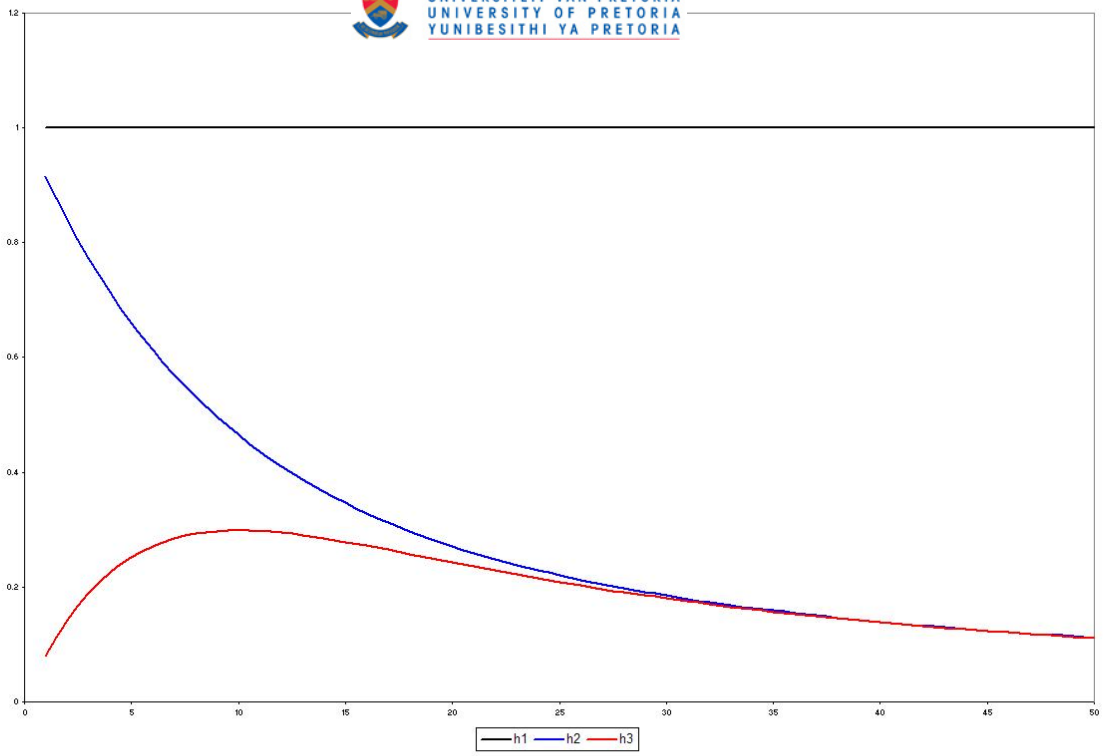


Fig. 2.3: Nelson-Siegel model component functions

This is a particularly attractive model as it can potentially produce a rich set of yield curves with relatively few parameters. This particular choice of λ results in maximum curvature at 10 years for component function h_3.

A more generalised version of the Nelson-Siegel model (de Pooter (2007)) specifies λ as a t-dependent parameter.

Svensson Approach Svensson (1994) proposed an extension of the Nelson-Siegel model by adding an additional hump-shaped element. It was intended that this model should be capable of producing a better fit to yield curve shapes with more than one local minimum or maximum. The model for forward rates at a given point in time (t) is:

$$f_t(\tau) = \beta_{1,t} + \beta_{2,t} e^{-\frac{\tau}{\lambda_1}} + \beta_{3,t}(\frac{\tau}{\lambda_1})e^{-\frac{\tau}{\lambda_1}} + \beta_{4,t}(\frac{\tau}{\lambda_2})e^{-\frac{\tau}{\lambda_2}},$$

where

- τ represents the term of the forward rate,
- $\beta_{1,t}$, $\beta_{2,t}$, $\beta_{3,t}$, $\beta_{4,t}$ represent parameters which determine the shape of the yield curve,
- λ is a shape parameter.

This leads to the following specification for the spot curve at time (t):

$$s_t(\tau) = \beta_{1,t} h_{1,t}(\tau) + \beta_{2,t} h_{2,t}(\tau) + \beta_{3,t} h_{3,t}(\tau) + \beta_{4,t} h_{4,t}(\tau),$$

where

$$h_{1,t}(\tau) = 1,$$

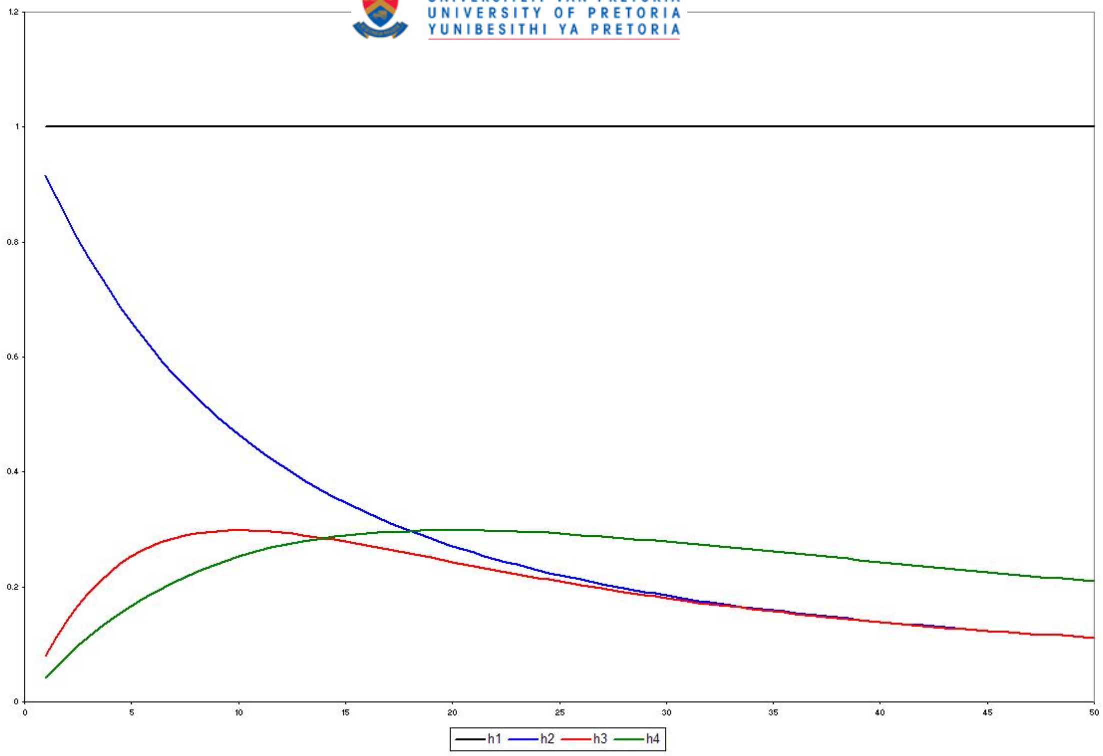


Fig. 2.4: Svensson model component functions

$$h_{2,t}(\tau) = \frac{1 - e^{-\frac{\tau}{\lambda_1}}}{\frac{\tau}{\lambda_1}},$$

$$h_{3,t}(\tau) = \frac{1 - e^{-\frac{\tau}{\lambda_1}}}{\frac{\tau}{\lambda_1}} - e^{-\frac{\tau}{\lambda_1}},$$

$$h_{4,t}(\tau) = \frac{1 - e^{-\frac{\tau}{\lambda_2}}}{\frac{\tau}{\lambda_2}} - e^{-\frac{\tau}{\lambda_2}}.$$

Therefore, the spot curve can be seen as a linear combination of four element shapes: a flat curve, a sloped curve, and two humped curves. These are depicted in Figure (2.4)

This is an improvement on the Nelson-Siegel approach as it allows a more diverse set of yield curves to be modelled. However, it can potentially introduce a large amount of multi-collinearity when fitting against actual yield curve data, especially if $\mid \lambda_1 - \lambda_2 \mid$ is relatively small.

Cairns Approach Cairns (1997) proposed an exponential type model similar to that of Nelson and Siegel. Cairns intended that this model should be able to produce a better fit for curves with multiple inflection points. The model for forward rates at a given point in time (t) is:

$$f_t(\tau) = \beta_{0,t} + \beta_{1,t}e^{-c_1\tau} + \beta_{2,t}e^{-c_2\tau} + \beta_{3,t}e^{-c_3\tau} + \beta_{4,t}e^{-c_4\tau},$$

where

- τ represents the term of the forward rate,
- $\beta_{0,t}$, $\beta_{1,t}$, $\beta_{2,t}$, $\beta_{3,t}$, $\beta_{4,t}$ represent parameters which determine the shape of the yield curve,

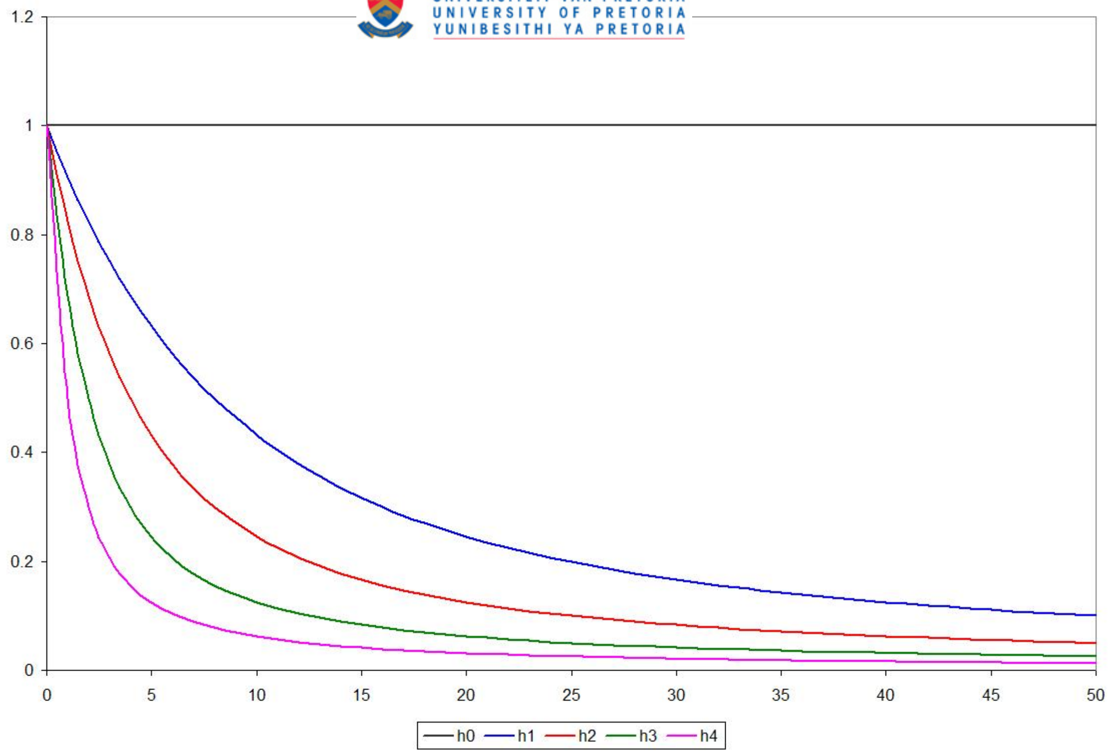


Fig. 2.5: Cairns model component functions

- λ is a shape parameter.

This leads to the following specification for the spot curve at time (t):

$$s_t(\tau) = \beta_{0,t}h_{0,t}(\tau) + \beta_{1,t}h_{1,t}(\tau) + \beta_{2,t}h_{2,t}(\tau) + \beta_{3,t}h_{3,t}(\tau) + \beta_{4,t}h_{4,t}(\tau),$$

where

$$h_{0,t}(\tau) = 1,$$

$$h_{i,t}(\tau) = \frac{1-e^{-c_i\tau}}{c_i\tau} \text{ for i = 1 to 4 .}$$

The c_i parameters govern the rate at which each h_i function reverts to zero. Therefore, the smallest c_i will be most relevant for modelling long term interest rates. To the extent that the c_i parameters are too close, this will introduce a measure of multicollinearity into the model. Cairns (1997) proposed a parameter set of (0.2, 0.4, 0.8, 1.6) which results in a set of reasonably spaced h_i functions and should reduce the risk of multicollinearity.

Figure 2.5 depicts the component functions for this model.

The Exponential-Polynomial Family The Cairns, Nelson-Siegel, and Svensson models are all part of a wider family of forward rate curves known as the Exponential-Polynomial family. Specifically:

DEFINITION. The forward curve manifold EP(K,n) is defined as the set of all curves of the form:

$$F(x) = \sum_{i=1}^{K} p_i(x)e^{-\alpha_i x}, \tag{2.39}$$

where

$x > 0$,

$\alpha_i \epsilon R \ \forall i$,

p_i is any polynomial with degree $\leqslant n_i \ \forall i$.

DEFINITION. The Exponential-Polynomial family is the family containing all sets of EP(K,n).

2.2.4 Comments on the Functional Form Approaches

Various researchers have analysed these functional form approaches, particularly regarding their ability to satisfy inter-temporal consistency. Some important contributions are summarised below:

The research of Dyvbig, Ingersoll and Ross Dyvbig, Ingersoll and Ross (1996) show that, in a liquid and arbitrage-free market without frictions, the long forward rate can never fall. They prove the following result:

Theorem 2.2.1: Let t < s, assume no arbitrage. Suppose that the long zero-coupon rate $z_L(t)$ exists at time t and that the long zero-coupon rate $z_L(s)$ exists (stochastically) at time s with probability 1. Then $z_L(t) \leqslant z_L(s, \omega)$ for a set of states ω at time s having probability 1.

Heuristically, the proof works as follows (assuming finitely many states at time s):

1. Suppose that there exists a state ω^* (with positive probability) such that $z_L(t) > z_L(s, \omega^*) = min_\omega z_L(s, \omega)$.
2. Consider the net trade of buying at t and selling at s, a bond maturing at T with face value $(1 + z_L(s, \omega^*))^{T-s}$.
3. The net cashflow generated by this transaction at time t is:
$$-\frac{(1 + z_L(s, \omega^*))^{T-s}}{(1 + z(t, T)^{T-t}}.$$
4. The net cashflow generated by this transaction at time $s > t$ is:
$$-\frac{(1 + z_L(s, \omega^*))^{T-s}}{(1 + z(t, T; \omega)^{T-s}}.$$
5. As $T \rightarrow \infty$, the cashflow at time t tends to zero because $lim_{T\rightarrow\infty} z(t, T) = z_L(t) > z_L(s, \omega^*)$.

6. However as $T \to \infty$, the cashflow at time s is contingent upon the state realised; it is 0 for states ω such that $z_L(s,\omega) > z_L(s,\omega^*)$, and 1 for states ω such that $z_L(s,\omega) = z_L(s,\omega^*)$.

7. Since there are finitely many states and ω^* has positive probability, this violates the no arbitrage condition and hence we have a contradiction.

Dyvbig, Ingersoll and Ross also prove the result for the continuous case. This result has implications for the models we have considered. Nelson-Siegel, Svensson and Cairns can all potentially allow the long forward rate to fall. This implies that, in order to ensure that the no-arbitrage condition is satisfied, it is necessary to impose sufficient restrictions on the processes governing movements in their parameters.

The research of Bjork and Christensen Suppose that we have a family of forward rate curves, e.g. the Nelson-Siegel family, denoted by χ. Suppose also that we have an interest rate model M, e.g. the Hull-White model, which represents behaviour of the financial markets. Bjork and Christensen (1999) define the concept of **consistency** as follows: the pair (M, χ) are consistent if all forward curves which may be produced by M are contained within the family χ. They identify three general problems around consistency:

1. Given an interest rate model M and a family of forward curves χ, what are necessary and sufficient conditions for consistency?

2. Take as given a specific family χ of forward curves (e.g. the Nelson-Siegel family). Does there exist any interest rate model M which is consistent with χ?

3. Take as given a specific interest rate model M (e.g. the Hull-White model). Does there exist any finitely parametrised family of forward curves χ which is consistent with M?

They fully explored the first question above by deriving necessary and sufficient conditions for consistency to exist between a Weiner-driven interest rate model and a family of forward curves. Applying the conditions to the Nelson-Siegel model they found the following important results:

1. The full Nelson-Siegel family is inconsistent with the Ho-Lee interest rate model.

2. The degenerate Nelson-Siegel family ($\lambda = 0$, $\beta_2 = 0$) is consistent with the Ho-Lee model.

3. The Hull-White model is inconsistent with the Nelson-Siegel family.

4. The Heath-Jarrow-Morton model is inconsistent with the Nelson-Siegel family.

Bjork and Christensen went further to explore the second and third questions above. Specifically, they derived some wider results for the Exponential-Polynomial family of forward rate curves under specific restrictions, but left the development of more general results as further research.

The research of Filipovic Filipovic (1999, 2000) carried the analysis further and examines the Exponential Polynomial family of models in detail. He addressed the second and third questions posed by Bjork and Christensen, specifically considering the Exponential-Polynomial family, and where interest rate models are assumed to be driven by Ito processes.

He proved the following two results:

Theorem 2.2.2: Let $K \epsilon N$, $\underline{n} \epsilon N_0^K$, $\Im_t$ is the diffusion for generic Ito process Z. If Z is consistent with the exponential-polynomial family, then the exponents are constant for $1 \leq i \leq K$.

Theorem 2.2.3: If Z is consistent with the exponential-polynomial family, then it is non-trivial only if there exists a pair of indices $1 \leq i < j \leq K$, i.e.

$$2\alpha_i = \alpha_j. \tag{2.40}$$

Therefore, where interest rates are assumed to follow an Ito process:

Theorem 2.2.2 tells us immediately that the exponential parameters in the Nelson-Siegel, Svensson, and Cairns models cannot be stochastic if intertemporal consistency is desired.

Theorem 2.2.3 tells us that there exists no non-trivial consistent process for the Nelson -Siegel model.

Theorem 2.2.3 also constrains the Svensson model exponential parameters to the following choices:

- $2\lambda_1 = \lambda_2 > 0$,
- $\lambda_1 = 2\lambda_2 > 0$.

In addition, Filipovic showed that the parameters $\beta_{1,t}$, $\beta_{3,t}$, $\beta_{4,t}$ for the Svensson family are necessarily deterministic functions of t. Hence $\beta_{2,t}$ is the only parameter with a non-trivial stochastic representation. Therefore, there exists a non-trivial diffusion process providing an arbitrage free model for the Svensson family, however the choice of parameters is extremely limited since all but one of the parameters are either constant or deterministic. This effectively means that the model, which has 6 parameters, is effectively reduced to a 1 factor model.

Theorem 2.2.3 places some broad limitations on the choice of exponential parameters in the Cairns model. However, provided that we choose these appropriately there will exist a consistent Ito process. We can also see that the parameterisation suggested by Cairns does indeed satisfy these requirements.

The research of Krippner Krippner (2005) showed that it is possible to modify the traditional Nelson-Siegel approach to obtain a model which is inter-temporally consistent. The model is defined as follows:

Assumption 1: At time t and as a function of future time t+m ($m \geq 0$), the expected path of the short rate $E_t[r(t+m)]$ under the physical measure is defined as:

$$E_t[r(t+m)] = \sum_{n=1}^{3} \lambda_n(t).g_n(\phi, m), \tag{2.41}$$

where

- $g_1(\phi, m) = 1$,
- $g_2(\phi, m) = -e^{-\phi m}$,
- $g_3(\phi, m) = -e^{-\phi m}(-2\phi m + 1)$,
- λ_n are time dependent coefficients.

Assumption 2: Instantaneous stochastic changes to the forward rate curve are as follows:

$$\sum_{n=1}^{3} \sigma_n g_n(\phi, m).dW_n(t), \tag{2.42}$$

where

- $dW_n(t)$ are Weiner increments under the physical measure.

Assumption 3: The market prices of risk (θ_n) are constants.

This results in a forward rate curve $f(t, m)$ defined as follows:

$$f(t, m) = \sigma_1\theta_1 m + \sum_{n=1}^{3} \beta_n(t) g_n(\phi, m) - \sum_{n=1}^{3} \sigma_n^2 h_n(\phi, m), \tag{2.43}$$

where

- $\beta_n(t) = \gamma_n + \lambda_n(t)$,
- $\gamma_1 = \frac{1}{\phi}(-\sigma_1\theta_2 + \sigma_2\theta_3)$,
- $\gamma_2 = \frac{1}{\phi}(-\sigma_2\theta_2 - 2\sigma_3\theta_3)$,
- $\gamma_3 = \frac{1}{\phi}\sigma_3\theta_3$,
- $h_1(\phi, m) = \frac{1}{2}m^2$,
- $h_2(\phi, m) = \frac{1}{2\phi^2}(1 - e^{-\phi m})^2$,
- $h_3(\phi, m) = \frac{1}{2\phi^2}(1 - e^{-\phi m} - 2\phi m e^{-\phi m})^2$.

2.2.5 Smith-Wilson Approach

Smith and Wilson (2000) proposed a class of models where the long forward rate is a fixed input parameter, and does not vary over time as bond prices change. The approach provides a stable method for extrapolating the yield curve, and is consistent with absence of arbitrage. An additional feature of this approach is that it is capable of exactly fitting to the initial term structure, based on a finite set of inputs.

They begin with the standard problem of yield-curve fitting, based on a set of prices for a number of bonds at a given point in time. Suppose that a bond market has I bonds of varying maturities and coupons. Denote the i^{th} bond's market value by m_i, with cash flows $c_{i,j}$ on future dates u_j. If we denote the term structure by P(τ), the price of a zero-coupon bond at time τ, then the term structure will be defined (not necessarily uniquely) as the solution of the following set of equations:

$$m_i = \sum_{j=1}^{J} c_{i,j} P(u_j).$$

However, when cash flows occur on different dates the $(c_{i,j})$ matrix will be sparse. Hence there is no guarantee that the solution of these equations will be sensible. In order ensure that a reasonable set of solutions is obtained, the following additional conditions are imposed:

- P(0) = 1,
- P(t) is a smooth function of t,
- P(t) is a positive decreasing function,
- P(t) tends exponentially to zero for large t.

It is shown separately that most stationary, arbitrage-free models of interest rates imply bond price behaviour (for large t) of the form:

$$P(t) \backsim X_0 e^{-f_\infty t} + X_1 e^{e^{-(f_\infty + \alpha)t}} + ...,$$

where

- X_0 and X_1 vary over time,
- f_∞ and α are constant over time.

Therefore, they propose a zero-coupon bond pricing function as follows:

$$P_t(\tau) = e^{-f_\infty \tau} + \sum_{i=1}^{I} \xi_{i,t} K_i(\tau),$$

where

- τ represents the term of the forward rate,

- f_∞ represents the infinite forward rate,
- I represents the number of observable bond prices used for fitting,
- ξ represents a series of time-varying parameters used to fit the actual yield curve.

The K_i's represent a set of kernel functions for each input observable bond price. This approach ensures that the pricing equations for determining the term structure are linear, which makes computation significantly simpler.

In order to ensure the asymptotic behaviour of the price function, the functional form of each kernel is chosen as follows:

$$K_i(\tau) = \sum_{j=1}^{J_i} c_{ij} W(\tau, u_j),$$

where

- c_{ij} is the cash flow in respect of bond i at time u_j,
- $W(\tau, u_j)$ is a symmetric function known as "Wilson's Function".

Wilson's function is defined as follows:

$$W(t,u) = e^{-f_\infty(t+u)}[\alpha.min\{t,u\} - e^{-\alpha.max\{t,u\}} sinh(\alpha.min\{t,u\})],$$

Figure 2.6 shows Wilson's function W(t,50), for $f_\infty = 0.05$, and $\alpha = 0.1$. Figure 2.7 shows Wilson's function W(t,u), we can see that the function is symmetric around $t = u$.

The use of Wilson's function allows the long term forward rates to converge towards the chosen infinite rate (f_∞), while the α-parameter controls the level of smoothness inherent in the extrapolation. High values of α will place greater emphasis on flatness of the forward curve beyond the longest observable term, while lower values will result in greater smoothness.

Hence, the fitted parameters ξ_i are determined by the solution to the set of equations:

$$m_i - \sum_{j=1}^{J} c_{ij} e^{-f_\infty t_j} = \sum_{s=1}^{I} \sum_{j=1}^{J_i} c_{ij} K_s(t_j) \xi_s,$$

Calculation of these various Kernel functions can be very tedious if a large number of coupon bearing bonds is used. Therefore, they propose an "approximate" kernel function which assumes coupons are paid continuously:

$$K_i^{approx}(t) = W(t, \tau_i) + g_i \int_0^{\tau_i} W(t,u) du,$$

where

- g_i represents the coupon percentage on the $i^t h$ bond,
- τ_i represents the maturity date of the $i^t h$ bond.

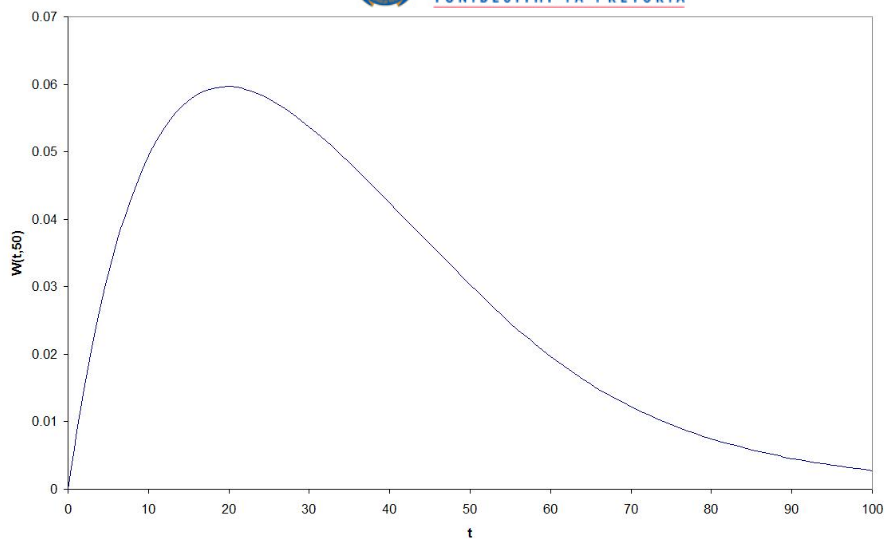


Fig. 2.6: Wilson's Function W(t,50) for $f_\infty = 0.05$*, and* $\alpha = 0.1$

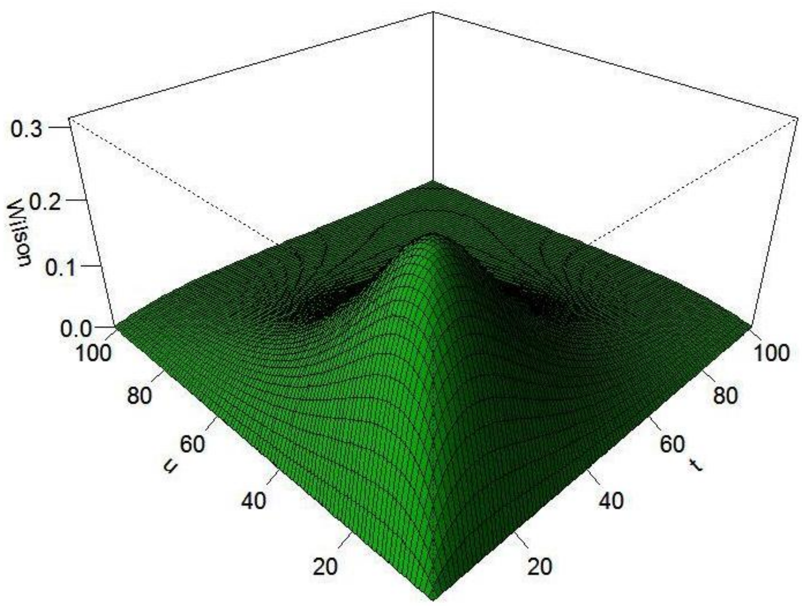


Fig. 2.7: Wilson's Function W(t,u) for $f_\infty = 0.05$*, and* $\alpha = 0.1$

Hence, using the approximate set of kernel functions, the fitted parameters ξ_i are determined by the solution to the set of equations:

$$m_i - \sum_{j=1}^{J_i} c_{ij} e^{-f_\infty t_j} = \sum_{s=1}^{I} \sum_{j=1}^{J_i} c_{ij} K_s^{approx}(t_j) \xi_s.$$

2.2.6 Modern Approaches for Managing Life Insurance / Pension Fund Risks

We have now examined various modern methods for hedging interest rate risk. We started by looking at the stochastic duration approach which attempts to quantify the sensitivity of one's interest rate exposure to the factors which drive changes in the yield curve, where yield curve behaviour is governed by a specified stochastic process. This is a significant step forward from the traditional approaches as it enables us to derive theoretical hedges for contingent interest rate risks, relative to the assumed stochastic process.

One key problem with the stochastic duration framework is that it is limited by the accuracy of the underlying stochastic interest rate process. Therefore, the nature of actual yield curve behaviour, and even the actual shape of the yield curve, may not be adequately captured by the stochastic process used. A further problem arises from the fact that the factors which drive the assumed stochastic interest rate process may not be observable (or sensible) quantities.

Our next step was to consider some of the functional form approaches to yield curve modelling. This specifically focused on the Nelson-Siegel, Svensson and Cairns approaches to modelling yield curve behaviour. Much research into the limitations of these approaches has been performed by various contributors, and we focus on some particularly interesting results from Dvybig, Ingersoll and Ross, Filipovic, and Krippner. Further, the structure of these models makes them easy to adapt for long term yield curve modelling as will be shown in the coming chapters / case studies.

We then turned our attention to a model designed by Smith and Wilson, who were specifically focused on developing a model for long term non-observable yield curve behaviour. This approach will also be used extensively in the coming chapters / case studies.

We now turn our attention to see how these approaches can be used for forecasting and hedging long term interest rate risk.

3. FORECASTING LONG TERM INTEREST RATES: YIELD CURVE EXTRAPOLATION PROCEDURES

In this chapter we investigate a number of approaches to extrapolating the yield curve beyond its maximum observable term. The selection of simple approaches considered is based on my knowledge of what has been done in practice previously, while the selection of more advanced approaches is based on the content of the last chapter.

We assume that the maximum observable and tradable term of the yield curve (M) is 30 years.

3.1 Simple Extrapolation Procedures

There are four forward rate extrapolation and four spot rate extrapolation procedures that have been included. Each of them performs a different extrapolation of the yield curve to determine the long term zero-coupon rates beyond M years. Extrapolations are performed at yearly intervals.

3.1.1 Simple Forward Rate Extrapolations

The following simple linear extrapolations of the forward curve are proposed:

Final Forward Rate Extrapolation This method assumes that the final observable forward rate prevails for each year beyond the maximum observable and tradable term, hence:

$$f_t(\tau) = f_t(M), \tau > M. \tag{3.1}$$

Linear Forward Rate Extrapolation This method assumes that the forward rates beyond M years follow a first order linear progression of the form:

$$f_t(\tau) = a + b \times \tau, \tau > M. \tag{3.2}$$

Exponential Forward Rate Extrapolation This method assumes that the forward rates beyond M years follow an exponential progression of the form:

$$f_t(\tau) = a \times e^{b \times \tau}, \tau > M. \tag{3.3}$$

Power Forward Rate Extrapolation This method assumes that the forward rates beyond M years follow a power progression of the form:

$$f_t(\tau) = a \times \tau^b, \tau > M. \tag{3.4}$$

The extrapolations will thus be performed as follows:

Final Forward Rate Extrapolation We know that, by definition:

$$P_t(\tau) = \prod_{s=1}^{\tau} (1 + f_t(s))^{-1}.$$

However, since forward rates are only observable up to M years, for $\tau > M$:

$$P_t(\tau) = P_M \times \prod_{s=M+1}^{\tau} (1 + f_t(s))^{-1},$$

so

$$P_t(\tau) = P_t(M) \times \prod_{s=M+1}^{\tau} (1 + f_t(M))^{-1}. \tag{3.5}$$

Linear Forward Rate Extrapolation Similarly to above, for $\tau > M$:

$$P_t(\tau) = P_t(M) \times \prod_{s=M+1}^{\tau} (1 + a + b \times s)^{-1}. \tag{3.6}$$

Exponential Forward Rate Extrapolation Similarly to above, for $\tau > M$:

$$P_t(\tau) = P_t(M) \times \prod_{s=M+1}^{\tau} (1 + a \times e^{b \times s})^{-1}. \tag{3.7}$$

Power Forward Rate Extrapolation Similarly to above, for $\tau > M$:

$$P_t(\tau) = P_t(M) \times \prod_{s=M+1}^{\tau} (1 + a \times s^b)^{-1}. \tag{3.8}$$

3.1.2 Simple Spot Rate Extrapolations

The following simple linear extrapolations of the spot curve are proposed:

Final Spot Rate Extrapolation This method assumes that the final observable forward rate prevails for each year beyond the maximum observable and tradable term, hence:

$$s_t(\tau) = s_t(M), \tau > M. \tag{3.9}$$

Linear Spot Rate Extrapolation This method assumes that the forward rates beyond M years follow a first order linear progression of the form:

$$s_t(\tau) = a + b \times \tau, \tau > M. \tag{3.10}$$

Exponential Spot Rate Extrapolation This method assumes that the forward rates beyond M years follow an exponential progression of the form:

$$s_t(\tau) = a \times e^{b \times \tau}, \tau > M. \tag{3.11}$$

Power Spot Rate Extrapolation This method assumes that the forward rates beyond M years follow a power progression of the form:

$$s_t(\tau) = a \times \tau^b, \tau > M. \tag{3.12}$$

The extrapolations will thus be performed as follows:

Final Spot Rate Extrapolation For $\tau > M$ we have:

$$P_t(\tau) = (1 + s_t(M))^{-\tau}. \tag{3.13}$$

Other Spot Rate Extrapolations For $\tau > M$ we have:

$$P_t(\tau) = (1 + s_t(M; a, b))^{-\tau}. \tag{3.14}$$

Figure 3.1 illustrates how the spot rate extrapolations are performed on a spot curve beyond 30 years.

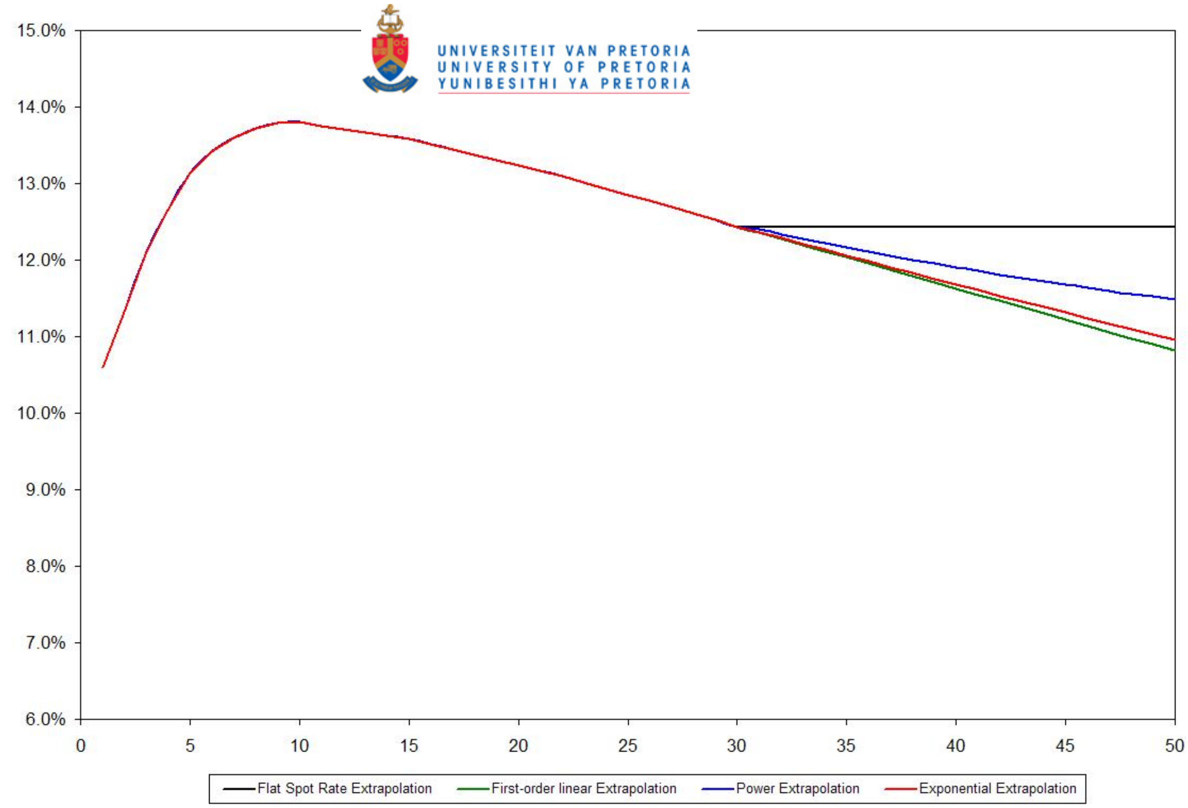


Fig. 3.1: Illustration of simple spot rate extrapolations

3.2 Advanced Extrapolation Procedures

The following, more advanced, extrapolation procedures will be used to extrapolate the yield curve beyond the maximum observable term (M):

3.2.1 Nelson-Siegel Approach

As described in the previous chapter, for $\tau > M$ the spot curve at time (t) is given by:

$$s_t(\tau) = \beta_{1,t}h_{1,t}(\tau) + \beta_{2,t}h_{2,t}(\tau) + \beta_{3,t}h_{3,t}(\tau), \tag{3.15}$$

where

$$h_{1,t}(\tau) = 1,$$

$$h_{2,t}(\tau) = \frac{1 - e^{-\frac{\tau}{\lambda}}}{\frac{\tau}{\lambda}},$$

$$h_{3,t}(\tau) = \frac{1 - e^{-\frac{\tau}{\lambda}}}{\frac{\tau}{\lambda}} - e^{-\frac{\tau}{\lambda}}.$$

3.2.2 Svensson Approach

As described in the previous chapter, for $\tau > M$ the spot curve at time (t) is given by:

$$s_t(\tau) = \beta_{1,t}h_{1,t}(\tau) + \beta_{2,t}h_{2,t}(\tau) + \beta_{3,t}h_{3,t}(\tau) + \beta_{4,t}h_{4,t}(\tau), \tag{3.16}$$

where

$$h_{1,t}(\tau) = 1,$$

$$h_{2,t}(\tau) = \frac{1 - e^{-\frac{\tau}{\lambda_1}}}{\frac{\tau}{\lambda_1}},$$

$$h_{3,t}(\tau) = \frac{1 - e^{-\frac{\tau}{\lambda_1}}}{\frac{\tau}{\lambda_1}} - e^{-\frac{\tau}{\lambda_1}},$$

$$h_{4,t}(\tau) = \frac{1 - e^{-\frac{\tau}{\lambda_2}}}{\frac{\tau}{\lambda_2}} - e^{-\frac{\tau}{\lambda_2}}.$$

3.2.3 Cairns Approach

As described in the previous chapter, for $\tau > M$ the spot curve at time (t) is given by:

$$s_t(\tau) = \beta_{0,t}h_{0,t}(\tau) + \beta_{1,t}h_{1,t}(\tau) + \beta_{2,t}h_{2,t}(\tau) + \beta_{3,t}h_{3,t}(\tau) + \beta_{4,t}h_{4,t}(\tau), \tag{3.17}$$

where

$$h_{0,t}(\tau) = 1,$$

$$h_{i,t}(\tau) = \frac{1-e^{-c_i\tau}}{c_i\tau} \text{ for i = 1 to 4.}$$

3.2.4 Smith-Wilson Approach

As described in the previous chapter, for term $\tau > M$ the price of a zero coupon bond at time (t) is given by:

$$P_t(\tau) = e^{-f_\infty \tau} + \sum_{i=1}^{I} \xi_{i,t} K_i(\tau), \tag{3.18}$$

where

$$K_i(\tau) = \sum_{j=1}^{J_i} c_{ij} W(\tau, u_j), \tag{3.19}$$

$$W(\tau, u) = e^{-f_\infty(\tau+u)}[\alpha.min\{\tau, u\} - e^{-\alpha.max\{\tau,u\}} sinh(\alpha.min\{\tau, u\})]. \tag{3.20}$$

Remember that $c_{i,j}$ in the above equations represents the j^{th} cash flow on the i^{th} bond used to calibrate the price function, while u_j represents the term of the respective cash flow. Therefore, if we assume a finite set of observable zero-coupon bond prices are used for calibration then equation (3.19) reduces as follows:

$$K_i(\tau) = W(\tau, u_i). \tag{3.21}$$

We will make use of this simplification in the work that follows as we will be assuming an observable spot curve at yearly intervals.

Notice that we can therefore re-write equation (3.18) as follows:

$$P_t(\tau) = e^{-f_\infty \tau} + \underline{\xi_t' W(\tau)}, \tag{3.22}$$

where

$$\underline{\xi}_t(n \times 1) = \begin{matrix} \xi_{1,t} \\ \xi_{2,t} \\ \dots \\ \xi_{n,t} \end{matrix}, \underline{W(\tau)}(n \times 1) = \begin{matrix} W(\tau, u_1) \\ W(\tau, u_2) \\ \dots \\ W(\tau, u_n) \end{matrix}$$

3.2.5 Bayes and Credibility Theory vs Smith-Wilson Approach

If we assume that there is only one observable and tradable zero-coupon bond (with term κ) in the market, then the $P_t(\tau)$ function reduces to:

$$P_t(\tau) = e^{-f_\infty \tau} + \frac{P_t(\kappa) - e^{-f_\infty \kappa}}{W(\kappa, \kappa)} W(\tau, \kappa). \tag{3.23}$$

This can be rewritten as:

$$P_t(\tau) = e^{-f_\infty \tau}(1 - \gamma(\tau, \kappa)) + (1 + X)e^{-f_\infty \tau}\gamma(\tau, \kappa), \tag{3.24}$$

where

$$\gamma(\tau, \kappa) = \frac{W(\tau, \kappa)}{W(\kappa, \kappa)e^{-f_\infty(\tau - \kappa)}} = \frac{\alpha.min(\tau, \kappa) - e^{-\alpha.max(\tau,\kappa)} sinh(\alpha.min(\tau, \kappa))}{\alpha.\kappa - e^{-\alpha.\kappa} sinh(\alpha.\kappa)}, \tag{3.25}$$

$$X = \frac{P_t(\kappa) - e^{-f_\infty \kappa}}{e^{-f_\infty \kappa}}. \tag{3.26}$$

X represents the percentage difference between $P_t(\kappa)$ and $e^{-f_\infty \kappa}$. For the case where we project long term interest rates beyond the term of the observable bond, i.e. $\tau > \kappa$, gamma reduces to:

$$\gamma(\tau, \kappa) = \frac{\alpha.\kappa - e^{-\alpha.\tau} sinh(\alpha.\kappa)}{\alpha.\kappa - e^{-\alpha.\kappa} sinh(\alpha.\kappa)}. \tag{3.27}$$

This result resembles the statistical problem of Credibility (described by Norberg (2004)) where a quantity is estimated based on a weighted average of two quantities. Further, the above approach is closely related to the Bayes framework. If we view $e^{-f_\infty \tau}$ as the prior estimate of the price of a ZCB of term τ, then formula (3.24) can be seen as a posterior estimate for $P_t(\tau)$, conditional upon the value of $P_t(\kappa)$.

We can therefore infer certain implicit characteristics regarding the behaviour of interest rates in the Smith-Wilson approach. Taking the appropriate derivatives, we see that the forward rates in the above process are given by:

$$f_t(\tau) = -\frac{dln(P_t(\tau))}{d\tau} = \frac{f_\infty e^{-f_\infty \tau} - \xi \frac{dW(\tau,\kappa)}{d\tau}}{P_t(\tau)}.$$

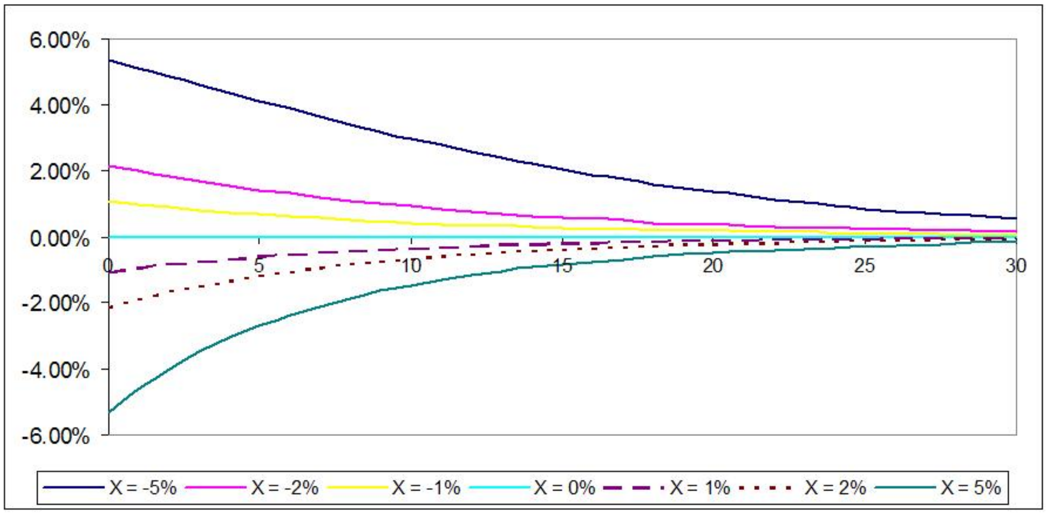


Fig. 3.2: Illustration of $f_t(\tau) - f_\infty$ for varying levels of X, $\kappa = 1$

Now for $\tau > \kappa$:

$$\frac{dW(\tau,\kappa)}{d\tau} = -f_\infty W(\tau,\kappa) + \alpha.e^{-f_\infty(\tau+\kappa)-\alpha\tau} sinh(\alpha\kappa).$$

Therefore

$$f_t(\tau) = f_\infty - \xi \frac{\alpha e^{-f_\infty(\tau+\kappa)-\alpha\tau} sinh(\alpha\kappa)}{e^{-f_\infty \tau} + \xi W(\tau,\kappa)}. \quad (3.28)$$

This expression gives us a very clear indication of the behaviour of forward rates around the long term mean of f_∞. Figure 3.2 shows the behaviour of the second component in this expression for various levels of X, where $\kappa = 1$.

3.3 Estimating Extrapolation Parameters: Simple Extrapolations

Estimation of the extrapolation parameters in the Simple Extrapolation Approaches (a *and* b) will be based on least squares estimates which are fitted to the observable interest rates. Hence, where necessary, a and b will be estimated from forward rates *f(M-q)* to *f(M)*. We will use q = 10 for estimating parameters a and b, as this is the range that seems to be most commonly used by practitioners for extrapolating the yield curve. We will illustrate the estimation process assuming q = 10 and M = 30.

Note that we will denote the input vectors as follows:

$$\underline{T}(10\times 1) = \begin{matrix} 21 \\ 22 \\ ... \\ 30 \end{matrix}, \underline{T^*}(10\times 1) = \begin{matrix} ln(21) \\ ln(22) \\ ... \\ ln(30) \end{matrix}, \underline{f}_t(10\times 1) = \begin{matrix} f_t(21) \\ f_t(22) \\ ... \\ f_t(30) \end{matrix}, \underline{f_t^*}(10\times 1) = \begin{matrix} ln(f_t(21)) \\ ln(f_t(22)) \\ ... \\ ln(f_t(30)) \end{matrix}$$

$$\underline{s}_t(10\times 1) = \begin{matrix} s_t(21) \\ s_t(22) \\ ... \\ s_t(30) \end{matrix}, \underline{s_t^*}(10\times 1) = \begin{matrix} ln(s_t(21)) \\ ln(s_t(22)) \\ ... \\ ln(s_t(30)) \end{matrix}$$

Result 3.3.1: Simple Forward Rate Extrapolations

Final Forward Rate Extrapolation No complex extrapolation is necessary as it is based solely on the M^{th} yearly forward rate - which we have assumed is directly observable.

Linear Forward Rate Extrapolation The parameters a and b are estimated as follows:

$$\hat{b} = \frac{\underline{T}'\underline{f}_t - \frac{\underline{T}'\underline{1}\times\underline{f}_t'\underline{1}}{n}}{\underline{T}'\underline{T} - \frac{(\underline{T}'\underline{1})^2}{n}}, \hat{a} = \frac{\underline{f}_t'\underline{1} - \hat{b}\times\underline{T}'\underline{1}}{n}. \tag{3.29}$$

Exponential Forward Rate Extrapolation The parameters a and b are estimated as follows:

$$\hat{b} = \frac{\underline{T^*}'\underline{f_t^*} - \frac{\underline{T^*}'\underline{1}\times\underline{f_t^*}'\underline{1}}{n}}{\underline{T^*}'\underline{T} - \frac{(\underline{T^*}'\underline{1})^2}{n}}, \hat{a} = \frac{\underline{f_t^*}'\underline{1} - \hat{b}\times\underline{T^*}'\underline{1}}{n}. \tag{3.30}$$

Power Forward Rate Extrapolation The parameters a and b are estimated as follows:

$$\hat{b} = \frac{\underline{T}'\underline{f_t^*} - \frac{\underline{T}'\underline{1}\times\underline{f_t^*}'\underline{1}}{n}}{\underline{T}'\underline{T} - \frac{(\underline{T}'\underline{1})^2}{n}}, \hat{a} = \frac{\underline{f_t^*}'\underline{1} - \hat{b}\times\underline{T}'\underline{1}}{n}. \tag{3.31}$$

Result 3.3.2: Simple Spot Rate Extrapolations

Final Spot Rate Extrapolation No complex extrapolation is necessary as it is based solely on the M-year spot rate - which we have assumed is directly observable.

Linear Spot Rate Extrapolation The parameters a and b are estimated as follows:

$$\hat{b} = \frac{\underline{T}'\underline{s}_t - \frac{\underline{T}'\underline{1} \times \underline{s}_t'\underline{1}}{n}}{\underline{T}'\underline{T} - \frac{(\underline{T}'\underline{1})^2}{n}}, \hat{a} = \frac{\underline{s}_t'\underline{1} - \hat{b} \times \underline{T}'\underline{1}}{n}. \tag{3.32}$$

Exponential Spot Rate Extrapolation The parameters a and b are estimated as follows:

$$\hat{b} = \frac{\underline{T}^{*\prime}\underline{s}_t^* - \frac{\underline{T}^{*\prime}\underline{1} \times \underline{s}_t^{*\prime}\underline{1}}{n}}{\underline{T}^{*\prime}\underline{T} - \frac{(\underline{T}^{*\prime}\underline{1})^2}{n}}, \hat{a} = \frac{\underline{s}_t^{*\prime}\underline{1} - \hat{b} \times \underline{T}^{*\prime}\underline{1}}{n}. \tag{3.33}$$

Power Spot Rate Extrapolation The parameters a and b are estimated as follows:

$$\hat{b} = \frac{\underline{T}'\underline{s}_t^* - \frac{\underline{T}'\underline{1} \times \underline{s}_t^{*\prime}\underline{1}}{n}}{\underline{T}'\underline{T} - \frac{(\underline{T}'\underline{1})^2}{n}}, \hat{a} = \frac{\underline{s}_t^{*\prime}\underline{1} - \hat{b} \times \underline{T}'\underline{1}}{n}. \tag{3.34}$$

3.4 Estimating Extrapolation Parameters: Advanced Extrapolations

Suppose that we have an observable set of zero coupon bond prices at time t, for terms of $\{T_1, T_2, ..., T_M\}$ with associated spot rates and prices given by $\{s_t(T_1), s_t(T_2), ..., s_t(T_M)\}$ and $\{P_t(T_1), P_t(T_2), ..., P_t(T_M)\}$. We will use the following notation:

$$\underline{T}(M \times 1) = \begin{matrix} T_1 \\ T_2 \\ ... \\ T_M \end{matrix}, \underline{s_t}(M \times 1) = \begin{matrix} s_t(T_1) \\ s_t(T_2) \\ ... \\ s_t(T_M) \end{matrix}, \underline{P_t}(M \times 1) = \begin{matrix} P_t(T_1) \\ P_t(T_2) \\ ... \\ P_t(T_M) \end{matrix}$$

Result 3.4.1: Nelson-Siegel Approach

Under the Nelson-Siegel approach we aim to select $\underline{\hat{\beta}_t}$ such that we minimize $\sum_{i=1}^{M} (s_t(T_i) - \hat{s}_t(T_i))^2$ where $\hat{s}_t(T_i)$ is defined by (3.15), with parameters $\underline{\hat{\beta}_t}$. We can express this as follows:

Estimation Objective Minimize $(\underline{s}_t - H_N\underline{\hat{\beta}_t})'(\underline{s}_t - H_N\underline{\hat{\beta}_t})$

where

$$H_N(M \times 3) = \begin{matrix} h_{1,t}(T_1) & h_{2,t}(T_1) & h_{3,t}(T_1) \\ h_{1,t}(T_2) & h_{2,t}(T_2) & h_{3,t}(T_2) \\ ... & & \\ h_{1,t}(T_M) & h_{2,t}(T_M) & h_{3,t}(T_M) \end{matrix},$$

and the $h_{i,t}$ functions are as defined in (3.15) above.

Estimation Solution Taking partial derivatives yields the following solution:

$$\hat{\beta}_t = (H_N' H_N)^{-1}(H_N' \underline{s}_t). \tag{3.35}$$

Result 3.4.2: Svensson Approach

Under the Svensson approach we aim to select $\underline{\hat{\beta}_t}$ such that we minimize $\sum_{i=1}^{M} (s_t(T_i) - \hat{s}_t(T_i))^2$ where $\hat{s}_t(T_i)$ is defined by (3.16), with parameters $\underline{\hat{\beta}_t}$. We can express this as follows:

Estimation Objective Minimize $(\underline{s}_t - H_S\underline{\hat{\beta}_t})'(\underline{s}_t - H_S\underline{\hat{\beta}_t})$

where

$$H_S(M \times 4) = \begin{matrix} h_{1,t}(T_1) & h_{2,t}(T_1) & h_{3,t}(T_1) & h_{4,t}(T_1) \\ h_{1,t}(T_2) & h_{2,t}(T_2) & h_{3,t}(T_2) & h_{4,t}(T_2) \\ ... & & & \\ h_{1,t}(T_M) & h_{2,t}(T_M) & h_{3,t}(T_M) & h_{4,t}(T_M) \end{matrix},$$

and the $h_{i,t}$ functions are as defined in (3.16) above.

Estimation Solution Taking partial derivatives yields the following solution:

$$\underline{\hat{\beta}_t} = (H_S' H_S)^{-1}(H_S' \underline{s}_t) \tag{3.36}$$

Result 3.4.3: Cairns Approach

Under the Cairns approach we aim to select $\underline{\hat{\beta}_t}$ such that we minimize $\sum_{i=1}^{M}(s_t(T_i) - \hat{s}_t(T_i))^2$ where $\hat{s}_t(T_i)$ is defined by (3.17), with parameters $\underline{\hat{\beta}_t}$. We can express this as follows:

Estimation Objective Minimize $(\underline{s}_t - H_C\underline{\hat{\beta}_t})'(\underline{s}_t - H_C\underline{\hat{\beta}_t})$

where

$$H_C(M \times 5) = \begin{matrix} h_{0,t}(T_1) & h_{1,t}(T_1) & h_{2,t}(T_1) & h_{3,t}(T_1) & h_{4,t}(T_1) \\ h_{0,t}(T_2) & h_{1,t}(T_2) & h_{2,t}(T_2) & h_{3,t}(T_2) & h_{4,t}(T_2) \\ \ldots & & & & \\ h_{0,t}(T_M) & h_{1,t}(T_M) & h_{2,t}(T_M) & h_{3,t}(T_M) & h_{4,t}(T_M) \end{matrix},$$

and the $h_{i,t}$ functions are as defined in (3.17) above.

Estimation Solution Taking partial derivatives yields the following solution:

$$\underline{\hat{\beta}_t} = (H_C' H_C)^{-1}(H_C' \underline{s}_t) \tag{3.37}$$

Result 3.4.4: Smith-Wilson Approach

Estimation Objective Under the Smith-Wilson approach we aim to select $\underline{\hat{\xi}_t}$ such that $\hat{P}_t(T_i) = P_t(T_i)$ for $i = 1, 2 \ldots M$, where $\hat{P}_t(T_i)$ is defined similarly to (3.22) above.

Estimation Solution Assuming values for f_∞ and α (such that W below is invertible); we have M equations and M unknowns giving us the following result:

$$\underline{\hat{\xi}_t} = W^{-1}[\underline{P_t} - \underline{P_t^*}] \tag{3.38}$$

where

$$\underline{P_t^*}(M \times 1) = \begin{matrix} e^{-f_\infty T_1} \\ e^{-f_\infty T_2} \\ \ldots \\ e^{-f_\infty T_M} \end{matrix}, W(M \times M) = \begin{matrix} W(T_1, T_1) & W(T_1, T_2) & \ldots & W(T_1, T_M) \\ W(T_2, T_1) & W(T_2, T_2) & \ldots & W(T_2, T_M) \\ \ldots & & & \\ W(T_M, T_1) & W(T_M, T_2) & \ldots & W(T_M, T_M) \end{matrix}.$$

4. HEDGING LONG TERM INTEREST RATES: SOME GENERAL RESULTS

In the previous two chapters we identified a number of approaches which can be used to forecast long term interest rates beyond the longest observable term of the yield curve. In this chapter we will use our knowledge of the various forecasting approaches to derive information that will be necessary to hedge long term interest rates. At this stage we will only focus on the simple case where we hedge a long term zero coupon bond. The first section derives hedging information with respect to the relevant extrapolation parameters, while the second derives information with respect to the observable forward rates. Relevant proofs are provided in the Appendix A.

4.1 Deriving the Greeks: Simple Extrapolation Parameters

We will start by calculating the relevant partial derivatives of the price of the zero coupon bond, with respect to the extrapolation parameters, for each extrapolation procedure.

However, in order to hedge a long term zero coupon bond, it is not sufficient to only have the partial derivative of the price w.r.t. the extrapolation parameters. We also need to have:

a. The partial derivatives of the extrapolation parameters w.r.t to each of the prevailing forward / spot rates.

b. The partial derivatives of the price w.r.t to the prevailing yield curve, or w.r.t. each of the prevailing forward / spot rates.

In this section we will derive these for each extrapolation procedure; but first, we will derive a result that is common to all extrapolation procedures.

Result 4.0.0 For any zero-coupon bond, we can write:

$$\frac{\partial P_t(\tau)}{\partial f_t(k)} = \begin{array}{ll} -P_t(\tau) \times (1 + f_t(k))^{-1}, & for\ k < \tau \\ 0 & , otherwise \end{array} \tag{4.1}$$

4.1.1 Final Forward Rate Extrapolation

Result 4.1.1A For $\tau > M$, we have:

$$P_t(\tau) = \frac{P_t(M)}{(1+\beta)^{x-M}},$$

where

$$\beta = f_t(M),$$

so

$$\frac{\partial P_t(\tau)}{\partial \beta} = \frac{-P_t(M) \times (\tau - M)}{(1+\beta)^{(\tau - M + 1)}}. \tag{4.2}$$

Result 4.1.1A gives us the partial relation between the price of the bond and the extrapolation parameter. Result 4.0.0 gives us the partial relation between the price of the bond and the observable forward rates. However, we still need to derive the relation between the extrapolation parameter and the observable forward rates:

Result 4.1.1B Since $\beta = f_M$

$$\frac{d\beta}{df_t(k)} = \begin{matrix} 1, & k = M \\ 0, otherwise \end{matrix} \tag{4.3}$$

Bringing together the Results 4.1.1A-B, we get:

Result 4.1.1C

$$dP_t(\tau) = \sum_{s=1}^{M} \frac{\partial P_t(\tau)}{\partial f_t(s)} df_t(s) + \frac{\partial P_t(\tau)}{\partial \beta} \times \frac{d\beta}{df_t(M)} df_t(M). \tag{4.4}$$

4.1.2 Linear Forward Rate Extrapolation

Result 4.1.2A Where the long term non-observable forward rates follow a linear progression of the form:

$$f_t(s) = a + b \times s;$$

Then we can write:

$$\frac{\partial P_t(\tau)}{\partial a} = -P_t(\tau) \times \sum_{i=1}^{\tau-M} \frac{P_t(M+i)}{P_t(M+i-1)}. \tag{4.5}$$

Result 4.1.2B Where the long term non-observable forward rates follow a linear progression of the form:

$$f_t(s) = a + b \times s;$$

Then we can write:

$$\frac{\partial P_t(\tau)}{\partial b} = -P_t(\tau) \times \sum_{i=1}^{\tau-M} \frac{P_t(M+i)}{P_t(M+i-1)} \times (M+i). \tag{4.6}$$

Results 4.1.2A and 4.1.2B give us the partial relation between the price of the bond and the extrapolation parameters. Result 4.0.0 gives us the partial relation between the price of the bond and the observable forward rates. Once again, we still need to derive the relation between the extrapolation parameters and the observable forward rates:

Result 4.1.2C Following Result 3.3.1; for $M - q + 1 <= k <= M$

$$\frac{db}{df_t(k)} = \frac{T_k - \frac{\sum_{s=M-q+1}^{M} T_s}{q}}{\sum_{s=M-q+1}^{M} T_s^2 - \frac{(\sum_{s=M-q+1}^{M} T_s)^2}{q}}. \tag{4.7}$$

Result 4.1.2D Following Result 3.3.1; for $M - q + 1 <= k <= M$

$$\frac{da}{df_t(k)} = \frac{(1 - \frac{db}{df_t(k)} \times \sum_{s=M-q+1}^{M} T_s)}{q}. \tag{4.8}$$

The proofs of these results have been excluded as they follow from the respective definitions.

Bringing together the Results 4.1.2A-D, we get:

Result 4.1.2E

$$dP_t(\tau) = \sum_{s=1}^{M} \frac{\partial P_t(\tau)}{\partial f_t(s)} df_t(s) + \frac{\partial P_t(\tau)}{\partial a} \times \sum_{s=M-q+1}^{M} \frac{da}{df_t(s)} df_t(s) + \frac{\partial P_t(\tau)}{\partial b} \times \sum_{s=M-q+1}^{M} \frac{db}{df_t(s)} df_t(s). \tag{4.9}$$

The result follows from the fact that:

$$P_t(\tau) = f^1(f_t(1), f_t(2), f_t(3), ..., f_t(M), a, b),$$

$$a = f^2(f_t(M-q), f_t(M-q+1), ..., f_t(M)),$$
$$b = f^3(f_t(M-q), f_t(M-q+1), ..., f_t(M)).$$

4.1.3 Exponential Forward Rate Extrapolation

Result 4.1.3A Where the long term non-observable forward rates follow an exponential progression of the form:

$$f_t(s) = a \times b^s;$$

Then we can write:

$$\frac{\partial P_t(\tau)}{\partial a} = -P_t(\tau) \times \sum_{i=1}^{\tau-M} \frac{b^{M+i}}{1 + a \times b^{M+i}}. \tag{4.10}$$

Result 4.1.3B Where the long term non-observable forward rates follow an exponential progression of the form:

$$f_t(s) = a \times b^s;$$

Then we can write:

$$\frac{\partial P_t(\tau)}{\partial b} = -P_t(\tau) \times \sum_{i=1}^{\tau-M} \frac{b^{M+i-1}}{1 + a \times b^{M+i}} \times (M+i). \tag{4.11}$$

Results 4.1.3A and 4.1.3B give us the partial relation between the price of the bond and the extrapolation parameters. Result 4.0.0 gives us the partial relation between the price of the bond and the observable forward rates. Once again, we still need to derive the relation between the extrapolation parameters and the observable forward rates:

Result 4.1.3C Following Result 3.3.1; for $M - q + 1 <= k <= M$

$$\frac{db}{df_t(k)} = \frac{1}{f_t(k)} \times \frac{T_k - \frac{\sum_{s=M-q+1}^{M} T_s}{q}}{\sum_{s=M-q+1}^{M} T_s^2 - \frac{(\sum_{s=M-q+1}^{M} T_s)^2}{q}}. \tag{4.12}$$

Result 4.1.3D Following Result 3.3.1; for $M-q+1 <= k <= M$

$$\frac{da}{df_t(k)} = a \times \frac{(\frac{1}{f_t(k)} - \frac{db}{df_t(k)} \times \sum_{s=M-q+1}^{M} T_s)}{q}. \tag{4.13}$$

The proofs of these results have been excluded as they follow from the respective definitions. Bringing together the Results 4.1.3A-D, we get:

Result 4.1.3E

$$dP_t(\tau) = \sum_{s=1}^{M} \frac{\partial P_t(\tau}{\partial f_t(s)} df_t(s) + \frac{\partial P_t(\tau)}{\partial a} \times \sum_{s=M-q+1}^{M} \frac{da}{df_t(s)} df_t(s) + \frac{\partial P_t(\tau)}{\partial b} \times \sum_{s=M-q+1}^{M} \frac{db}{df_t(s)} df_t(s). \tag{4.14}$$

The result follows from the fact that:

$$P_t(\tau) = f^1(f_t(1), f_t(2), f_t(3), ..., f_t(M), a, b),$$

$$a = f^2(f_t(M-q), f_t(M-q+1), ..., f_t(M)),$$
$$b = f^3(f_t(M-q), f_t(M-q+1), ..., f_t(M)).$$

4.1.4 Power Forward Rate Extrapolation

Result 4.1.4A Where the long term non-observable forward rates follow a power progression of the form:

$$f_t(s) = a \times s^b;$$

Then we can write:

$$\frac{\partial P_t(\tau)}{\partial a} = -P_t(\tau) \times \sum_{i=1}^{\tau-M} \frac{(M+i)^b}{1 + a \times (M+i)^b}. \tag{4.15}$$

Result 4.1.4B Where the long term non-observable forward rates follow a power progression of the form:

$$f_t(s) = a \times s^b;$$

Then we can write:

$$\frac{\partial P_t(\tau)}{\partial b} = -P_t(\tau) \times \sum_{i=1}^{\tau-M} \frac{a \times (M+i)^b \times ln(M+i)}{1 + a \times (M+i)^b}. \tag{4.16}$$

Results 4.1.4A and 4.1.4B give us the partial relation between the price of the bond and the extrapolation parameters. Result 4.0.0 gives us the partial relation between the price of the bond and the observable forward rates. Once again, we still need to derive the relation between the extrapolation parameters and the observable forward rates:

Result 4.1.4C Following Result 3.3.1; for $M-q+1 <= k <= M$

$$\frac{db}{df_t(k)} = \frac{1}{f_t(k)} \times \frac{ln(T_k) - \frac{\sum_{s=M-q+1}^{M} ln(T_s)}{q}}{\sum_{s=M-q+1}^{M} ln(T_s)^2 - \frac{(\sum_{s=M-q+1}^{M} ln(T_s))^2}{q}}. \tag{4.17}$$

Result 4.1.4D Following Result 3.3.1; for $M-q+1 <= k <= M$

$$\frac{da}{df_t(k)} = a \times \frac{(\frac{1}{f_t(k)} - \frac{db}{df_t(k)} \times \sum_{s=M-q+1}^{M} ln(T_s))}{q}. \tag{4.18}$$

The proofs of these results have been excluded as they follow from the respective definitions. Bringing together the Results 4.1.4A-D, we get:

Result 4.1.4E

$$dP_t(\tau) = \sum_{s=1}^{M} \frac{\partial P_t(\tau)}{\partial f_t(s)} df_t(s) + \frac{\partial P_t(\tau)}{\partial a} \times \sum_{s=M-q+1}^{M} \frac{da}{df_t(s)} df_t(s) + \frac{\partial P_t(\tau)}{\partial b} \times \sum_{s=M-q+1}^{M} \frac{db}{df_t(s)} df_t(s). \tag{4.19}$$

The result follows from the fact that:

$$P_t(\tau) = f^1(f_t(1), f_t(2), f_t(3), ..., f_t(M), a, b),$$

$$a = f^2(f_t(M-q), f_t(M-q+1), ..., f_t(M)),$$
$$b = f^3(f_t(M-q), f_t(M-q+1), ..., f_t(M)).$$

4.1.5 Flat Spot Rate Extrapolation

Result 4.1.5A Where the long term non-observable spot rates follow a progression of the form:

$$s_t(\tau) = s_t(M), \tau > M;$$

Then we can write:

$$\frac{\partial P_t(\tau)}{\partial s_t(M)} = -P_t(\tau) \times (1 + s_t(M))^{-1}. \tag{4.20}$$

Further, because $P_t(\tau) = f(s_t(M))$, we can write:

Result 4.1.5B Following Result 3.3.2; for $\tau > M$, we have:

$$dP_t(\tau) = \frac{\partial P_t(\tau)}{\partial s_t(M)} ds_t(M). \tag{4.21}$$

4.1.6 Linear Spot Rate Extrapolation

Result 4.1.6A Where the long term non-observable spot rates follow a linear progression of the form:

$$s_t(\tau) = a + b\tau, \tau > M;$$

Then we can write:

$$\frac{\partial P_t(\tau)}{\partial a} = -P_t(\tau) \times (1 + s_t(\tau))^{-1}. \tag{4.22}$$

Result 4.1.6B Where the long term non-observable spot rates follow a linear progression of the form:

$$s_t(\tau) = a + b\tau, \tau > M;$$

Then we can write:

$$\frac{\partial P_t(\tau)}{\partial b} = -\tau P_t(\tau) \times (1 + s_t(\tau))^{-1}. \tag{4.23}$$

We now need to derive the partial derivatives of the extrapolation parameters to the observable spot rates:

Result 4.1.6C Following Result 3.3.2; for $M - q + 1 <= k <= M$

$$\frac{db}{ds_t(k)} = \frac{T_k - \frac{\sum_{s=M-q+1}^{M} T_s}{q}}{\sum_{s=M-q+1}^{M} T_s^2 - \frac{(\sum_{s=M-q+1}^{M} T_s)^2}{q}}. \tag{4.24}$$

Result 4.1.6D Following Result 3.3.2; for $M - q + 1 <= k <= M$

$$\frac{da}{ds_t(k)} = \frac{(1 - \frac{db}{ds_t(k)} \times \sum_{s=M-q+1}^{M} T_s)}{q}. \tag{4.25}$$

The proofs of these results have been excluded as they follow from the respective definitions.

Bringing together the Results 4.1.6A-D, we get:

Result 4.1.6E

$$dP_t(\tau) = \frac{\partial P_t(\tau)}{\partial a} \times \sum_{s=M-q+1}^{M} \frac{da}{ds_t(s)} ds_t(s) + \frac{\partial P_t(\tau)}{\partial b} \times \sum_{s=M-q+1}^{M} \frac{db}{ds_t(s)} ds_t(s). \tag{4.26}$$

The result follows from the fact that:

$$P_t(\tau) = f^1(a, b),$$
$$a = f^2(s_t(M - q + 1), ..., s_t(M)),$$
$$b = f^3(s_t(M - q + 1), ..., s_t(M)).$$

4.1.7 Exponential Spot Rate Extrapolation

Result 4.1.7A Where the long term non-observable spot rates follow an exponential progression of the form:

$$s_t(\tau) = a.e^{b\times\tau}, \tau > M;$$

Then we can write:

$$\frac{\partial P_t(\tau)}{\partial a} = -P_t(\tau) \times (1 + s_t(\tau))^{-1} e^{b\times\tau}. \tag{4.27}$$

Result 4.1.7B Where the long term non-observable spot rates follow an exponential progression of the form:

$$s_t(\tau) = a.e^{b\times\tau}, \tau > M;$$

Then we can write:

$$\frac{\partial P_t(\tau)}{\partial b} = -\tau P_t(\tau) \times (1 + s_t(\tau))^{-1} \tau s_t(\tau). \tag{4.28}$$

We now need to derive the partial derivatives of the extrapolation parameters to the observable spot rates:

Result 4.1.7C Following Result 3.3.2; for $M - q + 1 <= k <= M$

$$\frac{db}{ds_t(k)} = \frac{1}{s_t(k)} \times \frac{T_k - \frac{\sum_{s=M-q+1}^{M} T_s}{q}}{\sum_{s=M-q+1}^{M} T_s^2 - \frac{(\sum_{s=M-q+1}^{M} T_s)^2}{q}}. \tag{4.29}$$

Result 4.1.7D Following Result 3.3.2; for $M - q + 1 <= k <= M$

$$\frac{da}{ds_t(k)} = a \times \frac{(\frac{1}{s_t(k)} - \frac{db}{ds_t(k)} \times \sum_{s=M-q+1}^{M} T_s)}{q}. \tag{4.30}$$

The proofs of these results have been excluded as they follow from the respective definitions.

Bringing together the Results 4.1.7A-D, we get:

Result 4.1.7E

$$dP_t(\tau) = \frac{\partial P_t(\tau)}{\partial a} \times \sum_{s=M-q+1}^{M} \frac{da}{ds_t(s)} ds_t(s) + \frac{\partial P_t(\tau)}{\partial b} \times \sum_{s=M-q+1}^{M} \frac{db}{ds_t(s)} ds_t(s). \tag{4.31}$$

The result follows from the fact that:

$$P_t(\tau) = f^1(a, b),$$
$$a = f^2(s_t(M - q + 1), ..., s_t(M)),$$
$$b = f^3(s_t(M - q + 1), ..., s_t(M)).$$

4.1.8 Power Spot Rate Extrapolation

Result 4.1.8A Where the long term non-observable spot rates follow a power progression of the form:

$$s_t(\tau) = a.\tau^b, \tau > M;$$

Then we can write:

$$\frac{\partial P_t(\tau)}{\partial a} = -P_t(\tau) \times (1 + s_t(\tau))^{-1}\tau^b. \tag{4.32}$$

Result 4.1.8B Where the long term non-observable spot rates follow an exponential progression of the form:

$$s_t(\tau) = a.\tau^b, \tau > M;$$

Then we can write:

$$\frac{\partial P_t(\tau)}{\partial b} = -\tau P_t(\tau) \times (1 + s_t(\tau))^{-1} s_t(\tau) ln(b). \tag{4.33}$$

We now need to derive the partial derivatives of the extrapolation parameters to the observable spot rates:

Result 4.1.8C Following Result 3.3.2; for $M - q + 1 <= k <= M$

$$\frac{db}{ds_t(k)} = \frac{1}{s_t(k)} \times \frac{ln(T_k) - \frac{\sum_{s=M-q+1}^{M} ln(T_s)}{q}}{\sum_{s=M-q+1}^{M} ln(T_s)^2 - \frac{(\sum_{s=M-q+1}^{M} ln(T_s))^2}{q}}. \tag{4.34}$$

Result 4.1.8D Following Result 3.3.2; for $M - q + 1 <= k <= M$

$$\frac{da}{ds_t(k)} = a \times \frac{(\frac{1}{s_t(k)} - \frac{db}{ds_t(k)} \times \sum_{s=M-q+1}^{M} ln(T_s))}{q}. \tag{4.35}$$

The proofs of these results have been excluded as they follow from the respective definitions.

Bringing together the Results 4.1.8A-D, we get:

Result 4.1.8E

$$dP_t(\tau) = \frac{\partial P_t(\tau)}{\partial a} \times \sum_{s=M-q+1}^{M} \frac{da}{ds_t(s)} ds_t(s) + \frac{\partial P_t(\tau)}{\partial b} \times \sum_{s=M-q+1}^{M} \frac{db}{ds_t(s)} ds_t(s). \tag{4.36}$$

The result follows from the fact that:

$$P_t(\tau) = f^1(a, b),$$
$$a = f^2(s_t(M - q + 1), ..., s_t(M)),$$
$$b = f^3(s_t(M - q + 1), ..., s_t(M)).$$

4.2 Deriving the Greeks: Advanced Extrapolation Parameters

We now calculate the relevant partial derivatives of the price of the zero-coupon bond, with repect to the extrapolation paramteres, for each "advanced" extrapolation procedure.

4.2.1 Nelson-Siegel Method

For $\tau > M$, we have

$$P_t(\tau) = (1 + s_t(\tau))^{-\tau}.$$

where

$$s_t(\tau) = \beta_{1,t}h_{1,t}(\tau) + \beta_{2,t}h_{2,t}(\tau) + \beta_{3,t}h_{3,t}(\tau), \tag{4.37}$$

and

$$h_{1,t}(\tau) = 1, h_{2,t}(\tau) = \frac{1 - e^{-\frac{\tau}{\lambda}}}{\frac{\tau}{\lambda}}, h_{3,t}(\tau) = \frac{1 - e^{-\frac{\tau}{\lambda}}}{\frac{\tau}{\lambda}} - e^{-\frac{\tau}{\lambda}}.$$

Result 4.2.1A For $\tau > M$ and $i\epsilon\{1, 2, 3\}$:

$$\frac{dP_t(\tau)}{d\beta_{i,t}} = \tau.P_t(\tau)(1 + s_t(\tau))^{-1}h_{i,t}. \tag{4.38}$$

From Result 3.4.1 we can derive the following:

Result 4.2.1B

$$\frac{d\hat{\beta}_t}{ds_t(T_i)} = (H'_N H_N)^{-1}(\underline{h}^i_N). \tag{4.39}$$

where

$$h^i_N(3 \times 1) = \begin{matrix} h_{1,t}(T_i) \\ h_{2,t}(T_i) \\ h_{3,t}(T_i) \end{matrix} .$$

Under the Nelson-Siegel framework we can write $P_t(\tau) = f(\beta_{1,t}, \beta_{2,t}, \beta_{3,t})$ and $\underline{\hat{\beta}_t} = f(s_t(T_1), s_t(T_2), ..., s_t(T_M))$. Bringing the above two results together we get:

Result 4.2.1C

$$dP_t(\tau) = \sum_{i=1}^{3} \frac{\partial P_t(\tau)}{\partial \hat{\beta}_{i,t}} \sum_{j=1}^{M} \frac{d\hat{\beta}_{i,t}}{ds_t(T_j)} ds_t(T_j). \tag{4.40}$$

4.2.2 Svensson Method

For $\tau > M$, we have:

$$P_t(\tau) = (1 + s_t(\tau))^{-\tau},$$

where

$$s_t(\tau) = \beta_{1,t}h_{1,t}(\tau) + \beta_{2,t}h_{2,t}(\tau) + \beta_{3,t}h_{3,t}(\tau) + \beta_{4,t}h_{4,t}(\tau), \tag{4.41}$$

and

$$h_{1,t}(\tau) = 1, h_{2,t}(\tau) = \frac{1 - e^{-\frac{\tau}{\lambda_1}}}{\frac{\tau}{\lambda_1}}, h_{3,t}(\tau) = \frac{1 - e^{-\frac{\tau}{\lambda_1}}}{\frac{\tau}{\lambda_1}} - e^{-\frac{\tau}{\lambda_1}},$$

$$h_{4,t}(\tau) = \frac{1 - e^{-\frac{\tau}{\lambda_2}}}{\frac{\tau}{\lambda_2}} - e^{-\frac{\tau}{\lambda_2}}.$$

Result 4.2.2A For $\tau > M$ and $i\epsilon\{1, 2, 3, 4\}$:

$$\frac{dP_t(\tau)}{d\beta_{i,t}} = \tau P_t(\tau)(1 + s_t(\tau))^{-1}h_{i,t}. \tag{4.42}$$

From Result 3.4.2 we can derive the following:

Result 4.2.2B

$$\frac{d\hat{\beta}_t}{ds_t(T_i)} = (H'_S H_S)^{-1}(\underline{h}^i_S). \tag{4.43}$$

where

$$h^i_S(4 \times 1) = \begin{matrix} h_{1,t}(T_i) \\ h_{2,t}(T_i) \\ h_{3,t}(T_i) \\ h_{4,t}(T_i) \end{matrix} .$$

Under the Svensson framework we can write $P_t(\tau) = f(\beta_{1,t}, \beta_{2,t}, \beta_{3,t}), \beta_{4,t})$
and $\underline{\hat{\beta}_t} = f(s_t(T_1), s_t(T_2), ..., s_t(T_M))$. Bringing the above two results together we get:

Result 4.2.2C

$$dP_t(\tau) = \sum_{i=1}^{4} \frac{\partial P_t(\tau)}{\partial \hat{\beta}_{i,t}} \sum_{j=1}^{M} \frac{d\hat{\beta}_{i,t}}{ds_t(T_j)} ds_t(T_j). \tag{4.44}$$

4.2.3 Cairns Method

For $\tau > M$, we have:

$$P_t(\tau) = (1 + s_t(\tau))^{-\tau},$$

where

$$s_t(\tau) = \beta_{0,t}h_{0,t}(\tau) + \beta_{1,t}h_{1,t}(\tau) + \beta_{2,t}h_{2,t}(\tau) + \beta_{3,t}h_{3,t}(\tau) + \beta_{4,t}h_{4,t}(\tau), \quad (4.45)$$

and

$$h_{0,t}(\tau) = 1,$$

$$h_{i,t}(\tau) = \tfrac{1-e^{-c_i\tau}}{c_i\tau} \text{ for i = 1 to 4.}$$

Result 4.2.3A For $\tau > M$ and $i \epsilon \{1, 2, 3, 4\}$:

$$\frac{dP_t(\tau)}{d\beta_{i,t}} = \tau P_t(\tau)(1 + s_t(\tau))^{-1}h_{i,t}. \quad (4.46)$$

From Result 3.4.3 we can derive the following:

Result 4.2.3B

$$\frac{d\hat{\beta}_t}{ds_t(T_i)} = (H_C'H_C)^{-1}(\underline{h}_C^i), \quad (4.47)$$

where

$$h_C^i(5 \times 1) = \begin{matrix} h_{0,t}(T_i) \\ h_{1,t}(T_i) \\ h_{2,t}(T_i) \\ h_{3,t}(T_i) \\ h_{4,t}(T_i) \end{matrix} .$$

Under the Cairns framework we can write $P_t(\tau) = f(\beta_{0,t}, \beta_{1,t}, \beta_{2,t}, \beta_{3,t}), \beta_{4,t})$ and $\underline{\hat{\beta}_t} = f(s_t(T_1), s_t(T_2), ..., s_t(T_M))$. Bringing the above two results together we get:

Result 4.2.3C

$$dP_t(\tau) = \sum_{i=1}^{5} \frac{\partial P_t(\tau)}{\partial \hat{\beta}_{i,t}} \sum_{j=1}^{M} \frac{d\hat{\beta}_{i,t}}{ds_t(T_j)} ds_t(T_j) \quad (4.48)$$

4.2.4 Smith-Wilson Method

Result 4.2.4A For $\tau > M$:

$$\frac{\partial P_t(\tau)}{\partial \xi_{i,t}} = W(\tau, u_i). \quad (4.49)$$

These results all follow directly from the respective definitions of $P_t(\tau)$.

From Result 3.4.4 we can derive the following result:

Result 4.2.4B

$$\frac{d\hat{\xi}_t}{\underline{df_t(T_i)}} = W^{-1}[\frac{dP_t}{\underline{df_t}}], \tag{4.50}$$

where

$$\frac{dP_t}{\underline{df_t(T_i)}}(M \times 1) = \begin{matrix} \frac{dP_t(T_1)}{df_t(T_i)} \\ \frac{dP_t(T_2)}{df_t(T_i)} \\ \dots \\ \frac{dP_t(T_M)}{df_t(T_i)} \end{matrix}, W(M \times M) = \begin{matrix} W(T_1,T_1) & W(T_1,T_2) & \dots & W(T_1,T_M) \\ W(T_2,T_1) & W(T_2,T_2) & \dots & W(T_2,T_M) \\ & \dots & & \\ W(T_M,T_1) & W(T_M,T_2) & \dots & W(T_M,T_M) \end{matrix}.$$

Under the Smith-Wilson framework we can write $P_t(\tau) = f(\underline{\hat{\xi}_t})$ and $\underline{\hat{\xi}_t} = f(f_t(T_1), f_t(T_2), ..., f_t(T_M))$. Bringing the above two results together we get:

Result 4.2.4C

$$dP_t(\tau) = \sum_{i=1}^{M} \frac{\partial P_t(\tau)}{\partial \hat{\xi}_{i,t}} \sum_{j=1}^{M} \frac{d\hat{\xi}_{i,t}}{df_t(T_j)} df_t(T_j). \tag{4.51}$$

4.3 Summary

In this chapter we have examined each of the proposed extrapolation methods and, in turn, we have expressed the dynamics of the pricing function in terms of the observable yield curve. This has been done as follows:

- We assume that (under each extrapolation method), information which influences the yield curve beyond the maximum term (M) is fully reflected in the observable yield curve.
- We derive the partial derivatives of the pricing / extrapolation function, with respect to the extrapolation parameters.
- We derive the partial derivatives of the pricing / extrapolation function, with respect to the observable points on the yield curve.
- We derive an equation which expresses continuous movements in the pricing / extrapolation function as a function of movements in the observable yield curve.

In the chapters which follow we will use the information that we have derived to construct dynamic hedges (replicating portfolios) of various long term interest rate risks.

5. HEDGING LONG TERM INTEREST RATES: CASE STUDY 1

In this chapter we build a case study where we look to hedge a long term hypothetical zero coupon bond by building a theoretical replicating portfolio from observable / tradable bonds. We base this study on observations of weekly data for the South African swap curve for the period from 21/8/2000 to 5/3/2007. This chapter is structured as follows:

- Description of the swap data for 21/08/2000 - 5/3/2007.
- Deriving annual spot rates from observable swap rates.
- Description of historical data: PCA Results.
- Results for the simple extrapolations.
- Illustration and discussion of hedging errors.
- Results for the advanced extrapolations.

5.1 Historical Swap Data

Historical South African swap curve data has been used, covering the period from 21/8/2000 - 5/3/2007. Weekly closing at-the-money annualised swap rates were obtained for terms of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, and 30 years. The following table provides descriptive statistics for the data:

Term	*Average*	*StdDev*	*Minimum*	*Maximum*	*Skewness*
1	9.28%	1.89%	6.76%	13.55%	43.55%
2	9.39%	1.68%	6.92%	13.06%	28.09%
3	9.56%	1.59%	7.07%	12.79%	21.56%
4	9.70%	1.58%	7.19%	12.93%	20.39%
5	9.80%	1.59%	7.25%	13.02%	22.46%
6	9.87%	1.60%	7.26%	13.25%	25.05%
7	9.90%	1.61%	7.26%	13.44%	27.74%
8	9.93%	1.62%	7.26%	13.57%	30.38%
9	9.93%	1.62%	7.26%	13.65%	32.24%
10	9.93%	1.63%	7.26%	13.68%	33.92%
15	9.75%	1.63%	7.17%	13.55%	44.59%
20	9.53%	1.64%	6.99%	13.46%	51.30%
25	9.35%	1.63%	6.86%	13.37%	58.14%
30	9.17%	1.63%	6.75%	13.30%	65.94%

Tab. 5.1: Descriptive statistics for weekly ZAR swap rates for 21 August 2000 - 5 March 2007

5.2 Interpolating Swap Data

Interpolation of the observed swap rates has been performed using the "Bessel cubic spline" approach as described by Hagan and West (2006). Although a number of approaches were tested:

5.2.1 Linear Interpolation

Firstly, a linear interpolation of swap rates was performed. However, this approach was found to generate irregularities (volatility / discontinuities) in forward rates across nodes. Figure 5.1 shows how this occurs when interpolating the observed swap rates at 26 December 2005. The effect of the irregularities at the nodes was found to produce spurious results at later stages of the analysis when hedging was performed.

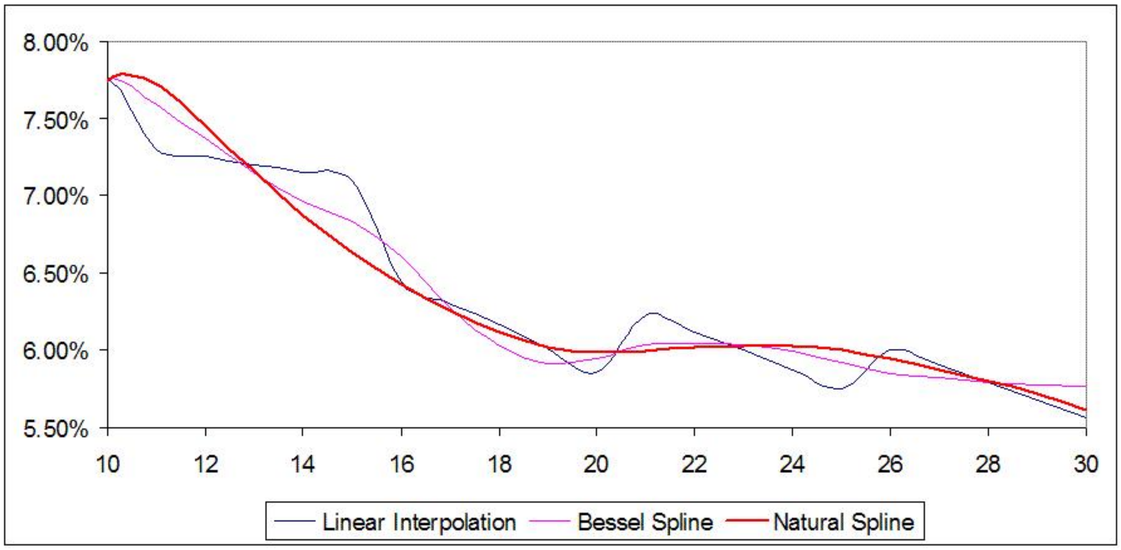


Fig. 5.1: Comparison of interpolated Forward Rates by Term

5.2.2 Bessel Cubic Spline

Secondly, a Bessel interpolation of swap rates was performed (as described by Hagan and West (2006)). This approach uses a quadratic function to make some clever estimates of the slope in the spot curve at each node. It has the advantage that it is both intuitive and easy to implement.

This approach is sometimes unreliable when data is sparse, though it seems to perform well for the range of data that we are using. Another flaw in this approach lies in the fact that it does not guarantee continuous forward rates which is also indicated by Figure 5.1.

5.2.3 Natural Cubic Spline

Thirdly, a Natural interpolation of swap rates was performed (as described by Hagan and West (2006)). This is a more complicated extrapolation based on the assumption that the second order derivative of the spot rate is continuous. The approach produced a much more stable set of forward rates as shown in Figure 5.1. This approach has been used in results which follow.

At this stage it is worth noting that the choice of interpolation procedure can have a significant impact on the results obtained from this analysis. It is particularly significant in influencing the results obtained from the Simple extrapolation methods. Results were generally improved by using either the Bessel or Natural spline interpolations, though the ultimate conclusions of the analysis did not change across the interpolation procedures tested.

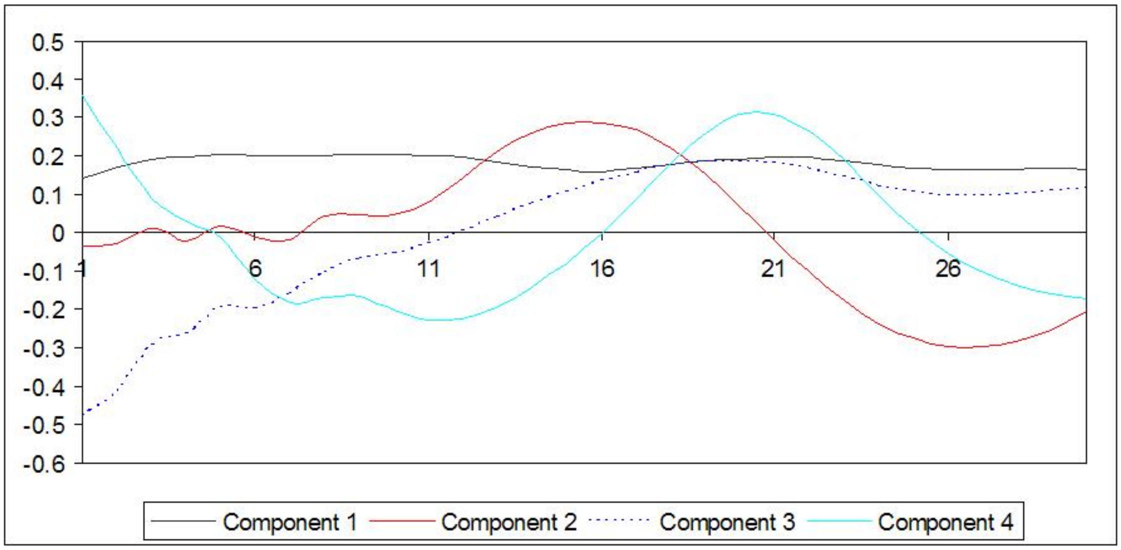


Fig. 5.2: Coefficients of first 4 principal components - Full Yield Curve PCA

5.3 Principal Component Analysis

Principal Component Analysis is often a useful tool in describing yield curve behaviour over a particular period. We therefore perform two principal component analyses on weekly spot rate movements over the historical period:

5.3.1 PCA on Full Curve

Firstly, we perform a PCA on the spot rates (at annual intervals) over the full term of the observable yield curve. Hence we include all spot rate movements from terms 1 to 30.

The following table shows the results of the analysis:

Principal Component	*Proportion Variability Explained*
1	73.95%
2	14.12%
3	5.16%
4	2.47%
5+	4.30%

Tab. 5.2: ZAR Principal Component Analysis - Proportion Variability Explained by each Component

The results in Table 5.2 indicate that over 93% of the movement in the full yield curve over the period 21 August 2000 - 5 March 2007 can be explained by the first 3 components.

Figure 5.2 indicates that the absolute coefficients for the first principal component are relatively flat over term, although they seem to increase initially and reduce at later durations. This component can be regarded as a factor affecting the yield curve approximately equally at all observable terms, i.e. a level shift. The second component has a seemingly spurious effect over the first 10 years, but beyond this represents an inversion / dis-inversion type

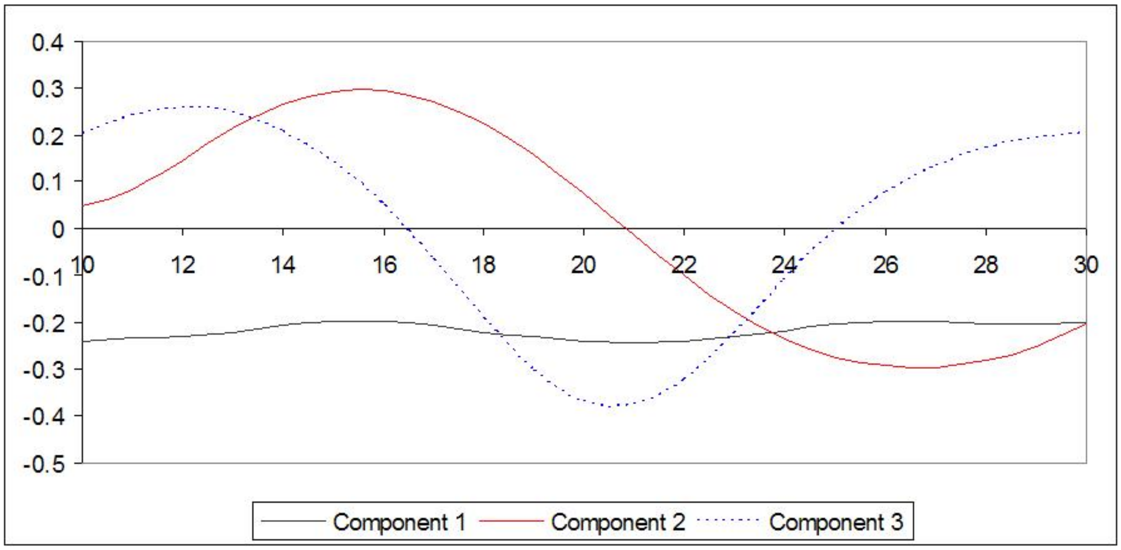


Fig. 5.3: Coefficients of first 3 principal components - Partial Yield Curve PCA

movement. This is not surprising since inversion of the yield curve has characterised the recent yield curve movements. The third component is initially negative and increases to positive at long terms, representing a twist in the yield curve. The unusual kinks in the curves are likely to be due to sampling error in deriving the correlation matrix.

5.3.2 PCA on Partial Curve

Secondly, we perform a PCA on the spot rates (at annual intervals) over the longer term of the observable yield curve. Hence we only include spot rates from terms 10 to 30. This analysis will exclude any factors which operate predominantly at the short end of the curve, since we are only interested in movements at the long end of the curve.

The following table shows the results of the analysis:

Principal Component	*Proportion Variability Explained*
1	72.96%
2	20.03%
3	3.29%
4+	3.72%

Tab. 5.3: ZAR Principal Component Analysis - Proportion Variability Explained by each Component

The results in Table 5.3 indicate that over 96% of the movement in the long end of the yield curve over the period 21 August 2000 - 5 March 2007 can be explained by the first 3 components.

Figure 5.3 indicates the coefficients for the first 3 principal components. Notice that the components are different from those in Figure 5.2. We still see the "flat" component dominating movements in spot rates, while the "inversionary" effect is still the second most important factor, albeit with a higher relevance for the long end. Interestingly, the third factor seems to correspond more closely to the fourth factor in the full yield curve analysis. It seems that analysing the long end of the curve has caused a reordering of the 3rd and 4th factors.

Note that these results are slightly different from those of Maitland (2002). Maitland's results were based on government bonds yields for the period of 1986 to 1998, and also stress the relevance of the first two (level and slope) principal components. However, the key difference in results relates to the more recent inversion that we have observed in the long end of the curve. It remains to be seen whether this inversion effect has been a temporary factor affecting the yield curve, or whether it will continue to play a role in future.

5.4 Simple Extrapolations: Hedging Results

In this section we present the results of a historical analysis on the effectiveness of the simple extrapolations / forecasting approaches. We begin with a brief description of the analysis performed to obtain the hedging results.

5.4.1 Methodology

The analysis works as follows:

Firstly, we have used the weekly spot curve as derived from observed swap rates (described above).

Secondly, at the start of each week we assume that there exists an entity that has a liability to pay a fixed amount of R N in 50 years, such that the discounted value of N is R 1 million. It is necessary to impose this requirement in order to standardise the results of the analysis across different time periods and different forecasting approaches, as these all place different present values on a 50 year zero coupon bond when notionals are equivalent.

Thirdly, we look to hedge this liability by implementing a hedge at the start of each week which is **rho neutral**. We derive the rho-neutral hedge (for each extrapolation procedure) as a set of bonds in the tradable section of the yield curve using the results of our analysis in Chapter 4.

Fourthly, at the end of each week the hedging error (surplus / strain) is calculated based on the updated yield curve. The error is then recorded, the term of the liability is reset to 50 years and the analysis is repeated for the next week. Once again, we have imposed this requirement (resetting the term of the liability to 50 years) in order to standardise results across each time step.

5.4.2 Illustration of Hedging Errors

An important question arises out of this approach: **If it is possible to build a replicating portfolio for all of these forecast interest rates, then why are these hedges not necessarily perfect?**

The answer is quite subtle. In order to perfectly hedge a zero coupon bond with outstanding maturity X, such that $X > M$, our hedge needs to satisfy the following two conditions continuously:

1. $\frac{\partial P_t(X)}{\partial s_t(q)} = \frac{\partial H_t}{\partial s_t(q)}, \forall 0 < q < M$, i.e. the hedge position is rho neutral

2. $\frac{\partial P_t(X)}{\partial t} = \frac{\partial H_t}{\partial t}$, i.e. the hedge position is theta neutral

The hedge portfolio that we have derived for each forecasting approach is not perfect because it does not necessarily give a theta neutral position. This is not easily rectified; in fact it seems reasonable to suggest that the position would only be theta neutral if the forecasting approach were able to consistently forecast the true yield curve beyond the maximum term of M years. However, this is very useful since it provides us with a means of quantitatively validating any particular forecasting approach. Our reasoning works as follows:

For any particular forecasting approach we can build a theoretical rho neutral hedging strategy. In continuous space the partially hedged position will yield surpluses / deficits (hedging errors), which can be attributed to the theta effect described above. In discrete space the theta effect will be confounded by second order and interactive effects. However, it seems reasonable to suggest that better forecasting approaches should yield consistently smaller hedging errors.

5.4.3 Results

We now provide the results of the hedging analysis for each of the simple extrapolation approaches described in Chapter 3. In order to provide a reasonable basis for comparison of results, we have also provided the hedging errors that arise from an approach that is commonly used in practice.

We show the hedging errors that arise from using a Flat Spot Rate extrapolation and hedging with a long position in a coupon bearing bond only. (We have used a 30 year 6% coupon bearing bond for the purpose of illustration.) We will refer to this as the Benchmark Approach

The results of the historical analysis (rounded to the nearest hundred rand) are as follows:

Statistic	*Benchmark*	*Flat Spot Rate*	*Lin. Spot Rate*	*Pwr. Spot Rate*	*Exp. Spot Rate*
95% VAR	(172 000)	(12 600)	(98 600)	(31 200)	(43 500)
CTE[85%]	(157 900)	(13 400)	(107 000)	(30 400)	(44 300)
Mean	(10 000)	(3 800)	(22 300)	(7 800)	(9 000)
Minimum	(360 200)	(88 900)	(747 900)	(138 700)	(247 300)
Maximum	416 900	(1 400)	8 500	2 700	11 000

Tab. 5.4: CS1: Results of simple spot rate extrapolations

Figure 5.4 graphically illustrates the results for the various simple extrapolation procedures.

These results seem to indicate that it is possible to achieve a more effective hedge than simply matching with a long dated coupon bearing bond. Creating a rho hedge of a Flat Spot Rate extrapolation provides a substantial reduction in historically based risk measures.

Most notably from the above results, it seems that extrapolation techniques which display greater stability tend to produce better results. This is not surprising because the impact of second order and interactive effects will be greater when the forecast approach is less stable.

Statistic	*Benchmark*	*Flat Fwd Rate*	*Lin. Fwd Rate*	*Pwr. Fwd Rate*	*Exp. Fwd Rate*
95% VAR	(172 000)	(216 600)	(1 079 900)	(89 900)	(163 300)
CTE[85%]	(157 900)	(189 000)	(1 130 300)	(253 200)	(625 400)
Mean	(10 000)	(35 300)	(173 700)	(34 800)	(81 700)
Minimum	(360 200)	(827 400)	(6 312 600)	(3 003 600)	(9 554 900)
Maximum	416 900	(0)	1 721 300	325 100	920 600

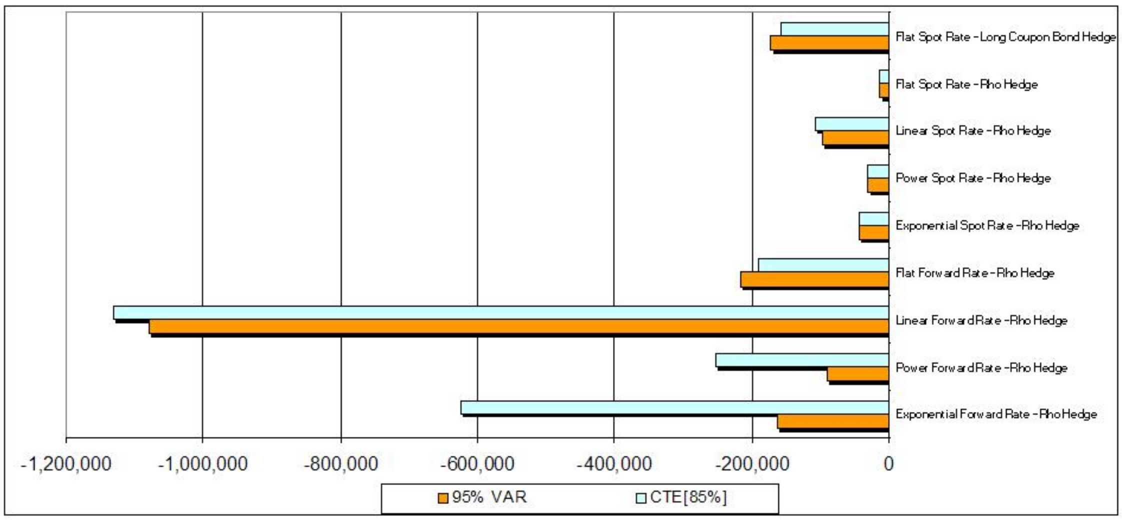


Fig. 5.4: CS1: VAR and CTE results for the simple extrapolation approaches

Linear extrapolation techniques are the most unstable from those considered and this can be seen from the relative size of the historically based risk measures. It is interesting that linear extrapolations of forward rates are not uncommon in practice! These results indicate that such a forecasting approach is not adequate for quantitative and hedging purposes.

Power and exponential extrapolations tend to perform better than linear extrapolations. Since the long end of the yield curve is generally downward sloping, these approaches are relatively more stable as the long rate tends towards zero.

The use of a Flat Spot Rate approach performs well in comparison to the rest of the approaches because its simplicity makes it easy to hedge. Therefore the combined position has relatively little exposure to second order interest rate risks and hedging errors tend to be primarily in respect of theta errors.

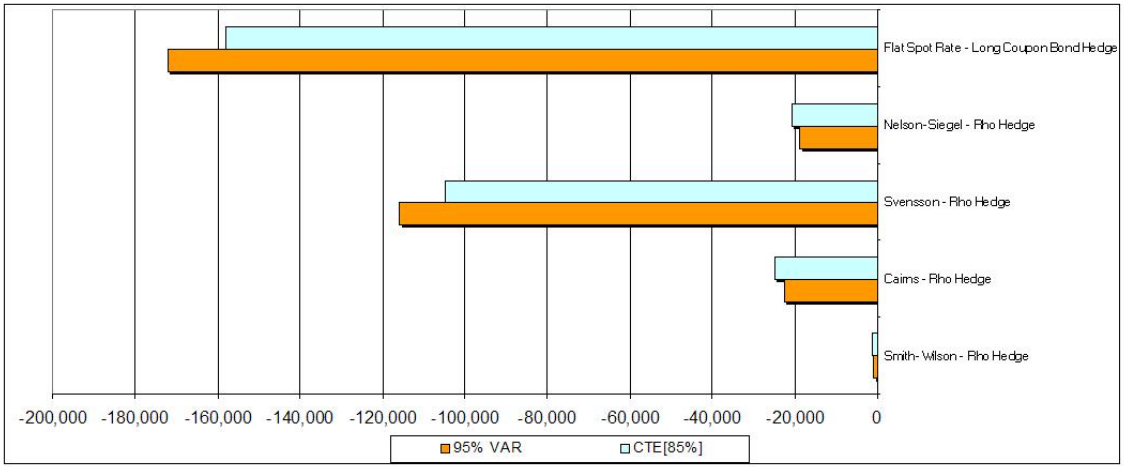


Fig. 5.5: CS1: VAR and CTE results for the advanced extrapolation approaches

5.5 Advanced Extrapolations: Hedging Results

We follow the same methodology described above to build and analyse hedges for the advanced extrapolation approaches. For ease of comparison we again show the results obtained from the Benchmark Approach described above. Results are as follows:

Statistic	*Benchmark*	*Nelson − Siegel*	*Svensson*	*Cairns*	*Smith − Wilson*
95% VAR	(172 000)	(18 800)	(115 900)	(22 600)	(1 000)
CTE[85%]	(157 900)	(20 800)	(104 800)	(24 900)	(1 000)
Mean	(10 000)	(4 900)	(27 800)	(5 500)	(400)
Minimum	(360 200)	(156 800)	(466 600)	(139 900)	(5 500)
Maximum	416 900	3 200	13 600	3 800	1 300

Tab. 5.6: CS1: Results of advanced extrapolations

Figure 5.5 graphically illustrates the results for the various advanced extrapolation procedures. Appendix B shows the distribution of residuals for the advanced hedging approaches. We will now discuss the results of each approach in more detail.

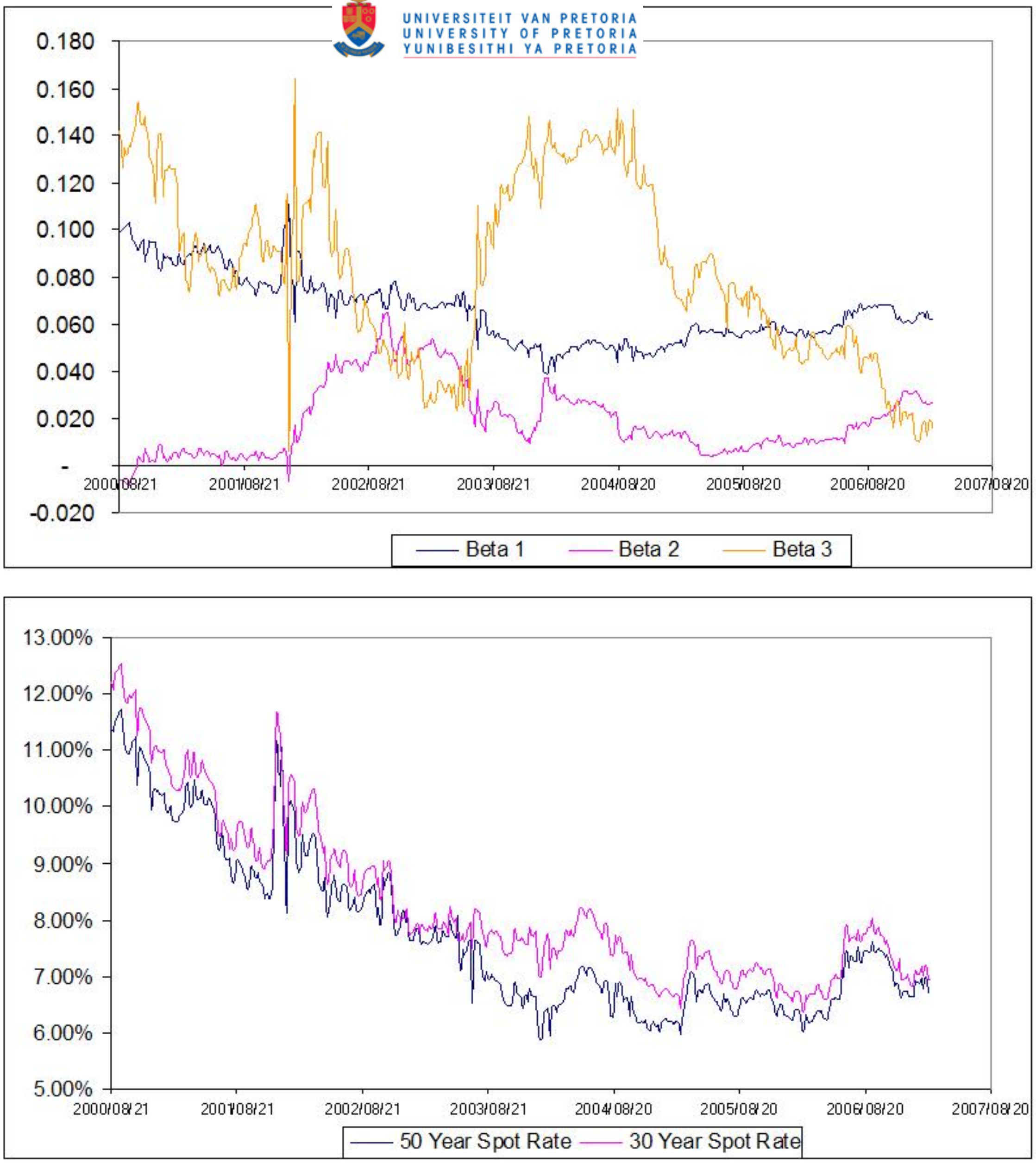
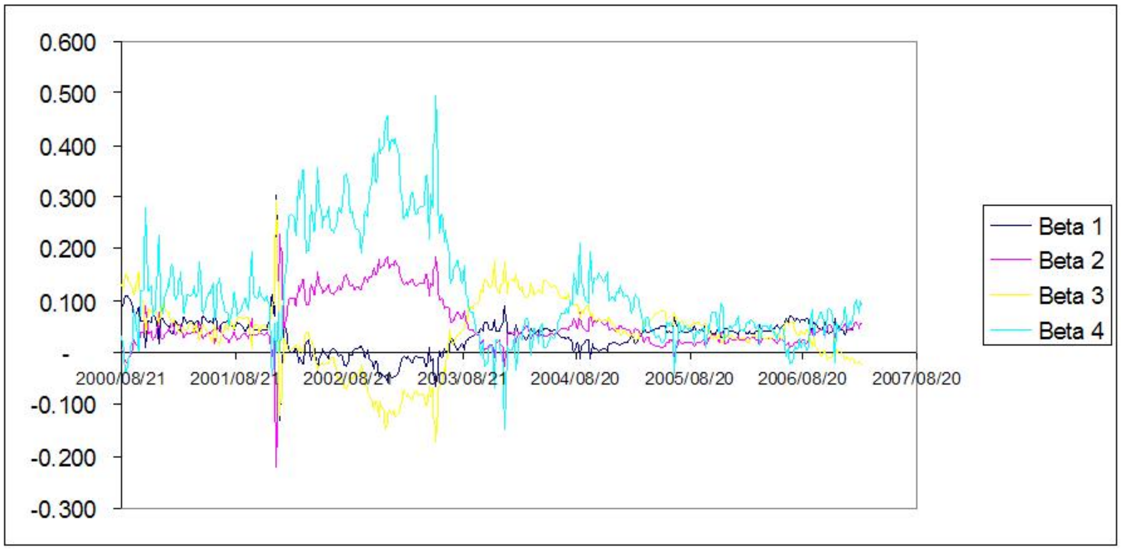


Fig. 5.6: CS1: Progression of Beta parameters, and 50 year vs 30 year spot rates, for Nelson-Siegel model

5.5.1 Nelson-Siegel Results

Figure 5.6 shows how the β parameters progress over the historical period of investigation.

Figure 5.6 also shows how the 50 year projected spot rate compares to the 30 year spot rate over the historical period of investigation. Notice that the margin between the 30 year and 50 year spot rates shows a large amount of volatility which is difficult to explain.



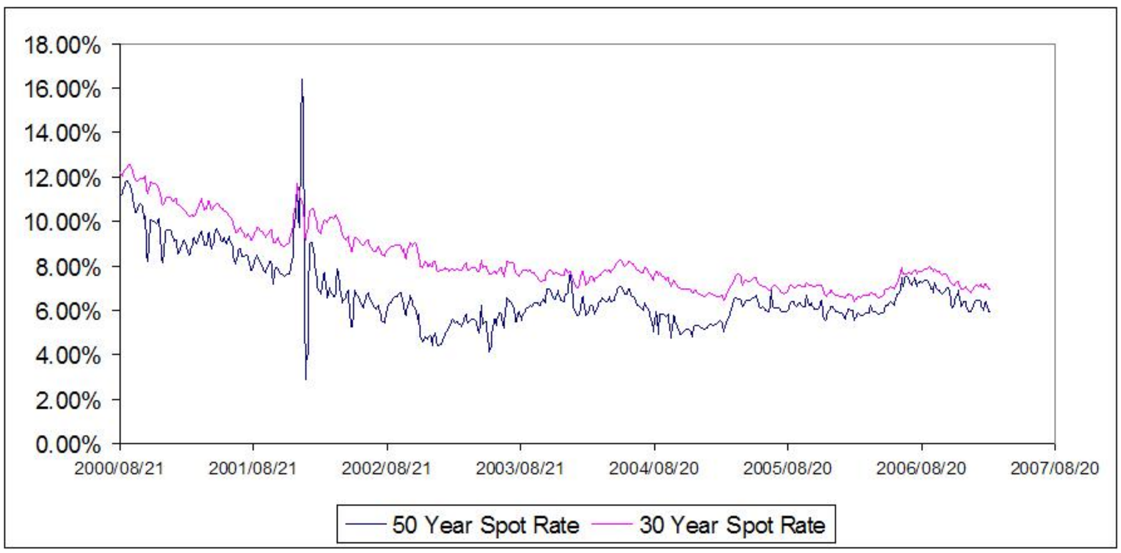


Fig. 5.7: CS1: Progression of Beta parameters, and 50 year vs 30 year spot rates, for Svensson model

5.5.2 Svensson Results

Figure 5.7 shows how the β parameters progress over the historical period of investigation.

Notice from Figure 5.7 that the margin between the 30 year and 50 year spot rates is more volatile than the margin for the Nelson-Siegel approach. This gives us some insight into the reasons for the poor hedging results in Table 5.6. Further, the volatility in beta parameters indicates a potentially large amount of multicollinearity between the 4 factors in the model.

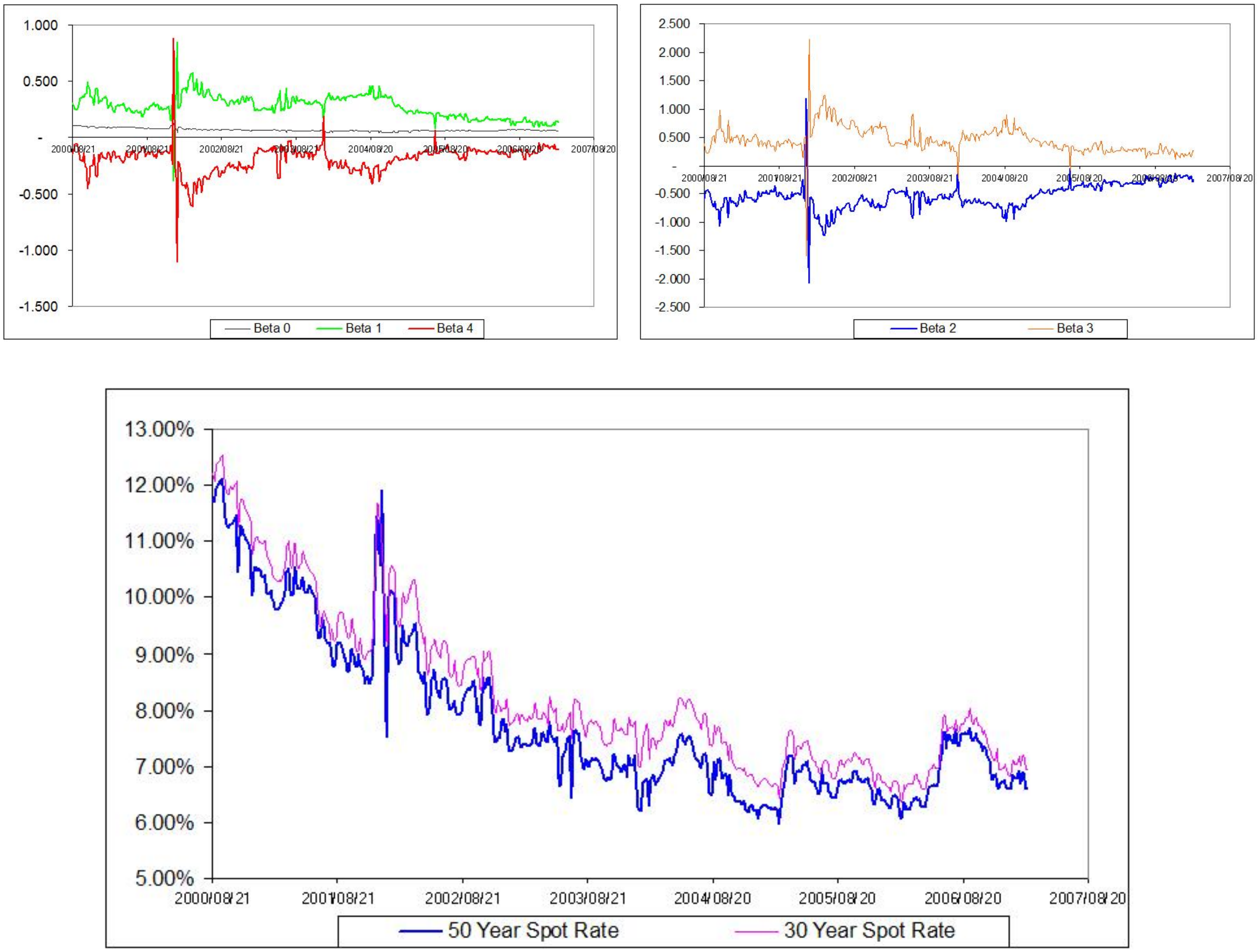


Fig. 5.8: CS1: Progression of Beta parameters, and 50 year vs 30 year spot rates, for Cairns model

5.5.3 Cairns Results

Figure 5.8 shows how the β parameters progress over the historical period of investigation.

Notice from Figure 5.8 that the margin between the 30 year and 50 year spot rates is more stable than the margin for the Nelson-Siegel and Svensson approaches. The β factors have been shown in separate figures for scaling purposes. β_0 is very stable as we would expect. Interestingly, β_1 and β_2 seem strongly positively correlated, but both strongly negatively correlated with β_3 and β_4. This suggests that the model could potentially be collapsed into fewer factors - at least over the historical period that we are analysing.

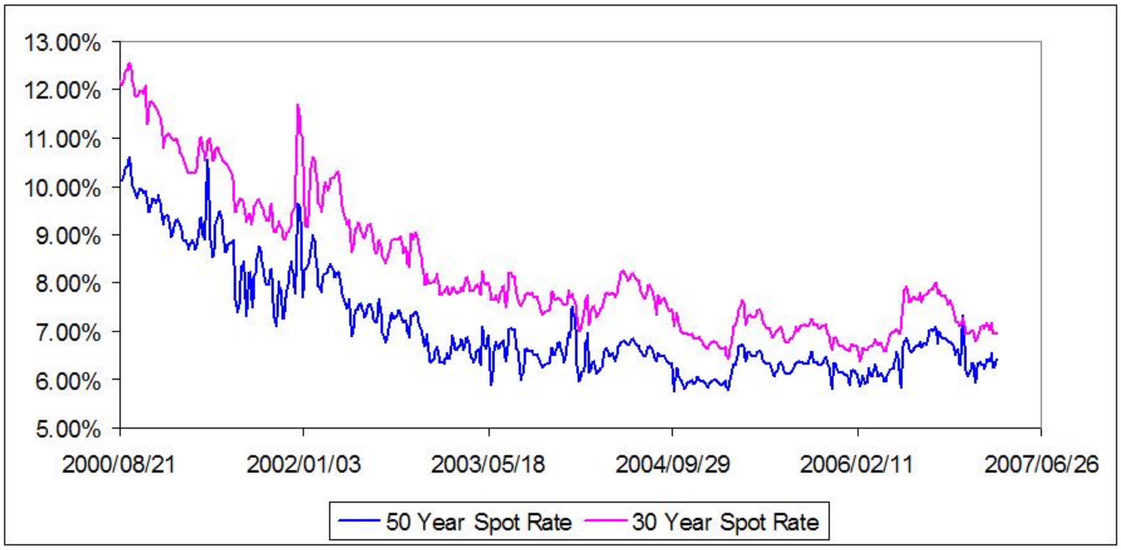


Fig. 5.9: CS1: 50 year vs 30 year spot rates, for Smith-Wilson model

5.5.4 *Smith-Wilson Results*

Figure 5.9 shows how the Smith-Wilson approach projects the 30 year and 50 year spot rates over the period of historical investigation. Despite the constraint that spot rates tend towards 5% as term increases towards infinity, the 50 year rate is able to vary substantially due to the high level of smoothness that we have imposed. Even in years where interest rates were substantially higher we find that the method still provides sensible results.

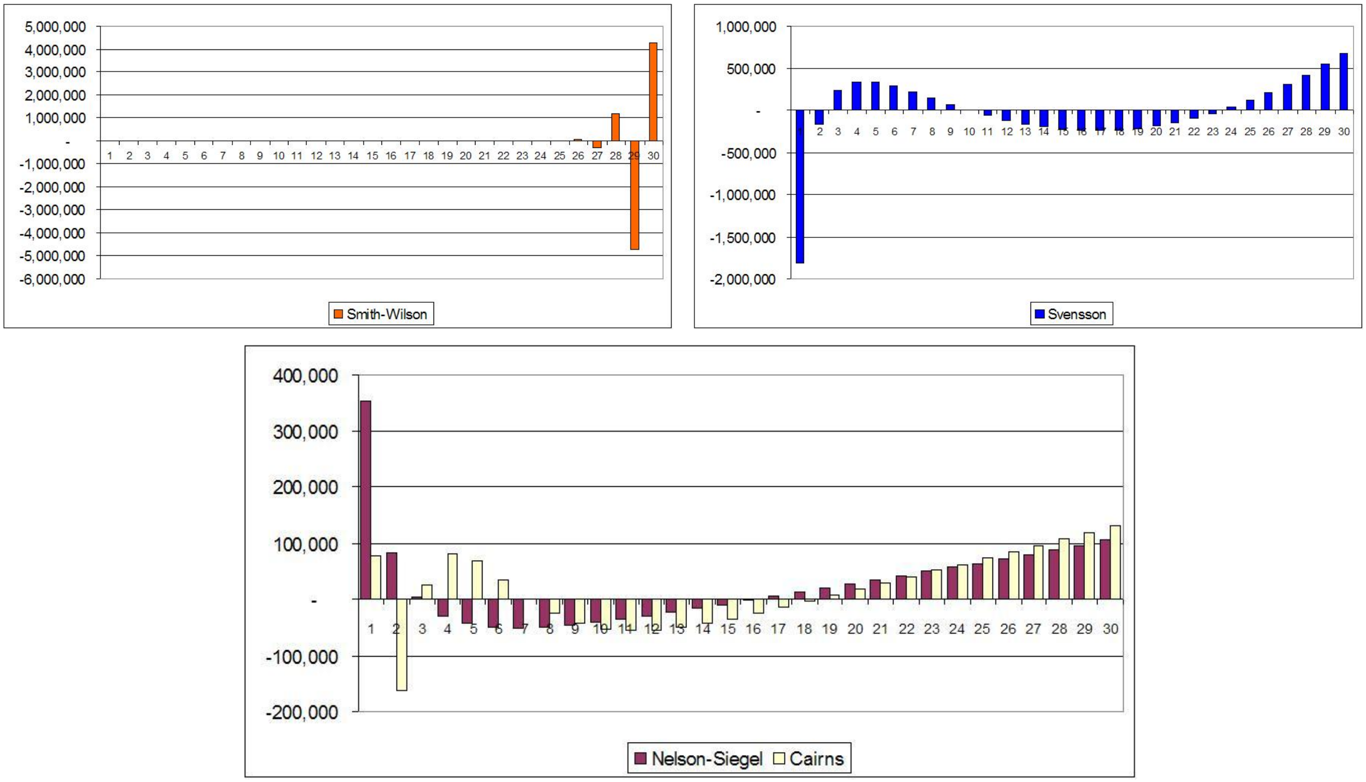


Fig. 5.10: CS1: Composition of hedge portfolio for (standardised) R1m 50 year ZCB, effective 26/2/2007

5.5.5 Comparison of Hedging Portfolios

Figure 5.10 illustrates the composition of the hedging portfolios for each of the advanced forecasting approaches. It is interesting to note that the functional form approaches, especially Nelson-Siegel and Cairns, have relatively "smooth" compositions when compared to the Smith-Wilson approach. By smooth we mean that there seems to be a high degree of correlation in the exposure of the hedge to adjacent zero coupon rates . This is a consistent feature of these approaches, regardless of the starting date / yield curve which is used. Further, the "shape" of the portfolios is relatively insensitive to the initial shape of the yield curve. More examples of the various hedging portfolios under alternative yield curves are provided in Appendix C.

The reason for this phenomenon is that the functional form approaches are less sensitive to changes at the long end of the observable term structure. This is because they represent a "fit" to the observable curve. However, the Smith-Wilson approach calibrates to achieve an exact fit to the observable input term structure, and therefore places a much higher reliance on the longest observable spot and forward rates.

From a practical perspective, it would be easier to implement the functional form approaches as less rebalancing is needed over time. A hedge based on a Smith-Wilson extrapolation requires a high amount of long exposure to the 30 year spot rate, with a high amount of short exposure to the 29 year spot rate. This is likely to require relatively more rebalancing than a hedge based on the functional form approaches, as their required exposure to spot rates tends to be a much smoother function of term.

5.6 Summary

We have previously identified three key problems faced by all entities who deal with long term interest rate risk:

- The non-observability of interest rates beyond the maximum term in the yield curve. Associated with this is the inability of entities to adequately quantify their interest rate exposure.
- The inadequacy of traditional methods (i.e. immunisation and bucketing) to mitigate long term interest rate risk.
- The lack of liquidity in long term interest rate markets. Associated with this is the inability of entities to adequately hedge their interest rate risk.

For companies who deal with long term interest rate risk, these problems are a reality. From a hedging perspective, many insurance entities have adopted the approach of holding the longest available coupon bearing bond to try hedge this interest rate risk. The analysis indicates that significant reductions in risk can be obtained through active hedging. However, it is necessary to adopt a forecasting approach in order to quantify exposure to long term interest rate risk. This case study is concerned with analysing a range of forecasting approaches (simple and advanced) in light of historical evidence.

The results obtained indicate that hedging based on more advanced forecasting approaches seems to yield substantial benefits from a risk perspective, particularly through the use of the Cairns and Smith-Wilson approaches, with Smith-Wilson yielding the largest reduction in risk measured by historical VAR.

Based on the historical risk measures it may seem that the Smith-Wilson approach provides an almost perfect solution for hedging long term interest rate risk. However, for hedging purposes, one weakness with the approach is its absolute reliance on the longest observable spot rate. This feature leads to hedging portfolios whose instruments are very heavily weighted toward the longest observable portion of the term structure.

Due to liquidity constraints at terms of 25 to 30 years, it may be argued that less reliance should be placed on these rates for hedging and risk management purposes. An advantage of the functional form approaches (Nelson-Siegel, Svensson, and Cairns) is the lower reliance on the 30 year spot rate for extrapolation purposes. As a result, the hedging portfolios derived using these methods tend to be more highly spread across the range of tradable and observable interest rates.

We now have evidence to suggest that fairly accurate hedging of long term (non-observable) interest rates can be achieved. However, we have already run into the question of cost / practicality vs reward as it seems that the cost (and complexity) of more accurate hedging strategies could be significantly higher.

5.7 Extension of Case Study: Hedging a 35 Year ZCB

Thus far we have only considered the case where we hedge a hypothetical 50 year zero coupon bond using the simple and advanced forecasting approaches described above. We have not considered whether different results would be obtained for different terms beyond the maximum tradable term of 30 years.

In order to address this concern, we perform an extension of the above analysis through hypothetical hedging of a 35 year zero coupon bond. (Results are included in Appendix D.) We note that there has been a slight reordering of results: the simple extrapolation approaches seem to perform relatively better as term shortens closer to 30 years, though the Smith-Wilson approach still seems to produce the best results.

6. HEDGING LONG TERM INTEREST RATES: CASE STUDY 2

In this chapter we build a second case study along similar lines to case study 1 above. However, in this case we use out-of-sample yield curve data which we generate from a multi-factor yield curve model described by Cairns (2004). This chapter is structured as follows:

- Description of the yield curve model used.
- Description of the data.
- Description of the methodology.
- Results for the simple extrapolations.
- Results for the advanced extrapolations.

6.1 Description of the Yield Curve Model

Cairns (2004) describes a family of term structure models which can be used for long term projections of the yield curve. The family is based on the Flesaker and Hughston (1996) positive-interest framework, where it is proposed that zero coupon bond prices are defined by:

$$P(t,T) = \frac{\int_T^\infty M(t,s)\phi(s)ds}{\int_t^\infty M(t,s)\phi(s)ds}, \tag{6.1}$$

for some deterministic function $\phi(s)$.

Cairns shows that the model for M(t,T) is defined by:

$$M(0,T) = 1, \forall\, T, \tag{6.2}$$

$$dM(t,T) = M(t,T)\sum_{i=1}^{n} \sigma_i(t,T)d\hat{Y}_i(t), \tag{6.3}$$

where

- $\hat{Y}(0) = 0$,
- $\hat{Y}_i(t)$ are n Brownian motions under a real world measure, with $d\hat{Y}_i(t)d\hat{Y}_j(t) = \rho_{ij}dt$.

Cairns then goes on to show that, by assuming $\sigma_i(t,T) = \sigma_i e^{-\alpha_i(T-t)}$, we can express P(t,T) as follows:

$$P(t,T) = \frac{\int\limits_{T-t}^{\infty} H(u, X(t))du}{\int\limits_{0}^{\infty} H(u, X(t))du}, \tag{6.4}$$

where

$$H(u,x) = exp[-\beta u + \sum_{i=1}^{n} \sigma_i x_i e^{-\alpha_i u} - \frac{1}{2}\sum_{i,j=1}^{n} \frac{\rho_{ij}\sigma_i\sigma_j}{\alpha_i + \alpha_j} e^{-(\alpha_i+\alpha_j)u}]. \tag{6.5}$$

Note that β represents the long term (infinite) rate in the above equation, while $X_i(t)$ is an Ornstein-Uhlenbeck process with $X_i(0) = \hat{x}_i$ and $dX_i(t) = \alpha_i(\mu_i - X_i(t))dt + d\hat{Y}_i(t)$

6.2 Parameterisation of the Yield Curve Model

Since we are planning to use this model to test the forecasting / hedging approaches under a wide range of circumstances, we require sufficient factors and a parameterisation which can generate a wide range of possible yield curves. Cairns (2004) showed that it is possible for a 2 factor parameterisation of the model to generate a large variety of curves. After testing a number of parameterisations we have decided to adopt a 3 factor parameterisation of the model:

$Parameter$	$Value$
β	5%
α_1	0.6
α_2	0.2
α_3	0.05
σ_1	0.7
σ_2	1.0
σ_3	0.6
μ_1	-1.5
μ_2	7
μ_3	-1.5

Tab. 6.1: Parameters for 3-factor Cairns model

We use the following correlation matrix:

	X_1	X_2	X_3
X_1	100%	30%	30%
X_2	30%	100%	30%
X_3	30%	30%	100%

Tab. 6.2: Correlations in 3-factor Cairns model

Choosing $X(0) = \mu$ gives a yield curve as shown in Figure 6.1. We have chosen these parameters to generate a shape that is broadly consistent to the South African swap curve over December 2007.

One further important note on this model relates to the assumption of risk premia. Assuming a non-zero μ vector effectively implies the existence of non-zero risk-premia in our model. We will therefore be simulating yield curves under the physical measure, not the risk-neutral measure.

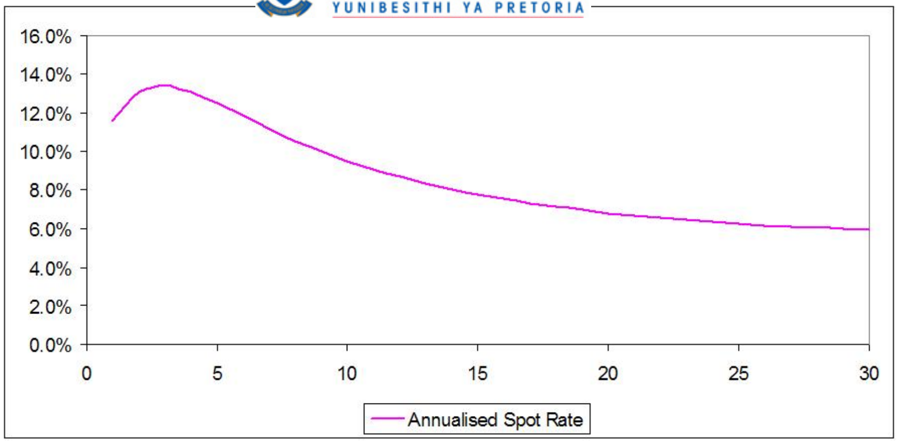


Fig. 6.1: Starting yield curve in Cairns 3 factor model

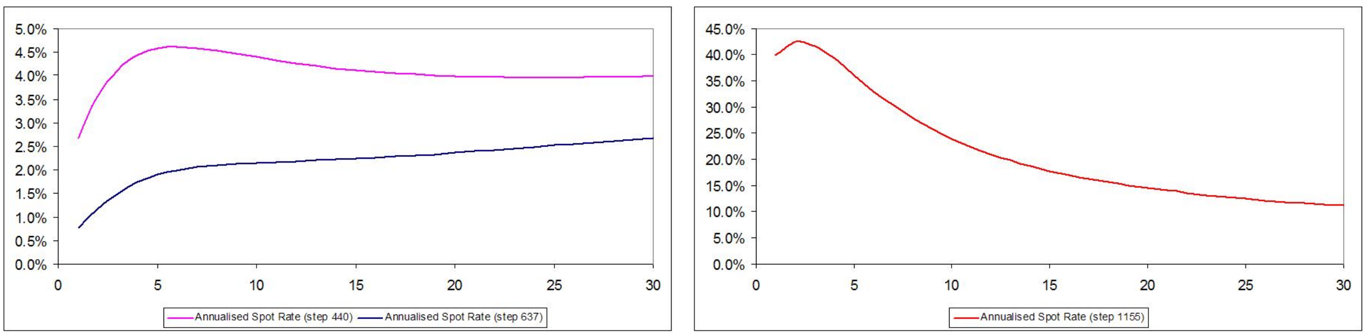


Fig. 6.2: CS2: Simulated Spot Curves using Cairns model and parameters defined above

6.3 Description of the Simulated Yield Curve Data

We use the above model to simulate a time series of 2000 weekly yield curve movements. Admittedly, the volatility parameters we have chosen could be regarded as high, however we are interested to see how the hedging approaches perform under extreme circumstances. Figure 6.2 shows some of the annualised spot curves generated in the time-series. We can see that there is a potentially wide range of curves that this model is able to generate.

6.4 Description of the Methodology

We use the same methodology as per Chapter 5 to create hypothetical hedging portfolios for a 50 year zero coupon bond. Surpluses which emerge are recorded and analysed.

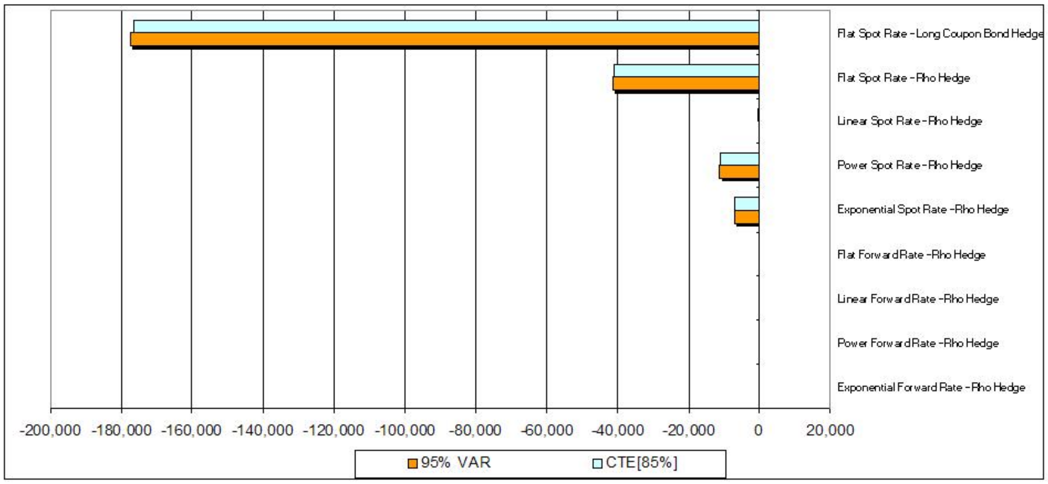


Fig. 6.3: CS2: VAR and CTE results for the simple extrapolation approaches

6.5 Simple Extrapolations: Hedging Results

We now provide the results of the hedging analysis for each of the simple extrapolation approaches described in Chapter 3. As in Chapter 5, we also provide the hedging errors that arise from an approach that is commonly used in practice.

We show the hedging errors that arise from using a Flat Spot Rate extrapolation and hedging with a long position in a coupon bearing bond only. (We have used a 30 year 6% coupon bearing bond for the purpose of illustration.) We will refer to this as the Benchmark Approach.

The results of the historical analysis (rounded to the nearest hundred rand) are as follows:

Statistic	*Benchmark*	*Flat Spot Rate*	*Lin. Spot Rate*	*Pwr. Spot Rate*	*Exp. Spot Rate*
95% VAR	(177 500)	(41 400)	(200)	(11 300)	(7 200)
CTE[85%]	(176 500)	(41 100)	(300)	(10 700)	(7 000)
Mean	(8 800)	(14 100)	2 200	(3 900)	(1 400)
Minimum	(855 300)	(219 200)	(8 200)	(16 300)	(12 100)
Maximum	274 700	400	13 500	7 000	30 100

Tab. 6.3: CS2: Results of simple spot rate extrapolations

Statistic	*Benchmark*	*Flat Fwd Rate*	*Lin. Fwd Rate*	*Pwr. Fwd Rate*	*Exp. Fwd Rate*
95% VAR	(177 500)	0	(100)	0	(300)
CTE[85%]	(176 500)	0	(0)	0	(200)
Mean	(8 800)	100	200	200	200
Minimum	(855 300)	(100)	(100)	(100)	(900)
Maximum	274 700	300	700	700	2 000

Tab. 6.4: CS2: Results of simple forward rate extrapolations

Figure 6.3 graphically illustrates the results for the various simple extrapolation procedures. Notice that the performance of the forward rate extrapolation approaches has dramatically improved. This raises an important question, one which is discussed in the coming sections.

6.6 Advanced Extrapolations: Hedging Results

As before, we follow the same methodology to build and analyse hedges for the advanced extrapolation approaches. For ease of comparison we again show the results obtained from the Benchmark Approach described above. Results are as follows:

Statistic	*Benchmark*	*Nelson − Siegel*	*Svensson*	*Cairns*	*Smith − Wilson*
95% VAR	(177 500)	(17 200)	(121 600)	(4 400)	(100)
CTE[85%]	(176 500)	(24 100)	(121 500)	(4 400)	(100)
Mean	(8 800)	4000	(28 500)	(1 800)	(0)
Minimum	(855 300)	(475 500)	(539 200)	(7 300)	(100)
Maximum	274 700	64 000	428 300	500	0

Tab. 6.5: CS2: Results of advanced extrapolations

Notice that the only two advanced approaches which seem to show a high level of performance are the Cairns approach and the Smith-Wilson approach. However, the Nelson-Siegel and Svensson approaches have maintained a similar level of performance to case study 1. This is a particularly interesting set of results, and they raise a number of important questions which we will discuss in the next section.

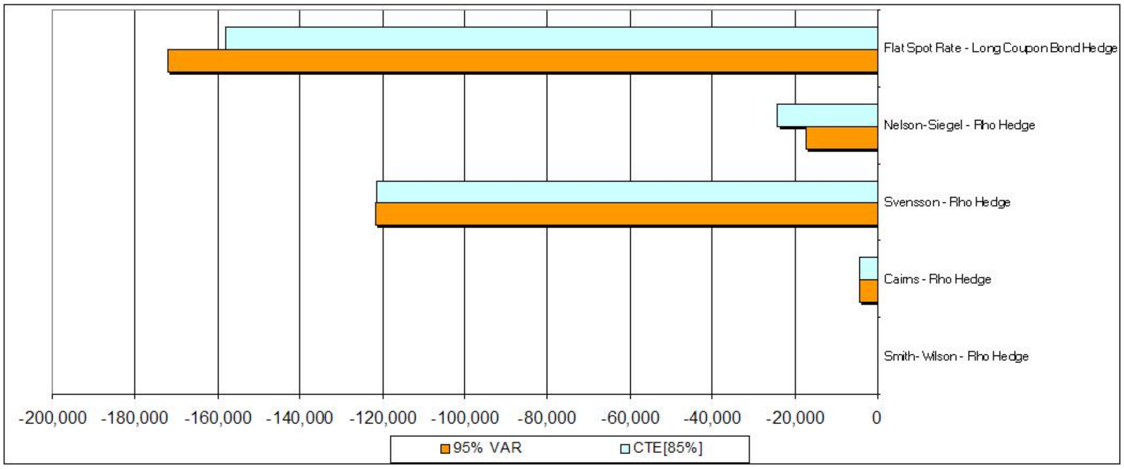


Fig. 6.4: CS2: VAR and CTE results for the advanced extrapolation approaches

6.7 Interpretation of Hedging Results

The results of the above analysis are remarkably different from those in case study 1. Key questions which arise are as follows:
• Why has the performance of the simple forward rate extrapolations improved so dramatically?
• How do we interpret the results for the advanced extrapolations?

We will deal with each of these questions separately.

6.7.1 Improved performance of simple forward rate extrapolations

Changes in the annual forward rate curve have been investigated. We have compared the volatility of movements in long forward rates between the above stochastic yield curve model and historical South African forward rates. The results are as follows:

$AnnualForwardRate$	$HistoricalStdDev$	$SimulatedStdDev$
21	0.94%	0.64%
22	1.06%	0.61%
23	1.20%	0.59%
24	1.24%	0.57%
25	1.08%	0.55%
26	0.75%	0.53%
27	0.71%	0.51%
28	1.01%	0.50%
29	1.31%	0.48%
30	1.49%	0.47%

Tab. 6.6: Comparison of historical vs simulated long forward rates

Under the Cairns model, we clearly see reducing volatility in forward rates as a function of outstanding term. This seems reasonable and fits perfectly with the above assumption that $\sigma_i(t,T) = \sigma_i e^{-\alpha_i(T-t)}$. However, historical data over the period in consideration seems to indicate a more erratic volatility profile in long term forward rates.

We can take this further and examine the correlation structure between changes in long term forward rates:

	21	22	23	24	25	26	27	28	29	30
21		99.99%	99.99%	99.97%	99.95%	99.92%	99.90%	99.86%	99.82%	99.77%
22			99.99%	99.99%	99.97%	99.95%	99.93%	99.90%	99.86%	99.82%
23				99.99%	99.99%	99.97%	99.95%	99.93%	99.90%	99.87%
24					99.99%	99.99%	99.97%	99.95%	99.93%	99.90%
25						99.99%	99.99%	99.97%	99.96%	99.94%
26							99.99%	99.99%	99.98%	99.96%
27								99.99%	99.99%	99.98%
28									99.99%	99.99%
29										99.99%
30										

Tab. 6.7: Cairns model - correlations between long term forward rates

	21	22	23	24	25	26	27	28	29	30
21		93.94%	84.20%	75.85%	67.74%	47.15%	-7.74%	-40.36%	-50.26%	-52.96%
22			97.50%	93.00%	86.33%	62.34%	-6.23%	-48.15%	-61.10%	-64.66%
23				98.71%	94.21%	71.23%	-0.78%	-46.72%	-61.25%	-65.23%
24					98.10%	78.89%	8.60%	-39.36%	-55.04%	-59.33%
25						88.91%	26.27%	-22.60%	-39.62%	-44.35%
26							67.44%	24.32%	6.59%	1.31%
27								88.01%	78.05%	74.49%
28									98.35%	97.12%
29										99.80%
30										

Tab. 6.8: Historical rates - correlations between long term forward rates

Based on these results, it seems that the nature of historical movements in long term forward rates is not fully captured by the yield curve simulator. In the South African context, it is quite possible that long term forward rates could display seemingly "spurious" changes as a result of supply / demand issues. Over the period of our historical investigation, the longest available government bond until mid-2006 was the R186 with a maturity of end-2027. Subsequently, the R209 was issued with a maturity of mid-2036. Therefore, the 25 - 30 year swap market, until mid-2006, has predominantly been an inter-bank market with relatively little liquidity. This could explain why we see a high amount of seemingly spurious volatility in long term forward rates over the period of historical investigation. This also explains, partly, why we see an improvement in the performance of the simple extrapolation approaches when moving into a more coherent environment as simulated using the Cairns model. However, in order to investigate these differences further we would need to use a simulator which allowed for more erratic movements in forward rates in the 20 - 30 year region of the yield curve. This will be a topic for further research.

6.7.2 Discussion of results for advanced extrapolations

As we have seen, the results for the advanced extrapolations are similar to those in case study 1. There has been notable increase in performance of the Cairns approach. We attribute this improvement to the greater flexibility of the approach as it is able to generate a variety of turning points and inflection points more easily.

Even in a simulated environment which is significantly different from that observed historically, we see that the Smith-Wilson approach again produces very small hedging errors consistent with our earlier results. This suggests a level of robustness within the approach.

6.8 Extension of Case Study 2: Hedging at monthly intervals

Thus far in our analyses we have only considered the case where rebalancing takes place at weekly intervals. In order to investigate sensitivity to the assumed hedging interval, we have extended the case study 2 investigation to the case where hedging takes place at monthly intervals. Results are included in Appendix E. As expected, the magnitude of hedging errors increases, however the ordering of results by approach is not significantly affected.

7. SUMMARY: ALTERNATIVE APPROACHES TO LONG TERM INTEREST RATE RISK

We began this dissertation with a statement of the following three problems facing entities with exposure to long term interest rates:

1. The inadequacy of traditional matching methods (i.e. immunisation and bucketing) to cope with the long term interest rate risks.
2. The non-observability of interest rate data beyond the maximum term in the yield curve. Associated with this is the inability to adequately quantify interest rate risk.
3. The lack of liquidity in long term interest rate markets. Associated with this is the inability to adequately hedge interest rate risk.

In order to address these problems, we have achieved the following:

We have examined various traditional methods used to hedge interest rate risks, and we have described why these approaches are not adequate for managing long term risks beyond the maximum tradable term (Chapter 2).

We have examined some modern methods to forecast and hedge interest rate risks and explored the possibility of their use in managing long term interest risk (Chapter 2).

On the back of these investigations, we proposed a number of possible yield curve extrapolation procedures and associated calibrations (Chapter 3). We then went further to derive generic theoretical hedging results relating to the proposed extrapolation procedures (Chapter 4).

Using the theoretical hedging results, we perform our first case study which involves deriving theoretical hedges over a historical period from October 2001 to March 2007. Weekly performance of the various extrapolation procedures is measured when used to forecast and hedge a 50 year zero coupon bond. An extension of the case study is performed by applying the same exercise to a 35 year zero coupon bond. The results indicate that extrapolation and hedging of the yield curve is able to significantly reduce Value-At-Risk of long term interest rate exposures. The Smith-Wilson approach gives the largest reduction in historical VAR, however it would also be the most exposive to implement as it requires consistent, large trades at relatively long terms (Chapter 5).

A second case study is then performed where weekly yield curve movements are simulated using a model proposed by Cairns (2004). The model produces yield curve behaviour which is inherently different from that in the observed historical period, however we find that our results and conclusions remain consistent (Chapter 6).

In conclusion, we find that there appears to be a significant benefit to the use of yield curve extrapolation techniques, particularly when used in conjunction with a hedging strategy. In some cases we find that the more simple extrapolation techniques actually increase risk (significantly) when used in conjunction with a hedging strategy. However, for some of the more advanced techniques, such as the functional form approaches and the Smith-Wilson approach, risk can be reduced significantly.

For an entity looking to deal with long term interest rate risk, we find that the choice of extrapolation technique and hedging strategy go hand-in-hand. For this reason the cost of hedging and reduction in risk are strongly correlated. The results obtained therefore suggest that it is necessary to weigh the benefits against the cost of hedging. However, this cost seems to increase with increasing reduction in risk. For example, we find that the functional form approaches are able to provide a moderate reduction in long term interest rate risk and do not require significant rebalancing over time. Conversely, the Smith-Wilson approach is (empirically) able to significantly reduce long term interest rate risk, although the extent of rebalancing is significantly increased.

APPENDIX A: PROOFS FOR SIMPLE HEDGING CASE

Proof of Result 4.0.0 By definition,

$$P_t(\tau) = \prod_{s=1}^{\tau} (1 + f_t(s))^{-1},$$

so for $k <= x$

$$\frac{\partial P_t(\tau)}{\partial f_t(k)} = -(\prod_{s=1}^{\tau} (1 + f_t(s))^{-1}) \times (1 + f_t(k))^{-1}.$$

Proof of Result 4.1.1C We know that

$$P_t(\tau) = f(f_t(1), f_t(2), f_t(3), ..., f_t(M), \beta).$$

Therefore

$$dP_t(\tau) = \sum_{s=1}^{M} \frac{\partial P_t(\tau)}{\partial f_t(s)} df_t(s) + \frac{\partial P_t(\tau)}{\partial \beta} d\beta.$$

So from (4.1.1B) it is obvious that result (4.1.1C) holds.

Proof of 4.1.2A By definition, we can write:

$$P_t(\tau) = \frac{P_t(\tau - 1)}{1 + f_t(\tau)}.$$

$$\Rightarrow P_t(\tau) \times (1 + a + b \times \tau) = P_t(\tau - 1).$$

$$\Rightarrow \frac{\partial P_t(\tau - 1)}{\partial a} = \frac{\partial P_t(\tau)}{\partial a} \times (1 + a + b \times \tau) + P_t(\tau).$$

so

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial a} = (\frac{\partial P_t(\tau - 1)}{\partial a} - P_t(\tau)) \times \frac{1}{1 + a + b \times \tau},$$

or

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial a} = (\frac{\partial P_t(\tau - 1)}{\partial a} - P_t(\tau)) \times \frac{P_t(\tau)}{P_t(\tau - 1)},$$

similarly

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial a} = \frac{\partial P_t(\tau - 2)}{\partial a} \times \frac{P_t(\tau)}{P_t(\tau - 2)} - P_t(\tau - 1) \times \frac{P_t(\tau)}{P_t(\tau - 2)} - P_t(\tau) \times \frac{P_t(\tau)}{P_t(\tau - 1)},$$

and by backwards induction we can show that, for $k < \tau - M$:

$$\frac{\partial P_t(\tau)}{\partial a} = \frac{\partial P_t(\tau - k)}{\partial a} \times \frac{P_t(\tau)}{P_t(\tau - k)} - P_t(\tau) \times \sum_{i=1}^{k} \frac{P_t(\tau - k + i)}{P_t(\tau - k + i - 1)},$$

however $P_t(M)$ is not dependent on a, so

$$\frac{\partial P_t(M)}{\partial a} = 0,$$

hence,

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial a} = -P_t(\tau) \times \sum_{i=1}^{\tau - M} \frac{P_t(M + i)}{P_t(M + i - 1)}.$$

Proof of 4.1.2B By definition, we can write:

$$P_t(\tau) = \frac{P_t(\tau - 1)}{1 + f_t(\tau)}.$$

$$\Rightarrow P_t(\tau) \times (1 + a + b \times \tau) = P_t(\tau - 1).$$

$$\Rightarrow \frac{\partial P_t(\tau - 1)}{\partial b} = \frac{\partial P_t(\tau)}{\partial b} \times (1 + a + b \times \tau) + P_t(\tau) \times \tau.$$

Hence

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial b} = (\frac{\partial P_t(\tau - 1)}{\partial b} - P_t(\tau) \times \tau) \times \frac{1}{1 + a + b \times \tau},$$

or

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial b} = (\frac{\partial P_t(\tau - 1)}{\partial b} - P_t(\tau) \times \tau) \times \frac{P_t(\tau)}{P_t(\tau - 1)},$$

similarly

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial b} = \frac{\partial P_t(\tau - 2)}{\partial b} \times \frac{P_t(\tau)}{P_t(\tau - 2)} - P_t(\tau - 1) \times (\tau - 1) \times \frac{P_t(\tau)}{P_t(\tau - 2)} - P_t(\tau) \times \tau \times \frac{P_t(\tau)}{P_t(\tau - 1)},$$

and by backwards induction we can show that, for $k < \tau - M$:

$$\frac{\partial P_t(\tau)}{\partial b} = \frac{\partial P_t(\tau - k)}{\partial b} \times \frac{P_t(\tau)}{P_t(\tau - k)} - P_t(\tau) \times \sum_{i=1}^{k} \frac{P_t(\tau + 1 - i)}{P_t(\tau - i)} \times (\tau + 1 - i).$$

However $P_t(M)$ is not dependent on b, so

$$\frac{\partial P_t(M)}{\partial b} = 0,$$

hence,

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial b} = -P_t(\tau) \times \sum_{i=1}^{\tau - M} \frac{P_t(M + i)}{P_t(M + i - 1)} \times (M + i).$$

Proof of 4.1.3A By definition, we can write:

$$P_t(\tau) = \frac{P_t(\tau-1)}{1+f_t(\tau)}.$$

$$\Rightarrow P_t(\tau) \times (1 + a \times b^\tau) = P_t(\tau-1).$$

$$\Rightarrow ln(P_t(\tau-1)) = ln(P_t(\tau)) + ln(1 + a \times b^\tau).$$

$$\Rightarrow \frac{1}{P_t(\tau-1)} \times \frac{\partial P_t(\tau-1)}{\partial a} = \frac{1}{P_t(\tau)} \times \frac{\partial P_t(\tau)}{\partial a} + \frac{b^\tau}{1 + a \times b^\tau}.$$

Hence

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial a} = P_t(\tau) \times (\frac{1}{P_t(\tau-1)} \times \frac{\partial P_t(\tau-1)}{\partial a} - \frac{b^\tau}{1 + a \times b^\tau}),$$

similarly

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial a} = P_t(\tau) \times (\frac{1}{P_t(\tau-2)} \times \frac{\partial P_t(\tau-2)}{\partial a} - \frac{b^\tau}{1 + a \times b^\tau} - \frac{b^{\tau-1}}{1 + a \times b^{\tau-1}}),$$

and by backwards induction we can show that, for $k < \tau - M$:

$$\frac{\partial P_t(\tau)}{\partial a} = P_t(\tau) \times (\frac{1}{P_t(\tau-k)} \times \frac{\partial P_t(\tau-k)}{\partial a} - \sum_{i=1}^{k} \frac{b^{\tau+1-i}}{1 + a \times b^{\tau+1-i}}).$$

However $P_t(M)$ is not dependent on a, so

$$\frac{\partial P_t(M)}{\partial a} = 0,$$

hence,

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial a} = -P_t(\tau) \times \sum_{i=1}^{\tau-M} \frac{b^{M+i}}{1 + a \times b^{M+i}}.$$

Proof of 4.1.3B By definition, we can write:

$$P_t(\tau) = \frac{P_t(\tau-1)}{1+f_t(\tau)}.$$

$$\Rightarrow P_t(\tau) \times (1 + a \times b^\tau) = P_t(\tau-1).$$

$$\Rightarrow ln(P_t(x-1)) = ln(P_t(\tau)) + ln(1 + a \times b^\tau).$$

$$\Rightarrow \frac{1}{P_t(\tau-1)} \times \frac{\partial P_t(\tau-1)}{\partial b} = \frac{1}{P_t(\tau)} \times \frac{\partial P_t(\tau)}{\partial b} + \frac{a \times b^{\tau-1} \times \tau}{1 + a \times b^\tau}.$$

Hence

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial b} = P_t(\tau) \times (\frac{1}{P_t(\tau-1)} \times \frac{\partial P_t(\tau-1)}{\partial a} - \frac{a \times b^{\tau-1} \times \tau}{1 + a \times b^\tau}),$$

similarly

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial b} = P_t(\tau) \times (\frac{1}{P_t(\tau-2)} \times \frac{\partial P_t(\tau-2)}{\partial a} - \frac{a \times b^{\tau-1} \times \tau}{1 + a \times b^{\tau}} - \frac{a \times b^{\tau-2} \times (\tau-1)}{1 + a \times b^{\tau-1}}),$$

and by backwards induction we can show that, for $k < \tau - M$:

$$\frac{\partial P_t(\tau)}{\partial b} = P_t(\tau) \times (\frac{1}{P_t(\tau-k)} \times \frac{\partial P_t(\tau-k)}{\partial b} - \sum_{i=1}^{k} \frac{a \times b^{\tau-i} \times (\tau+1-i)}{1 + a \times b^{\tau+1-i}}).$$

However $P_t(M)$ is not dependent on b, so

$$\frac{\partial P_t(M)}{\partial b} = 0,$$

hence,

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial b} = -P_t(\tau) \times \sum_{i=1}^{\tau-M} \frac{a \times b^{M+i-1} \times (M+i)}{1 + a \times b^{M+i}}.$$

Proof of 4.1.4A By definition, we can write:

$$P_t(\tau) = \frac{P_t(\tau-1)}{1 + f_x}.$$

$$\Rightarrow P_t(\tau) \times (1 + a \times \tau^b) = P_t(\tau-1).$$

$$\Rightarrow ln(P_t(\tau-1)) = ln(P_t(\tau)) + ln(1 + a \times \tau^b).$$

$$\Rightarrow \frac{1}{P_t(\tau-1)} \times \frac{\partial P_t(\tau-1)}{\partial a} = \frac{1}{P_t(\tau)} \times \frac{\partial P_t(\tau)}{\partial a} + \frac{\tau^b}{1 + a \times \tau^b}.$$

Hence

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial a} = P_t(\tau) \times (\frac{1}{P_t(\tau-1)} \times \frac{\partial P_t(\tau-1)}{\partial a} - \frac{\tau^b}{1 + a \times \tau^b}),$$

similarly

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial a} = P_t(\tau) \times (\frac{1}{P_t(\tau-2)} \times \frac{\partial P_t(\tau-2)}{\partial a} - \frac{\tau^b}{1 + a \times \tau^b} - \frac{(\tau-1)^b}{1 + a \times (\tau-1)^b}),$$

and by backwards induction we can show that, for $k < \tau - M$:

$$\frac{\partial P_t(\tau)}{\partial a} = P_t(\tau) \times (\frac{1}{P_t(\tau-k)} \times \frac{\partial P_t(\tau-k)}{\partial a} - \sum_{i=1}^{k} \frac{(\tau+1-i)^b}{1 + a \times (\tau+1-i)^b}).$$

However $P_t(M)$ is not dependent on a, so

$$\frac{\partial P_t(M)}{\partial a} = 0,$$

hence,

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial a} = -P_t(\tau) \times \sum_{i=1}^{\tau-M} \frac{(M+i)^b}{1 + a \times (M+i)^b}.$$

Proof of 4.1.4B By definition, we can write:

$$P_t(\tau) = \frac{P_t(\tau-1)}{1+f_t(\tau)}.$$

$$\Rightarrow P_t(\tau) \times (1 + a \times \tau^b) = P_t(\tau-1).$$

$$\Rightarrow ln(P_t(\tau-1)) = ln(P_t(\tau)) + ln(1 + a \times \tau^b).$$

$$\Rightarrow \frac{1}{P_t(\tau-1)} \times \frac{\partial P_t(\tau-1)}{\partial b} = \frac{1}{P_t(\tau)} \times \frac{\partial P_t(\tau)}{\partial b} + \frac{a \times \tau^b \times ln(\tau)}{1 + a \times \tau^b}.$$

Hence

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial b} = P_t(\tau) \times (\frac{1}{P_t(\tau-1)} \times \frac{\partial P_t(\tau-1)}{\partial b} - \frac{a \times \tau^b \times ln(\tau)}{1 + a \times \tau^b}),$$

similarly

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial b} = P_t(\tau) \times (\frac{1}{P_t(\tau-2)} \times \frac{\partial P_t(\tau-2)}{\partial b} - \frac{a \times \tau^b \times ln(\tau)}{1 + a \times \tau^b} - \frac{a \times (\tau-1)^b \times ln(\tau-1)}{1 + a \times (\tau-1)^b}),$$

and by backwards induction we can show that, for $k < \tau - M$:

$$\frac{\partial P_t(\tau)}{\partial b} = P_t(\tau) \times (\frac{1}{P_t(\tau-k)} \times \frac{\partial P_t(\tau-k)}{\partial b} - \sum_{i=1}^{k} \frac{a \times (M+1-i)^b \times ln(M+1-i)}{1 + a \times (M+1-i)^b}).$$

However $P_t(M)$ is not dependent on b, so

$$\frac{\partial P_t(M)}{\partial b} = 0,$$

hence,

$$\Rightarrow \frac{\partial P_t(\tau)}{\partial b} = -P_t(\tau) \times \sum_{i=1}^{\tau-M} \frac{a \times (M+i)^b \times ln(M+i)}{1 + a \times (M+i)^b}.$$

APPENDIX B: RESIDUAL HEDGING ERRORS - CASE STUDY 1

Figures 1 - 4 below provide the empirical frequencies of the hedging errors obtained from the historical hedging analysis for the advanced forecasting approaches.

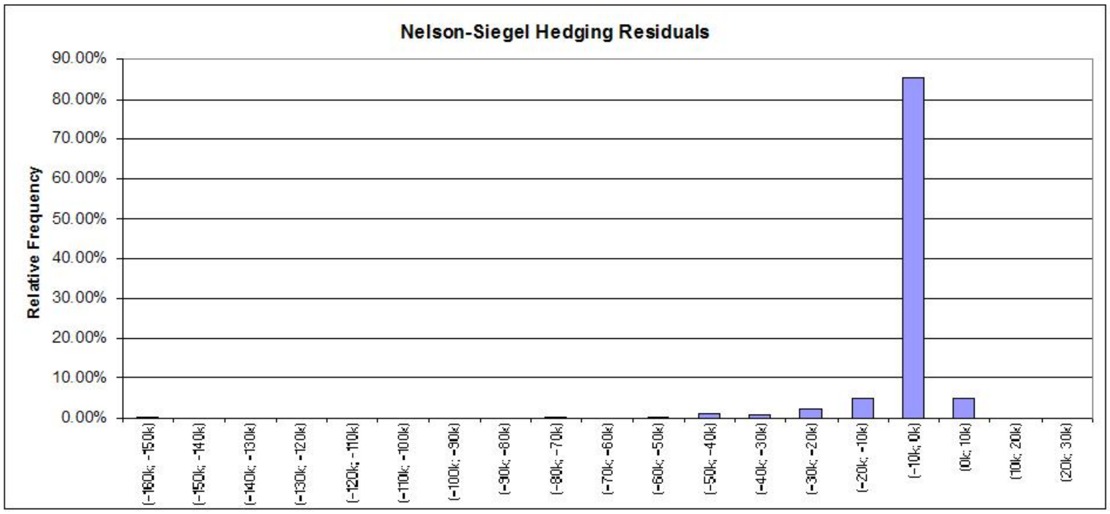


Fig. 10.1: CS1: Hedging Residuals for Nelson-Siegel Historical Analysis

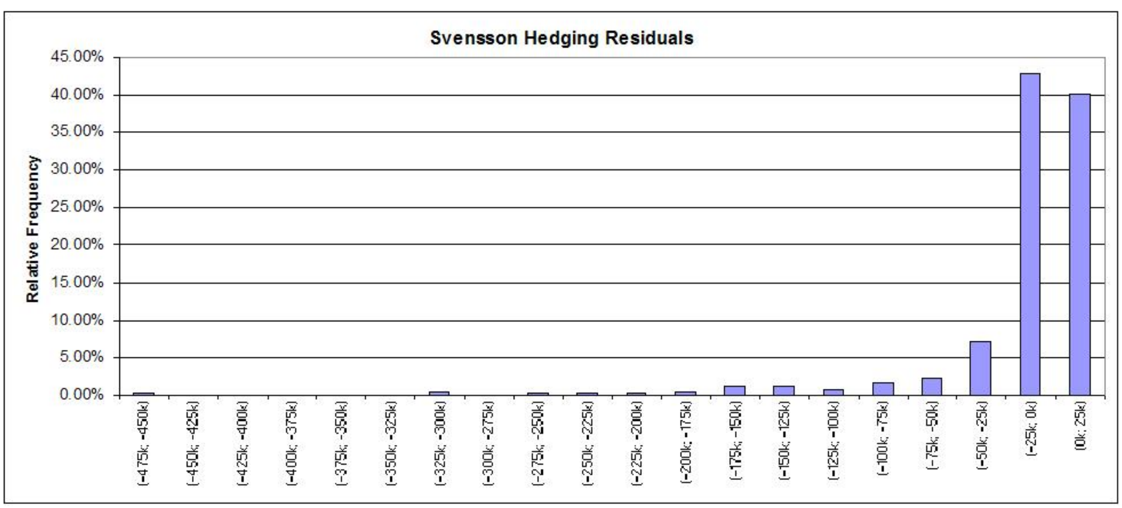


Fig. 10.2: CS1: Hedging Residuals for Svensson Historical Analysis

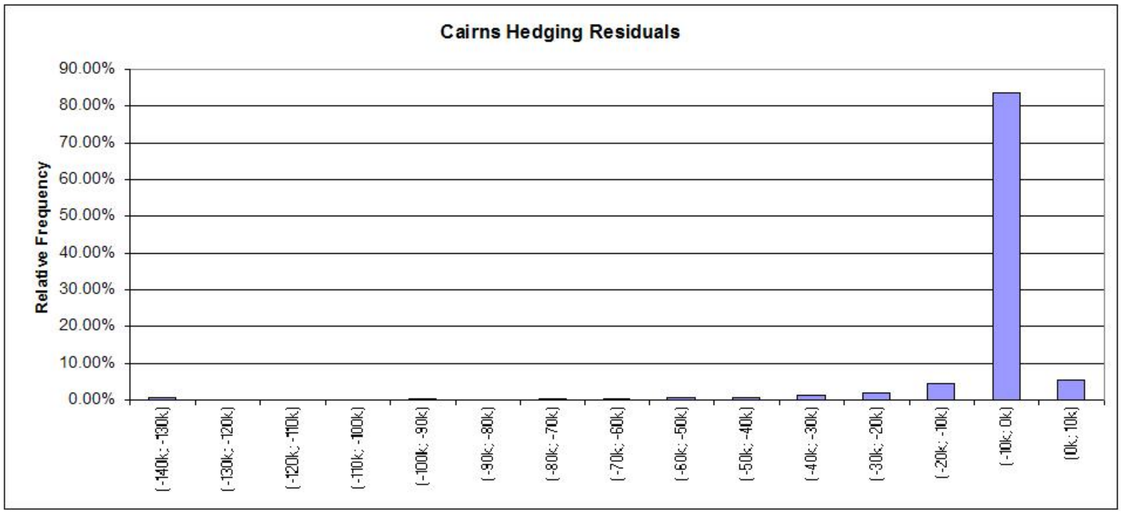


Fig. 10.3: CS1: Hedging Residuals for Cairns Historical Analysis

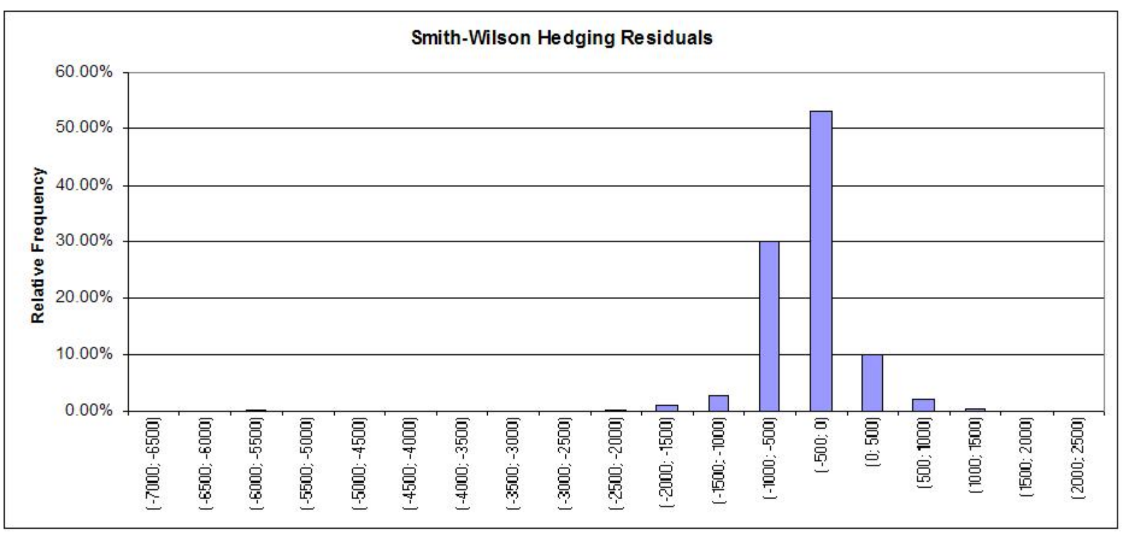


Fig. 10.4: CS1: Hedging Residuals for Smith-Wilson Historical Analysis

APPENDIX C: COMPOSITION OF HEDGING PORTFOLIOS - CASE STUDY 1

Figures 1 - 4 below show the composition of the hedging portfolios in Case Study 1 for various starting yield curves. Notice that the relative "shape" (composition) of the hedges is stable, although the absolute levels of hedging notionals do change.

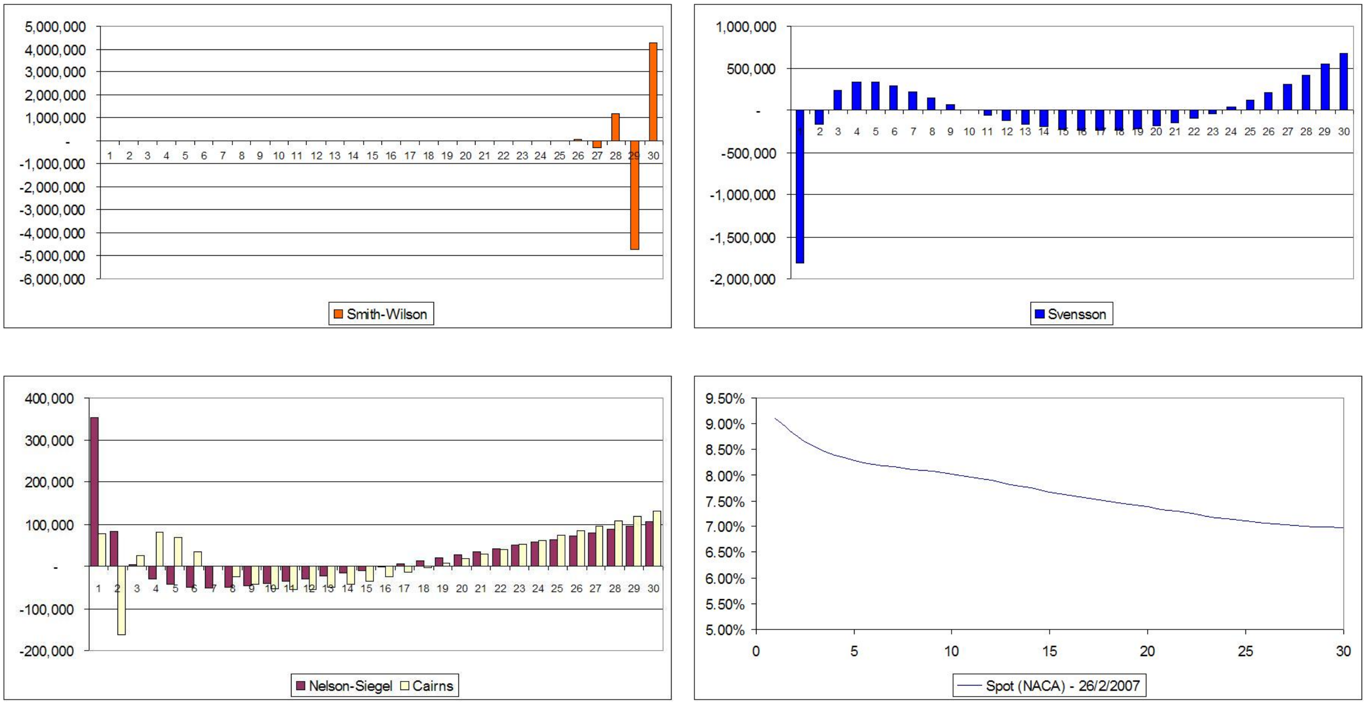


Fig. 11.1: *CS1: Composition of hedge portfolio for (standardised) R1m 50 year ZCB, effective 26/2/2007*

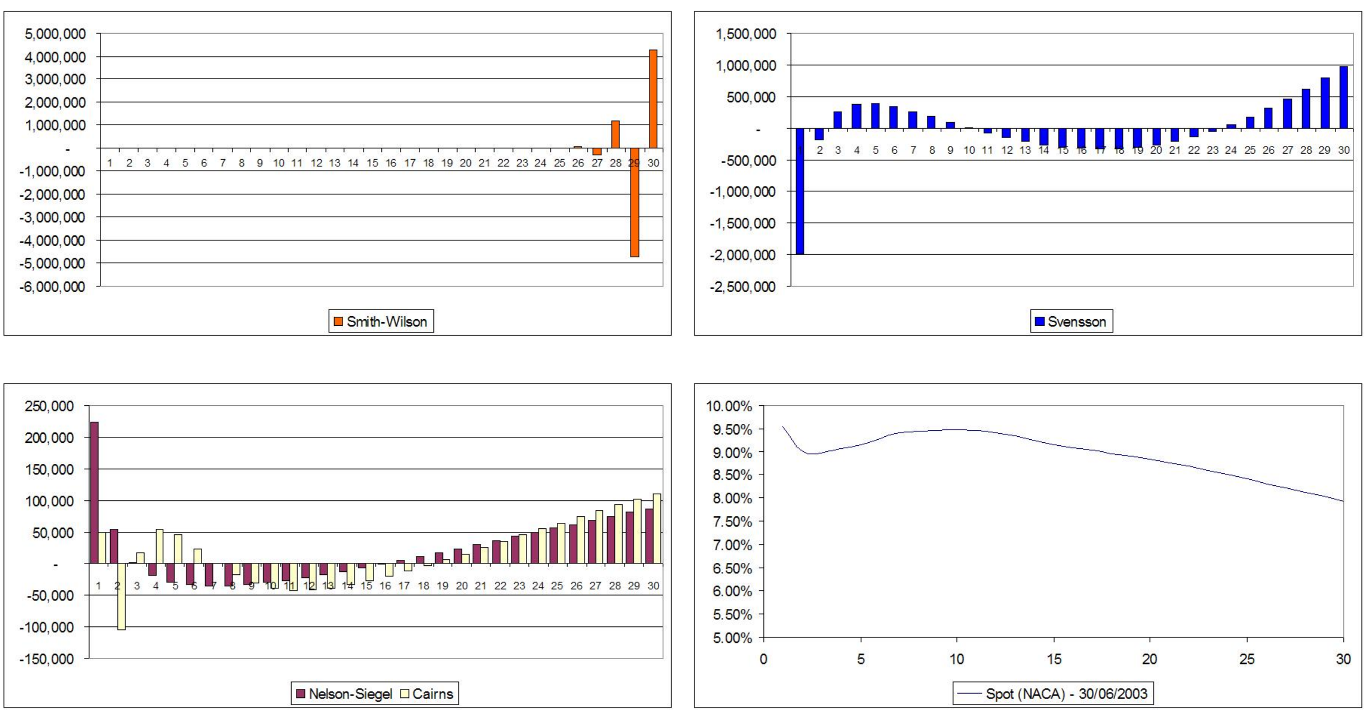


Fig. 11.2: *CS1: Composition of hedge portfolio for (standardised) R1m 50 year ZCB, effective 30/6/2003*

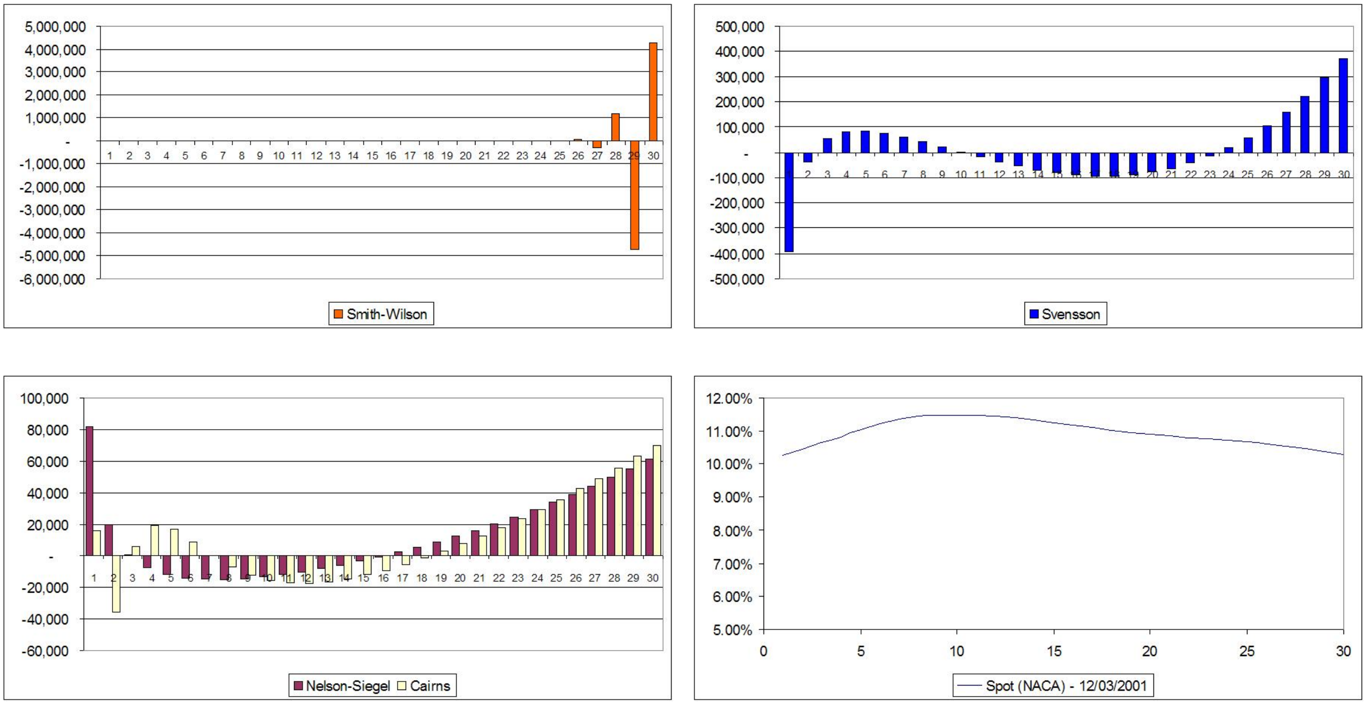


Fig. 11.3: CS1: Composition of hedge portfolio for (standardised) R1m 50 year ZCB, effective 12/3/2001

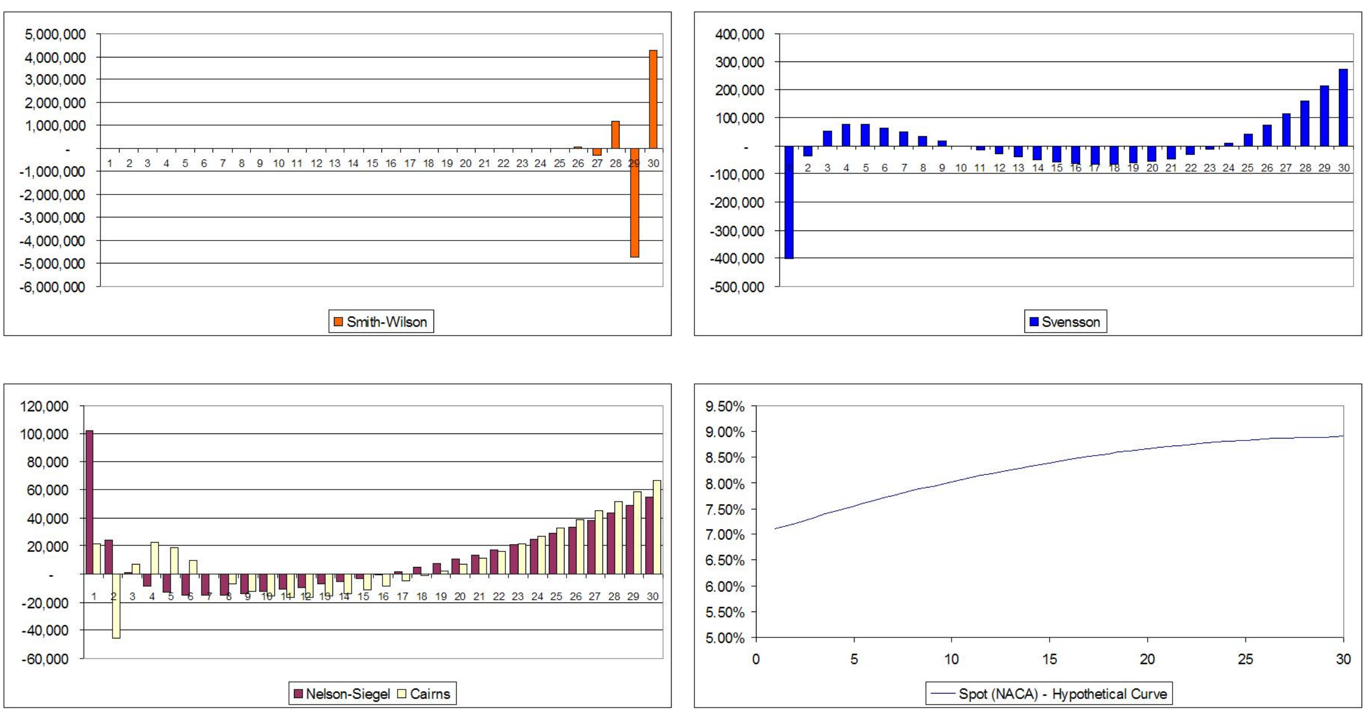


Fig. 11.4: CS1: Composition of hedge portfolio for (standardised) R1m 50 year ZCB, based on a hypothetical upward sloping curve

APPENDIX D: HEDGING A 35 YEAR BOND - EXTENSION OF CASE STUDY 1

In this appendix we present an extension of Case Study 1. Instead of hedging a 50 year hypothetical zero coupon bond, we now hedge a 35 year hypothetical zero coupon bond. Results of the simple forecasting approaches are presented separately from the advanced forecasting approaches.

Simple Extrapolations: Hedging Results

The results of the historical analysis for a hypothetical 35 year bond (rounded to the nearest hundred rand) are as follows:

Statistic	*Benchmark*	*Flat Spot Rate*	*Lin. Spot Rate*	*Pwr. Spot Rate*	*Exp. Spot Rate*
95% VAR	(108 200)	(4 400)	(8 400)	(5 200)	(4 900)
CTE[85%]	(98 100)	(4 600)	(8 500)	(5 200)	(5 300)
Mean	(5 300)	(2 700)	(2 900)	(2 500)	(2 100)
Minimum	(219 100)	(18 500)	(55 500)	(14 600)	(21 700)
Maximum	294 200	(1 300)	300	(600)	600

Tab. 12.1: CS1: Results of simple spot rate extrapolations (35 Year ZCB Hedge)

Statistic	*Benchmark*	*Flat Fwd Rate*	*Lin. Fwd Rate*	*Pwr. Fwd Rate*	*Exp. Fwd Rate*
95% VAR	(108 200)	(14 500)	(33 600)	(2 800)	(48 800)
CTE[85%]	(98 100)	(13 000)	(33 300)	(3 700)	(44 800)
Mean	(5 300)	(2 600)	(5 500)	200	(4 000)
Minimum	(219 100)	(45 300)	(236 600)	(19 700)	(225 000)
Maximum	294 200	(100)	86 000	32 500	114 200

Tab. 12.2: CS1: Results of simple forward rate extrapolations (35 Year ZCB Hedge)

Figure 1 graphically illustrates the results for the various simple extrapolation procedures.

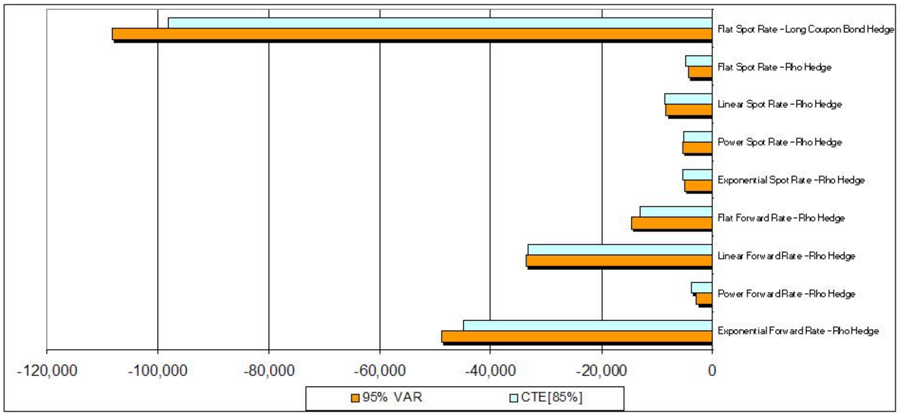


Fig. 12.1: CS1: VAR and CTE results for the simple extrapolation approaches (35 Year ZCB Hedge)

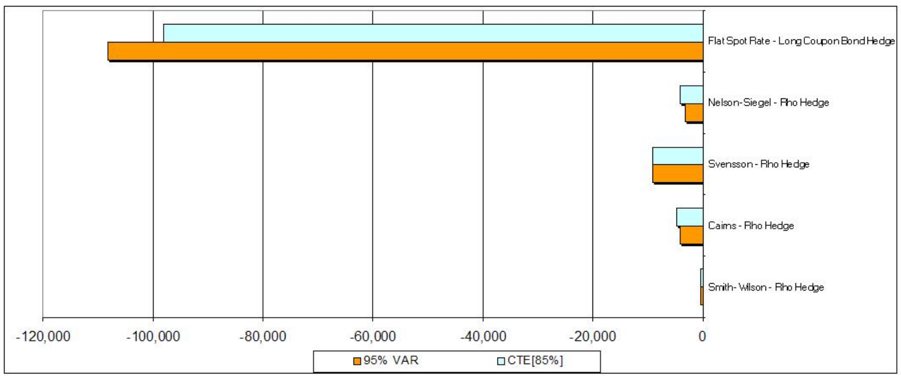


Fig. 12.2: CS1: VAR and CTE results for the advanced extrapolation approaches (35 Year ZCB Hedge)

Advanced Extrapolations: Hedging Results

Figure 2 graphically illustrates the results for the various advanced extrapolation procedures.

Statistic	*Benchmark*	*Nelson − Siegel*	*Svensson*	*Cairns*	*Smith − Wilson*
95% VAR	(108 200)	(3 300)	(9 100)	(4 200)	(400)
CTE[85%]	(98 100)	(4 200)	(9 100)	(4 900)	(400)
Mean	(5 300)	(1 200)	(1 100)	(1 200)	(200)
Minimum	(219 100)	(30 500)	(36 000)	(29 100)	(2 000)
Maximum	294 200	1 800	4 700	2 000	500

Tab. 12.3: CS1: Results of advanced extrapolations (35 Year ZCB Hedge)

APPENDIX E: HEDGING AT MONTHLY INTERVALS - EXTENSION OF CASE STUDY 2

We perform a similar analysis to that of Case Study 2. All elements of the analysis are identical to Case Study 2, except for the following changes:

- We simulate 2000 **monthly** yield curve movements from the given starting point.
- We calcualte hedging error and rebalance at the end of each simulated month.

Simple Extrapolations: Hedging Results

We provide the results of the hedging analysis for each of the simple extrapolation approaches described in Chapter 3. As in Chapter 5, we also provide the hedging errors that arise from an approach that is commonly used in practice.

We show the hedging errors that arise from using a Flat Spot Rate extrapolation and hedging with a long position in a coupon bearing bond only. (We have used a 30 year 6% coupon bearing bond for the purpose of illustration.) We will refer to this as the Benchmark Approach

The results of the historical analysis (rounded to the nearest hundred rand) are as follows:

Statistic	*Benchmark*	*Flat Spot Rate*	*Lin. Spot Rate*	*Pwr. Spot Rate*	*Exp. Spot Rate*
95% VAR	(447 100)	(102 800)	(3 000)	(10 100)	(5 900)
CTE[85%]	(439 400)	(119 400)	(4 300)	(9 600)	(5 600)
Mean	(40 900)	(29 000)	2 700	(1 200)	(4 700)
Minimum	(2 802 600)	(1 164 800)	(16 200)	(16 300)	(11 400)
Maximum	398 600	500	61 600	48 700	193 400

Tab. 13.1: CS2: Results of simple spot rate extrapolations (hedging at monthly intervals)

Figure 1 graphically illustrates the results for the various simple extrapolation procedures.

Advanced Extrapolations: Hedging Results

Results are given in Table 3 and Figure 2:

Statistic	*Benchmark*	*Flat Fwd Rate*	*Lin. Fwd Rate*	*Pwr. Fwd Rate*	*Exp. Fwd Rate*
95% VAR	(447 100)	(200)	(100)	0	(600)
CTE[85%]	(439 400)	(200)	(0)	0	(600)
Mean	(40 900)	0	200	300	400
Minimum	(2 802 600)	(1 100)	(100)	(100)	(2 200)
Maximum	398 600	400	3 200	3 200	7 900

Tab. 13.2: CS2: Results of simple forward rate extrapolations (hedging at monthly intervals)

Statistic	*Benchmark*	*Nelson – Siegel*	*Svensson*	*Cairns*	*Smith – Wilson*
95% VAR	(447 100)	(128 800)	(259 500)	(5 500)	(600)
CTE[85%]	(439 400)	(164 400)	(287 900)	(5 800)	(600)
Mean	(40 900)	(24 300)	(39 200)	(1 700)	(200)
Minimum	(2 802 600)	(1 857 400)	(3 136 600)	(35 300)	(900)
Maximum	398 600	59 900	1 901 800	4 300	300

Tab. 13.3: CS2: Results of advanced extrapolations (hedging at monthly intervals)

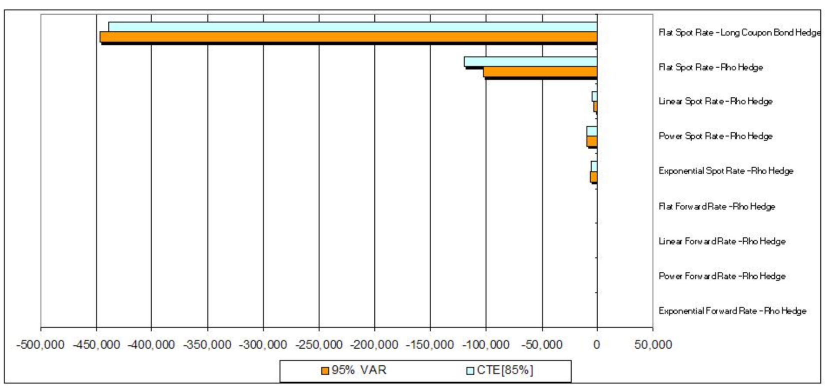


Fig. 13.1: CS2: VAR and CTE results for the simple extrapolation approaches (hedging monthly)

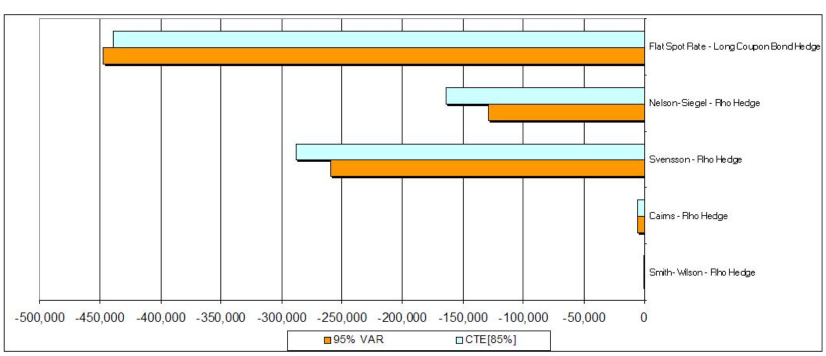


Fig. 13.2: CS2: VAR and CTE results for the advanced extrapolation approaches (hedging monthly)

REFERENCES

Agca, S. (2002): "The Performance of Alternative Interest Rate Risk Measures And Immunization Strategies Under A Heath-Jarrow-Morton Framework", Virginia Polytechnic Institute

Bierwag, G. (1977): "Immunization, Duration and the Term Structure of Interest Rates", The Journal of Financial and Quantitative Analysis, Vol. 12, No. 5, pp. 725 - 742

Bierwag, G. (1978): "Duration and Bond Portfolio Analysis: An Overview", The Journal of Financial and Quantitative Analysis, Vol. 13, No. 4, pp. 671 - 678

Bierwag, G., Fooladi, I. and Roberts, G. (1993): "Designing an immunized portfolio: Is M-squared the key?", Journal of Banking and Finance, Vol. 17: 1147-1170

Bierwag, G., Kaufman, G. and Toevs, A. (1983): "Immunization Strategies for Funding Multiple Liabilities", The Journal of Financial and Quantitative Analysis, Vol. 18, No. 1, pp. 113 - 123

Bjork, T. and Christensen, B. (1999): "Interest Rate Dynamics and Consistent Forward Rate Curves", Mathematical Finance, Vol. 9, No. 4, pp. 323 - 348

Bjork, T. (2004): "Arbitrage Theory in Continuous Time", Oxford University Press

Bliss, R. and Fama, E. (1987): "The Information in Long Maturity Forward Rates", An American Economic Review, Vol. 77, No. 4, pp. 680 - 692

Blocher, E. and Stickney, C. (1978): "Duration and Risk Assessments in Capital Budgeting", The Accounting Review

Brennan, M. and Schwartz, E. (1977): "Savings Bonds, Retractable Bonds and Callable Bonds", Journal of Financial Economics, Vol. 5, pp. 67 - 88

Boyle, P. P. (1980): "Immunization under Stochastic Models of the Term Structure", Volume 1, Actuarial Research Clearing House

Cairns, A. (1997): Descriptive Bond-Yield and Forward-Rate Models for the British Government Securities Market, Faculty and Institute of Actuaries

Cairns, A. (2004): A Family of Term Structure Models for long term Risk Management and Derivative Pricing, Mathematical Finance, Vol. 14, No. 3, pp. 415 444

Campolongo, F. and Rossi, A. (2002): "Sensitivity Analysis and the Delta Hegding Problem", Working Paper, IPSC

Carverhill, A. (2002): "Predictability and the Dynamics of Long Forward Rates", University of Hong Kong

Chambers, D. and Carleton, W. (1981): "A More General Duration Approach", Research In Finance

Chambers, D., Carleton, W. and McEnally, R. (1988): "Immunizing Default-Free Bond Portfolios with a Duration Vector", The Journal of Financial and Quantitative Analysis, Vol. 23, No. 1, pp. 89 - 104

Cooper, I. (1977): "Asset Values, Interest Rate Changes, and Duration", The Journal of Financial and Quantitative Analysis, Vol. 12, No. 5, pp. 701 - 723

Cox, J., Ingersoll, J. and Ross, S. (1979): "Duration and the Measurement of Basis Risk", Journal of Business, Vol. 52, No. 1, pp. 51 - 61

Cox, J., Ingersoll, J. and Ross, S. (1985): "A Theory of the Term Structure of Interest Rates", Econometrica, Vol. 53, No. 2, pp. 385 - 408

Chittim, D. (2006): "Long Duration Strategies for Pension Funds", Mellon Asset Management

Clewlow, L. and Strickland, C. (1998): "Implementing Derivatives Models", John Wiley and Sons

Dahlquist, M. and Svensson, L. (1996): "Estimating the Term Structure of Interest Rates for Monetary Policy Analysis", The Scandinavian Journal of Economics, Vol. 98, No. 2, pp. 163 - 183

de Pooter, M. (2007): "Examining the Nelson-Siegel Class of Term Structure Models", Discussion Paper, Erasmus University Rotterdam

Diebold, F., Ji, L. and Li, C. (2004): "A Three Factor Yield Curve Model: Non-Affine Structure, Systematic Risk Sources, and Generalised Duration", Working Paper, University Of California

Diebold, F. and Li, C. (2006): "Forecasting the Term Structure of Goverment Bond Yields", Journal of Econometrics, Vol. 130, pp. 337 - 364

Duffee, G. (2002): "Term Premia and Interest Rate Forecasts in Affine Models", Journal of Finance, Vol. 57, No. 1, pp. 405 - 443

Durand, D. (1974): "Payout Period, Time Spreads, and Duration: Aids to Judgement in Capital Budgeting", Journal of Bank Research

Dyvbig, P., Ingersoll J. and Ross, S. (1996): "Long Forward and Zero Coupon Rates Can Never Fall", Journal of Business, Vol. 69, No. 1

Elton, E., Gruber, M. and Michaely, R. (1990): "The Structure of Spot Rates and Immunization", Journal of Finance, June 1990

Filipović, D. (1999): "A Note on the Nelson Siegel Family", Mathematical Finance, Vol. 9, No. 4, pp. 349 - 359

Filipović, D. (2000): "Exponential-Polynomial Families and the Term Structure of Interest Rates", Bernoulli Journal, Vol. 6, No. 6, pp. 1081 - 1107

Fisher, L. and Weil, R. (1971): "Coping with the Risk of Interest Rate Fluctuations: Returns to Bondholders from Naive and Optimal Strategies", Journal of Business, pp. 408 - 431

Fishburn, P. (1977): "Mean-Risk Analysis with Risk Associated with Below-Target Returns", American Economic Review

Flesaker, B., and Hughston, L. (1996): Positive Interest, Risk 9(1), 46 - 49. Fong, H. and Vasicek, O. (1984): "A Risk Minimizing Strategy for Portfolio Immunization", Journal of Finance, Vol. 39, No. 5, pp. 1541 - 1546

Hagan, P. and West, G. (2006): "Interpolation Methods for Curve Construction", Applied Mathematical Finance, Vol. 13, No. 2, pp. 89 - 129

Haugh, M. (2005): "Deterministic Cash Flows and the Term Structure of Interest Rates"

Heath, D., Jarrow, R. and Morton, A. (1992): "Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claim Valuation" Econometrica, Vol. 60, pp. 77 - 105

Hicks, J. R. (1946): "Value and Capital", Oxford University Press

Ho, T. (1992): "Key Rate Duration Measures of Interest Rate Risk", Journal of Fixed Income (September), 29-44 Ingersoll, J., Skelton, J., and Weil R. (1978): "Duration Forty Years Later", Journal of Financial and Quantitative Analysis, Vol. 13, No. 4, pp. 627 - 650

Ingersoll, J. (1983): "Is immunisation feasible? Evidence from CRSP data", Innovations in Bond Portfolio Management, Greenwich, CT:JAI Press, p. 163-182

Keintz, R. and Stickney C. (1977): "Immunization of Pension Funds from Interest Rate Changes", Dartmouth College Working Paper

Khang, C. (1979): "Bond Immunization When Short-Term Rates Fluctuate More Than Long Term Rates", Journal of Financial and Quantitative Analysis

Krippner L. (2006): "A Theoretically Consistent Version of the Nelson and Siegel Class of Yield Curve Models", Applied Mathematical Finance, Vol. 13, No. 1, pp. 39 - 59

Longstaff, F. and Schwartz, E. (1992): "Interest Rate Volatility and the Term Structure: A Two-Factor General Equilibriam Model", The Journal of Finance, Vol. 97, No. 4, pp. 1259 - 1282

Lustig, H., Sleet, C. and Yeltekin, S. (2006): "Fiscal Hedging and the Yield Curve", Working Paper, UCLA

Maitland, J. (2002): "Interpolating the South African Yield Curve", Department of Statistics and Actuarial Science, University of Witwatersrand

Macaulay, F. R. (1938): "Some Theoretical Problems Suggested by the Movements of Interest Rates, Bond Yields, and Stock Prices in the U.S. since 1856", National Bureau of Economic Research

Nelson, C. and Siegel, A. (1987): "Parsimonious modelling of yield curves", Journal of Business, October, pp. 473 - 489

Norberg, R. (2004): "Credibility Theory", London School of Economics, London, United Kingdom

Niffikeer, C., Hewins, R. and Flavell, R. (2000): "A synthetic factor approach to the estimation of value-at-risk of a portfolio of interest rate swaps", Journal of Banking and Finance, Vol. 24, Issue 12, December 2000, pp. 1903-1932

Pelsser, A. (2003): "Pricing and hedging guaranteed annuity options via static option replication", Insurance: Mathematics and Economics, Vol. 33, Issue 2, pp. 283-296

Phoa, W. (2000): "Yield Curve Risk Factors: Domestic and Global Contexts", The Practitioner's Handbook of Financial Risk Management

Poitras, G. (2005): "Immunization Bounds, Time Value and Non-Parallel Yield Curve Shifts", Working Paper, Simon Fraser University

Rebonato, R. (1998): "Interest-Rate Option Models", Second Edition, John Wiley and Sons

Reddington, F. (1952): "Review of the Principles of Life Office Valuations", Journal of the Institute of Actuaries, Vol. 78, pp. 286 - 340

Santa-Clara, P. and Sornette, D. (2001): "The Dynamics of the Forward Interest Rate Curve with Stochastic String Shocks", The Review of Financial Studies, Vol. 14, No. 1, pp. 149 - 185

Smith, A. and Wilson, T. (2001): "Fitting Yield Curves with Long Term Constraints", Research Notes, Bacon and Woodrow

Svensson, L. (1994): "Estimating and interpreting forward interest rates", Centre for Economic Policy Research

Thomas, M. and Maré, E. (2007): "Long Term Forecasting and Hedging of the South African Yield Curve", 2007 ASSA Convention, The Actuarial Society of South Africa

Tito, D. and Wagner, W. (1977): "Definitive New Measures of Bond Performance and Risk", Wilshire Associates

Vasicek, O. (1977): "An Equilibrium Characterisation of the Term Structure", Journal of Financial Economics, Vol. 5

Wilkie, A., Waters, H. and Yang, S. (2003): "Reserving, pricing and hedging for policies with guaranteed annuity options", British Actuarial Journal, Vol. 9, No. 2, 2003 , pp. 263-391