

Market Valuation of Pension Liabilities

MARK GREENWOOD

8807484/F

Dissertation for the Degree of Master of Science

Department of Statistics and Actuarial Science

Faculty of Science

University of the Witwatersrand

Supervisor: Prof. R.J. Thomson

July 2008

I declare that this dissertation is my own, unaided work. It is being submitted for the Degree of Master of Science to the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination to any other University.



(Signature of candidate)

28 JUNE 2008

(Date)

Acknowledgements

For Simona, who has shown absolute faith in me, inspired me and supported me through our adventures of the past five years and for my parents, Vic and Merle, who have always been an limitless source of help and encouragement to their sons.

I wish to thank my supervisor, Professor Rob Thomson, for his insights and suggestions and for allowing me to draw on his experience and expertise in pensions and investment.

I also wish to thank my former colleagues at Rand Merchant Bank: Gill Raine for introducing me to inflation as an asset class in 1999 and helping me to understand its significance, Gill Baker who I learned much from when we worked together, Wayne Dennehy and Alan Chown for their thoughts on the SA inflation market, Steve Potter and Alan Joffe for data and advice on the South African interest rate volatility market. I am very grateful to Simona Svoboda for her insights into interest rate modelling and guidance through the related literature.

I found the PCA and Cholesky spreadsheet models of Graeme West of Financial Modelling Agency to be useful checks for my calculations. I also acknowledge Sergei Kucherenko of BRODA for the use of their high-dimension Sobol code and for advice on the application of Sobol sequences for generating quasi-random numbers for the purposes of testing the analytic formulae in the dissertation.

The comments and suggestions of the external examiners is gratefully acknowledged.

Mark Greenwood,
June 2008

Abstract

This dissertation analyses the market-consistent valuation of liabilities for defined-benefit pensions in payment. Models from the actuarial and financial economics literature for valuation of inflation options embedded in typical liabilities in the South African market are considered. The feasibility of the assumptions underpinning these models is then appraised and it is concluded that, while existing models may produce reasonable market valuations for pension liabilities, these models are unable to capture important aspects of the dynamics of the interest rate and inflation markets. The market for pension liabilities is incomplete due to background risks such as mortality, credit, regulatory and tax risks. Stochastic mortality models are considered and it is described how an incomplete-market valuation of liabilities using risk-adjusted pricing principles may be produced by extending an investment model to include mortality.

Contents

List of Figures	iii
List of Tables	v
Chapter 1. Objective and overview of research	1
1. Research objective	1
2. Introduction and overview	1
Chapter 2. Literature Survey	5
1. Actuarial literature	5
2. Financial economics literature	8
Chapter 3. Market valuations and pension liabilities	9
1. Market values and actuarial practice	10
2. Market values and financial economics theory	11
3. Financial and insurance instruments	14
4. Liability discount rate for market valuation	16
Chapter 4. Linear inflation liabilities	19
1. Notation	20
2. Breakeven inflation and real yield valuation	21
3. Inflation markets	23
4. Market disequilibrium	25
5. Curve modelling	26
6. The bond-swap inflation basis	31
7. Inflation seasonality	33
8. Inflation forecasts	38
9. Inflation risk premium	39
10. Curve extension	40
11. Hedge risk for linear liabilities	47
12. CPI revisions	58
13. Market completeness	59
Chapter 5. Non-Linear inflation liabilities	60

1. Forms of pension indexation	61
2. Inflation derivatives and pension indexation	62
3. The Black–Scholes world	63
4. Derivatives pricing theory	64
5. Black–Scholes currency model for inflation	65
6. Condition 1: Perfect Markets	71
7. Condition 2: Inflation driven by Brownian motion	80
8. Inflation targeting and mean reversion	82
9. Condition 3: Deterministic interest rates	90
10. Condition 4: Deterministic volatility	94
11. Financial market models for inflation derivatives	94
Chapter 6. Background risks	100
1. Background risks for a DB pension fund	101
2. Valuation of background risks	101
3. Longevity risk	102
4. Credit risk	115
5. Regulatory and tax effects	116
6. Conclusion	117
Appendix A. Specimen real annuity	119
Bibliography	121

List of Figures

1	SA nominal bond swap spreads	17
2	Illustration of variable indexation lag for monthly pension payments with annual increases	21
3	Composition of SA bond market	23
4	Real cashflows for 2008, 2013, 2023 and 2033 SA inflation-linked bonds	24
5	SA inflation-linked bond real yields	25
6	SA zero-coupon nominal bond-curves	28
7	SA nominal and inflation-linked bond-swap spreads assuming no inflation-linked bond-swap basis	33
8	Underlying SA CPIx trend excluding seasonality produced using X-11 preliminary stage and linear regression	35
9	Empirical seasonality profiles	36
10	Estimated monthly seasonality factors and confidence intervals for SA CPI	37
11	Hedging risk from seasonal variation in inflation	38
12	Stylised spot and forward curves showing various extrapolation methods	42
13	SA nominal, breakeven and implied 4-month lag real bond curves, 26 June 2006	46
14	Example of immunisation failing	52
15	Correlation of SA nominal bond curve 8-year rate with other maturity rates	53
16	SA nominal bond forward-rate correlation matrix for monthly absolute changes in rates	54
17	Factor loadings for first three principal components for the PCA using SA nominal bond forward-rates in Table 3	55

18	Factor loadings for first three principal components for SA nominal bond rate PCA using spot rates of Table 4	56
19	Factor loadings for first three principal components for the PCA using SA breakeven spot rates of Table 5	56
20	Indexed pension payments for the different forms of indexation	62
21	Published and unknown inflation at valuation date.	67
22	Lag between the date of the payout and indexation date	67
23	Monthly and daily hedge ratios for a 3-year type-2 inflation cap	74
24	Distribution of hedge error for Black–Scholes continuous hedge and minimum-variance discrete hedge	76
25	Distribution of hedge error for type-2 cap for various rehedging intervals	77
26	Histogram of month-on-month SA headline CPI	81
27	Q-Q and P-P plot for the absolute change in the SA headline inflation rates	82
28	SA CPI and CPIX since introduction of inflation targeting	83
29	Annual absolute change for nominal, breakeven and CPI inflation rates	85
30	Correlogram for monthly change in CPI index, ΔC_t	87
31	Histogram of daily absolute changes in SA 2013 bond breakeven	92
32	Q-Q and P-P plot of daily absolute changes in the SA 2013 bond breakeven rates	92
33	SA 2013 bond breakeven normal volatility versus breakeven inflation rate	93
34	SA 2015 nominal and 2013 breakeven rates normal volatilities	95
35	Real cashflows for specimen 60-year male real annuity at 26 June 2006, discounted with best-estimate mortality.	119
36	Market rates for SA bond and swap market, 26 June 2006	120

List of Tables

1	Pricing discrepancy between BEASSA perfect-fit and published nominal zero-coupon bond curves	29
2	Estimated SA CPI seasonality factors	36
3	First seven eigenvalues for the PCA of the correlation matrix of SA nominal bond forward-rates depicted in Figure 16	54
4	Factor loadings for first three principal components of the correlation matrix of spot rates for SA nominal bonds, June 2004 to June 2006 monthly rate moves	55
5	Factor loadings for first three principal components of the correlation matrix of spot rate changes for the SA bond breakeven curve, June 2004 to June 2006 monthly rate moves	56
6	Calculation of pension increases for different forms of indexation	62
7	10-year atm type-2 inflation cap cumulative hedge error for various δt	78
8	Annual absolute change for nominal, breakeven and CPI inflation rates	84
9	Serial-correlations for the absolute change for nominal, breakeven and CPI inflation in rates, ΔL_t , ΔY_t and ΔC_t , shown in Table 8	85
10	Significance p -values for correlations in Table 9	86
11	Serial-correlations in Table 10 corrected for bias	87

CHAPTER 1

Objective and overview of research

1. Research objective

The object of this dissertation is to determine market-consistent valuation of inflation-linked pensions in payment in the context of the South African market. The development of the inflation-linked bond and swap market provides pension funds with a risk-free means to match substantially such pension liabilities, while the law of one price dictates that the market value of the liabilities is the market value of these hedge assets. In practice pension liabilities present issues relating to the completeness of the market and, even in an idealised complete market, the complexity of the pension promise in South Africa often necessitates a valuation based on a dynamic hedging strategy. This research examines both financial-economics theory developed for inflation and interest-rate derivatives pricing and actuarial theory for the market valuation of insurance liabilities. We draw on relevant international experience, particularly that of the advanced UK market.

2. Introduction and overview

March 2000 saw the birth of a new asset class in the South African investment industry when National Treasury issued the first RSA inflation-linked bond. Following the experience of established markets, such as the UK, long-term institutional investors have gained appreciation for the asset-liability matching benefits of inflation-linked bonds. Derivatives, such as inflation-linked swaps, have emerged in response to the specific requirements of investors such as pension funds.

This is a timely development. The introduction into a market such as South Africa of inflation-linked bonds and derivatives with (virtually) risk-free inflation-linked returns provides pension funds with the opportunity to more precisely match assets to pensions liabilities. A debate is raging within the actuarial profession regarding the optimal investment strategy and valuation methodology for defined benefit (DB) pension funds. This debate has

been sparked and fuelled by ideas from financial economics. These ideas have had a profound impact on the risk management of banks and their corporate clients and hold out similar prospects for long-term institutions such as pension funds. The emergence of a risk-free inflation-linked asset class allows the actuarial profession to adopt these ideas in the valuation of inflation-linked pension liabilities.

At the heart of the debate is the most effective means of managing the investment risks of a pension fund. Risk mitigation through diversification of investments and smoothing of investment returns over time, what might be called the ‘insurance approach’, had until recent years been fairly successful in cost-effective provision of pension benefits. Equity markets have in recent years yielded returns below valuation assumptions. This has however cast doubt over the insurance approach and the assumptions needed in practice to apply it. The insurance approach underpins most actuarial valuation work.

The alternative ‘financial approach’, as exemplified by investment banks, manages risk through hedging transactions where risk is neutralised by transferring it to a natural counterpart. Risk transfer increasingly takes place via the capital markets in order to lower costs and provide increased liquidity.

This distinction between the actuarial and financial approaches to risk management carries through into valuation. Actuarial valuation often seeks to explain asset prices and risk premiums “in an absolute manner, in terms of the so-called fundamentals” (Jensen & Nielsen (1995)). A funding valuation for pension assets and liabilities is a prominent example. By contrast financial economics seeks to explain asset prices and risk premiums in a relative manner “in terms of other, given and observable prices” (Embrechts (1996)).

Over time, the markets for pension risks have become more complete and efficient through financial innovations such as derivatives and securitisation. There is now a greater universe of reliable market prices for these risks and actuarial valuations may become more consistent with financial economics pricing. This trend was heralded by Buhlmann (1987) in his famous ASTIN editorial urging the actuarial profession to become “actuaries of the third kind”, proficient in financial economics.

In the discussion of Smith (1996), the author asserts that “greater integration (between short-term derivative pricing models and long-term actuarial models) would be of considerable benefit to those offering and receiving guarantees”. However, he goes on to caution that the valuation results of long-term actuarial forecasts are very sensitive to subjective assumptions and motivates the use of arbitrage arguments as a sound basis for development of market-consistent prices.

The market valuation of pension liabilities has attracted increased attention in response to accounting developments and the shift to liability-driven investment strategies. Chapter 2 surveys the actuarial literature addressing market valuations for pensions and other long-dated guarantees of insurance nature and the financial economic literature relating to inflation- and interest-rate derivatives pricing.

Chapter 3 examines the notion of a market value in more detail. The key attributes of a market valuation of a pension liability are identified and the market conditions required to develop such a valuation are introduced. The nature of SA DB pensions in payment, where increases often have an explicit link to the CPI, is considered.

Chapter 4 considers the market valuation of linear pension liabilities, which we define as liabilities perfectly indexed with inflation and not subject to credit, longevity and other background risks. The valuation of these liabilities requires methods for interest-rate and inflation curve building and extrapolation which account for seasonality and the inflation indexation lag. Forward starting liabilities, where indexation commences at a future date, introduce further complications.

A model for interest rates and inflation may allow more precise hedging of linear liabilities and is essential for the dynamic hedging of liabilities where the inflation link is nonlinear (such as increases based on the inflation rate floored at 0%) as is frequently the case. Chapter 5 discusses the models proposed in the financial economics and actuarial literature. The conditions underpinning the perfect market of Black & Scholes (1973) are appraised.

Chapter 6 addresses the impact on the valuation of the background risks implicit in pension liabilities, primarily longevity and credit risk. Stochastic

mortality models and the relationship between longevity and investment risk are examined. It is considered how the market-price information from insurance and reinsurance prices and mortality securitisation and derivatives will inform the market price of risk factors and the significance of the mortality projection is noted. The dissertation concludes with a discussion of the challenges of extending a model for interest-rates and inflation to include longevity.

CHAPTER 2

Literature Survey

This chapter surveys literature relevant to the research objective, the market valuation of inflation-linked pensions in payment. There has been considerable debate in the actuarial literature regarding what constitutes a market valuation for pension liabilities. Much of the literature addresses the nature and purposes of the valuation rather than the methods that can be used to produce market values.

The financial economics literature on inflation-linked derivatives is surveyed, as inflation-linked pensions in payment may be viewed as inflation-linked derivatives subject to additional background risks such as mortality.

1. Actuarial literature

1.1. Market valuation of pension liabilities

1.1.1. Exley, Mehta & Smith (1997) is considered a landmark as it was the first to address comprehensively a theory of market valuation of pension liabilities. Since then Gordon (1999), Chapman, Gordon & Speed (2001) and Wise (2004) have elaborated on investment aspects of the theory.

1.1.2. Head et al. (2000) considered the different forms of pension fund valuation that might be termed market valuations, including valuations based on financial-economics principles.

1.2. Actuarial models for inflation-linked pension liabilities

1.2.1. Cairns (2004a) and Thomson & Gott (2007) have produced equilibrium models that are market consistent. Palin & Speed (2003) and Chapman et al. (2001) produce equilibrium models using the deflator technique advocated by Jarvis, Southall & Varnell (2001). The equilibrium model of Thomson (2005) uses mean-variance optimisation to produce market values for inflation-linked pension liabilities in incomplete markets.

1.2.2. A real and nominal interest-rate model is required to price pension liabilities that contain inflation derivatives, such as annual inflation-linked increases floored at 0%. Considerable research has been directed at interest-rate modelling in recent years. The primary application has been to short-term

interest-rate derivatives and the models have been criticised for having unrealistic long-term dynamics. Real and nominal yield models for long-term liabilities have been developed by Van Bezooyen, Exley & Smith (1997), Smith (1999) and Huang & Cairns (2004).

1.2.3. The dissertation undertakes a critical appraisal of the Black & Scholes (1973) methodology to consider whether it can be used to value the inflation-derivative features of pension liabilities.

1.3. Dynamic hedging of pension liabilities

1.3.1. Embrechts (1996), Cairns (2001), Moller (2002), Wang (2002) and Thomson (2005) survey the financial economics theories developed to value insurance liabilities in incomplete markets. Cairns notes the parallel between the theory and actuarial development of risk minimising reserves (e.g. Wilkie (1984)) to define market values.

1.3.2. Van Bezooyen et al. (1997) and Smith (1999) argue against reserve-based approaches in favour of arbitrage-free dynamic pricing in a perfect market. Exley (2006) uses a similar approach to address the issue of forward-indexed liabilities. Huang & Cairns (2004) build a dynamic hedging model in discrete time based on mean-variance optimisation and a VAR(1) representation of a more complex asset-liability model (ALM) such as the Wilkie model. These papers all address specifically Limited Price Indexation U.K. pension liabilities, where the pension is a non-linear function of the CPI.

1.3.3. Cairns (2001) produces a market valuation model for the price of a new security (such as an actuarial liability) when introduced into a perfect asset market. He shows the price he derives is a good approximation to the equilibrium price.

1.4. Background risks to pension liabilities

1.4.1. Background risks for pensions in payment, such as longevity, credit and regulatory risks, cannot be perfectly hedged. Pension liabilities form an incomplete market since there is no unique market value consistent with the market prices for the risks inherent in the liabilities.

1.4.2. Longevity risk is generally regarded as the most significant background risk for pensions in payment. Cairns, Blake & Dowd (2006) provide a comprehensive survey of mortality models and techniques for valuing liabilities subject to mortality risk.

1.4.3. Utility theory can be used to identify a unique market price under certain conditions for a given utility function. Cardinale et al. (2006) use a simple model to explore how the market value of pension liabilities changes

with background risks of different forms when representative agents have Cobb-Douglas utility functions.

1.4.4. Market values can be also derived without explicitly drawing on utility theory. Lin & Cox (2005) use the Wang transform to estimate the market price of mortality risk from the US life annuity market. The market price of mortality risk can then be included within a risk-adjusted valuation of pension liabilities. Alternatively, the risk-neutral efficiency measure of Smith (2001*a*) can be used to define a unique incomplete-market value for a liability using a maximum-efficiency criterion.

1.5. Actuarial asset–liability models

1.5.1. Thomson (2002) notes that an asset–liability model can be used to value risks that cannot be hedged provided it can be extended to include them. Thomson (2005) uses this approach to establish a fair value for pension liabilities in an incomplete market that is in equilibrium.

1.5.2. Traditional actuarial ALMs, such as Wilkie (1986) and Wilkie (1995), do not give market valuations as these models are not intended to be market consistent. Thomson (1996) published an ALM available for the South African market using the same ARIMA framework used for the first Wilkie model.

1.5.3. Smith (1996) considers the full spectrum of models, from the deterministic chaotic model of Clark (1992) to the random-walk model of Kemp (1997), that constitute the state of the art of actuarial ALM at the time. The author also introduces a jump equilibrium model with a mixture of independent gamma increments and Poisson jumps.

1.5.4. Lee & Wilkie (2000) in their comparison of ALMs also include Yakoubov, Teeger & Duval (1999) and the market consistent ALMs of Smith (1996) and Cairns (2004*a*) in their list of public ALMs that include both U.K. nominal and inflation-linked bonds among the asset classes modelled.

1.5.5. Whitten & Thomas (1999) build threshold autoregressive (TAR) features into a model similar to the Wilkie AR model. Such TAR models allow for switching between different regimes of autoreversion and can account for many of the features of the inflation process not captured by a simple AR model (such as the Wilkie AR model). De Gooijer & Vidiella-i-ANGEURA (2003) demonstrate promising fits of self-exciting threshold autoregressive models to several inflation series.

1.5.6. A number of proprietary ALMs such as Smith (2003) exist but the details published are limited to marketing material.

2. Financial economics literature

2.1. Valuation of pension liabilities

2.1.1. The irrelevance of asset allocation for DB pensions was first considered by the prominent financial economists Treynor (1977) and Black (1980). This allows a financial economic valuation to focus on the liabilities and second-order asset effects (such as pension tax benefits).

2.1.2. Albrecht (1991) considers the assumptions underlying the financial economics valuation methodology and questions whether it can be used to value liabilities with insurance characteristics. Reitano (1997) and Babbel, Gold & Merrill (2002) counter many of these concerns in their comparison of the actuarial and financial approaches for life annuity and other insurance liabilities.

2.2. Inflation-derivatives models

2.2.1. The inflation derivatives market was the precursor for the development of models for real rates and inflation in mathematical finance. Jarrow & Yildirim (2003) published the first model using the currency analogy for inflation and the HJM multicurrency framework of Amin & Jarrow (1991).

2.2.2. Belgrade & Benhamou (2004*b*), Belgrade, Benhamou & Khelif (2004) and Mercurio (2005*a*) subsequently published inflation-market models which are more easily calibrated to traded inflation derivatives. Mercurio & Moreni (2006) extended the latter model to account for volatility smile in the inflation caps and floors market. Lioui & Poncet (2005) derive an equilibrium model for the prices of inflation derivatives.

CHAPTER 3

Market valuations and pension liabilities

“A cynic is a man who knows the price of everything but the value of nothing.”

Oscar Wilde

We define the market value of liabilities used in this research as the price at which the liability would trade in an active market. This market valuation differs from actuarial valuations conventionally used for setting the pace of funding, but may be used for this purpose. A market valuation also has merit for accounting and commercial transaction purposes as its subjectivity is constrained by market prices of relevant instruments.

The market value defined in this research is motivated as the relevant value to use for trading liabilities in a nascent market for pensions in payment linked to the Consumer Price Index (CPI) and mortality. The circumstances in which such pension rights may come to be traded are discussed. The financial economics theory for the price of these rights is outlined.

1. Market values and actuarial practice

1.0.1. In recent years there has been increased interest in valuations of pension liabilities that are consistent with the market values of related assets and liabilities. Discretion regarding DB pension benefits has been eroded by regulation and reinforcement of member expectations. There is a growing acceptance that corporate pension liabilities should be treated in a consistent manner with the corporate's debt.

1.0.2. This research will define the market value of a pension liability as the estimated value at which the liability would trade in an active market. The definition of market value used in this research is similar to Reitano (1997) and Cairns (2001) in their research on the fair value of insurance liabilities.

1.0.3. Head et al. (2000) consider four main purposes for actuarial valuations of pension liabilities. Funding valuations are used to set the sponsor contribution rate for prefunding of liabilities. Regulatory valuations are usually used to assess pension fund solvency. Accounting valuations are used to publish pension liabilities and expenses in the accounts of the sponsor. Valuations are also required commercial transactions, generally where there is a bulk or individual transfer of a pension liability for cash (or equivalent assets) or there is an adjustment to benefits.

1.0.4. The market value as defined above will not necessarily correspond to an actuarial valuation of pension liabilities produced for these purposes. Fair-value accounting valuations are similar in principle to a market value, but may ignore the credit risk of the liability. The international accounting standard for pensions, IAS19, and the UK accounting standard, FRS17, use a general AA corporate-bond discount rate to value liabilities, while this research will motivate the use of the swap curve with suitable adjustments. These adjustments relate to the basis between the government bond market and the swap market and the credit risk to the fund from the sponsor in respect of unfunded liabilities in respect of pensions in payment.

1.0.5. Commercial transactions resulting in bulk transfers of pension liabilities may take place close to market value, although the transfer will in general not be limited to pensions in payment and there may be other considerations relating to the broader transaction. Smith (1996) notes that an anomaly often becomes apparent between other actuarial values of the pension liability and the market value that crystallises following a commercial transaction or imposition of a fair-value accounting standard.

1.0.6. The seminal paper of Exley et al. (1997) advocated the use of market values for all pension valuations. The authors used the Modigliani & Miller

(1958) corporate finance theory to argue that the fund assets are irrelevant, to first order, for the market value of liabilities. Second-order effects, such as credit and tax, will usually have only a limited effect on the valuation.

1.0.7. This research considers the valuation of DB pensions in payment. Section 4.1 argues that SA pensions in payment are sufficiently well defined in real terms to be treated as a stream of payments linked to the CPI and subject to background risks (such as mortality and credit risk) for which at least a partial hedge exists. This is not necessarily the case for pension liabilities in respect of DB actives still accruing benefits or deferreds due to receive a pension from a future retirement date. These liabilities are complicated by the difference between earnings and CPI inflation, decrements (withdrawal, ill-health early retirement, etc) and options at retirement (retirement age and options to commute benefits for cash). Reference to the pension liabilities in this research is taken to mean the liabilities in respect of pensions in payment unless otherwise specified.

2. Market values and financial economics theory

2.0.1. This section introduces the basic concepts in financial economics theory relating to market values. Financial economics theory usually assumes perfect markets, which are frictionless in that positions of unlimited size can be bought or sold instantaneously without incurring transaction costs, crossing bid-offer spreads or market impact. Perfect markets for pension liabilities, which are a fiction but nonetheless a useful concept for developing theory, are appraised in chapter 5.

2.0.2. Arbitrage is defined as the possibility of riskless profit. The law of one price (LOOP) states that if a tradable liability (or asset) can be decomposed into a set of tradable cashflows for which there is a market price then the market price of the liability is the market price of such cashflows. Any other market price ascribed to the liability allows an arbitrage by selling the more expensive portfolio and buying the cheaper.

2.0.3. Arbitrage-free pricing principles invoke the LOOP to price liabilities with reference to the market price of a replicating portfolio of tradable assets or liabilities. Under certain conditions perfect replication is possible and the resulting market price is preference independent (beyond the basic preference for exploiting an arbitrage, i.e. reducing the cost of defeasing the pension liability without risk). The price is independent of individual investor preferences because arbitrage will enforce a market price that is independent of the utility functions of market agents.

2.0.4. A complete market is defined as a market where any asset or liability is replicable. In chapter 4 it is considered whether linear pension liabilities free of background risks constitute a complete market for market valuation purposes. Since a pension liability is then replicable, the LOOP can be applied to determine the market value of the liability.

2.0.5. The fundamental theorem of finance states that the absence of arbitrage in a complete market implies the existence of a unique risk-neutral measure. This risk-neutral probability measure is equivalent to the original (often termed real-world) probability measure (Harrison & Pliska 1981). Two measures are equivalent if and only if they have the same null set of events. The fundamental theorem can also be formulated in terms of the existence of a unique deflator process used in some of the actuarial literature.

2.0.6. Later chapters address the elements of pension liabilities that cannot be perfectly replicated and so render the market incomplete. This leads to the existence of many alternative risk-neutral measures and their corresponding deflator processes.

2.0.7. A further condition that could be required of this arbitrage-free incomplete-market price is that it is consistent with an efficient market. The original definition of a (weak form) efficient market is a market where prices “fully reflect” all historical price information (Fama 1970). Historical price information cannot be the basis for trading strategies that earn excessive profits in relation to the risk of the trade.

2.0.8. The no-arbitrage condition precludes riskless profits. The market-efficiency condition is more general than the no-arbitrage condition, since it precludes excessive profits in relation to the risk of the trade. Cochrane (2001), quoting from Grossman & Stiglitz (1980), points out that efficient markets are a paradox in that efficient markets need traders and investment managers to exploit inefficiencies to keep them efficient. Market values derived from efficient-market models may nevertheless be considered to produce more robust and realistic market values than alternative models which do not use a market-efficiency criterion.

2.0.9. The economic condition of a market being in equilibrium requires that the total supply of the asset or liability must equal total demand at the market price in an efficient market (Duffie 2001). Equilibrium is a stronger market condition than efficiency. Equilibrium may be evident in a liquid market for a commoditised financial instrument (e.g. government bonds). This may not be the case in highly regulated, illiquid insurance markets.

2.0.10. The utility function of an assumed representative agent may be used to assess whether the risk-versus-return performance of trades is consistent with an efficient market. An alternative measure of market efficiency may be based on the difference between the real-world and risk-neutral distributions (Smith 2001*a*).

2.0.11. In an incomplete market, utility theory can be used to specify the deflator process from the marginal utility of a representative agent and an assumed real-world probability distribution of cashflows and hence specify a market price (Jarvis et al. 2001).

2.0.12. Utility theory can account for why risk-averse insureds accept an insurance premium greater than the (assumed real-world) expected value of claims. However, a unique incomplete market value can be identified using alternative approaches not based explicitly on utility theory. We consider such approaches in relation to incomplete markets for the longevity and investment risk of annuitants in chapter 6.

2.0.13. The market price of risk for each source of risk in the incomplete market can be used to price the instrument using a risk-adjusted pricing measure. This is a natural extension to risk-neutral pricing theory in complete markets.

2.0.14. The value of the instrument may be based on the cost of the self-financing dynamic hedging strategy that is optimal in some sense, such as the mean-variance hedging strategy of Duffie & Richardson (1991). There has been extensive debate in the literature about whether this cost is the market value at which the instrument would trade if a market existed (Cairns 2001).

2.0.15. Some financial economics theories for incomplete markets coincide with actuarial pricing principles. Standard premium-loading principles for insurance instruments are consistent with prices implied by specific utility functions. The Esscher transform frequently arises as the preferred risk-neutral measure for insurance instruments priced within the standard financial economics framework (Embrechts 1996). The Esscher market price P is based on an exponential tilt of the real-world payoff X ,

$$P = \mathbb{E}[Xe^{\delta X}] / \mathbb{E}[e^{\delta X}]$$

Wang (2002) proposed an alternative transform which is consistent with risk-neutral pricing in the idealised Black–Scholes world.

2.0.16. In both the financial and actuarial approaches to pricing there is no substitute for the accurate modelling of the liability cashflows. The financial

approach may not rely on some aspects of the cashflow process (such as the real-world drift of the Black–Scholes model) and other parameters (such as implied volatilities) may be directly calibrated to market but the resulting market value will be relevant only if the cashflows have been modelled accurately. Consequently, much of this dissertation focusses on models for investment and background risks underlying pensions in payment.

3. Financial and insurance instruments

3.1. Tradability of pension rights

3.1.1. Wilkie in the discussion of Head et al. (2000) argues that pension rights are “not like the assets usually dealt with in financial economics” since a third party does not wish to own a pension that depends on a member’s salary, starts when the member retires and ends when the member dies. This may be true for the liabilities in respect of active and deferred members, but pensions in payment are frequently insured with life annuities and in some cases (e.g. small funds, bulk annuities), the pension liability is transferred to the insurer. The credit risk faced by the pensioner is transferred from the fund to the insurer.

3.1.2. If pension increases are well-defined by law, pension fund rules or reasonable expectations and there is no prospect of benefit improvements, then the transfer of pension rights to a third party with a sound credit rating will be satisfactory to the members, their trustees and the regulators. Bulk transfers of pension liabilities make issues such as moral hazard and antiselection less problematic for the parties concerned.

3.2. Insurance securitisation

3.2.1. The securitisation of life annuity business, where blocks of liabilities are financed by debt securities, creates securities with payoffs similar to pension liabilities. The market value of pension liabilities may be based on the market value of life insurance securities with similar characteristics. The securitisation market may grow to the point where PWC conceive investors in securitised life portfolios “could become the arbiters of the fair valuation”¹. Insurance accounts for only a fraction of the assets and liabilities securitised to date. Cowley & Cummins (2005) divide life-insurance securitisations into five groups:

- (1) Blocks of insurance policies. Securitisation is used to release embedded value and capital or disinvest from a business. These deals have

¹ ‘Innovative financing - life insurance securitisation’, PWC Insurance Services, 2006.

been directed at linked and protection life business, although equity-release and reverse mortgage business for retirees has also been securitised.

- (2) Financing new business strain and conservative statutory reserves. These deals finance persistency and mortality risk for term-assurance products.
- (3) Conventional asset securitisations by insurers. Mortgage-backed loans by insurers are one such example.
- (4) Transfer of adverse mortality and longevity risk to the capital markets. The payments on life catastrophe and longevity bonds can be considered as mortality derivatives.
- (5) Life settlements and viaticals. These are life insurance policies sold by the insured to the secondary market for an amount greater than the surrender value but less than the face value.

The implications of life securitisation for the hedging and pricing of pension liabilities are discussed in the final chapter on mortality background risk.

3.3. Price and value

3.3.1. Investors account for and transact holdings at market price rather than perceived fundamental value. This research equates the market value for the pension liabilities with the estimated price in a hypothetical active market. It is suggested that this definition of market value is relevant for accounting or commercial transactions because the market price must be anticipated in the nascent market for pension liabilities. Failure to do so will result in losses or gains when the liabilities are transferred or realised as payments to pensioners.

3.3.2. Reitano (1997) draws an analogy between the incomplete markets for insurance liabilities and private debt placements of new debt issues. In the latter, the financial markets price the primary debt placement with reference to the price of instruments with similar credit and interest rate risk that trade actively in the secondary market. In time, an active secondary market develops for the new debt and the true market price manifests. This valuation process is to an extent self-fulfilling because valuations of untraded assets influence perceptions of reasonable market prices.

3.3.3. The market value envisaged by this research would correspond to the average of the bid and offer prices for an actively traded asset. This value is not biased towards the buyer or the seller and is used as the mid-market price for transactions in replicating assets or liabilities when a frictionless market is assumed.

3.3.4. We note that this implicitly assumes the market is in equilibrium, as does any valuation that is equated with a market price. The spread between bid and offer price for a financial asset will reflect the liquidity of the market. The bid–offer spread for illiquid life securitisation paper will be relatively wide while the asset class of life insurance securitisations is still immature, so this market will only provide weak information about the market value of pension liabilities.

3.3.5. Bid and offer prices for life securitisations will also apply in limited volumes. The market impact of transactions in larger volumes over a short period makes it likely that larger transactions will be executed outside this bid–offer spread. This is explored in more detail in the next chapter where it is argued that market impact is relevant only if the intention is to transfer the pension liabilities to a third party.

3.3.6. It is noted that the individual and bulk annuity market also conveys information regarding the market price of the risks pertaining to the pension liabilities. However, these are offer levels that need to be adjusted for margins for profit and capital support in order to infer mid prices. This topic is addressed in the final chapter.

3.3.7. A pension liability valuation based on market prices brings the discipline of the market to bear on the valuation. Subjectivity is constrained through the requirement that the model recovers market prices where possible. However, considerable actuarial judgment is still required in the construction of the valuation model.

4. Liability discount rate for market valuation

4.0.1. Head et al. (2000) consider the slightly broader concept of a market-consistent pension valuation that is “consistent with a feasible range of market prices, if a true market were actually to exist”. This acknowledges that pension liabilities are subject to longevity and credit risks for which the market is still incomplete. Financial economics theory admits a range of market values free from arbitrage in an incomplete market, even before transaction costs and bid–offer spreads for hedging instruments are factored in.

4.0.2. While Exley, Mehta & Smith (1997) demonstrated asset irrelevance to first order for liability valuation, credit risk is a second-order effect that depends on the assets held by the fund. Chapter 6 analyses the sponsor credit spread in respect of the unfunded pension liabilities that sponsor and trustees acting as fiduciaries for pensioners may agree is fair.

4.0.3. Chapter 6 motivates the use of the nominal and inflation swap curves, adjusted for the average government bond-swap spread, for discounting pension liabilities free of credit risk. Chapters 4 and 5 address the inflation projection and chapter 6 addresses the mortality projection.

4.0.4. The inflation and nom-

inal swap curves are used to generate the shape of the discount curve because it is swaps that complete the market for interest and inflation rate risks. In many markets swaps are more transparent and liquid

<i>RSA bond</i>	<i>R194</i>	<i>R153</i>	<i>R157</i>	<i>R186</i>
26 Apr 2006	-0.17%	-0.23%	-0.14%	-0.16%
26 May 2006	-0.21%	-0.29%	-0.17%	-0.16%
26 Jun 2006	-0.32%	-0.64%	-0.48%	-0.42%

Figure 1: SA nominal bond swap spreads (continuously compounded rates)

than government bonds and are less susceptible to disruptions in supply. The government bond-swap spread may be unstable over time. This is illustrated by Figure 1 which shows SA spreads increased in the months leading up to 26 June 2006. The effect of a relative supply squeeze on R153 2010 and R157 2015 bonds is also apparent as the swap spreads on these issues increased in comparison with the other issues. In markets with active bond futures contracts there may be irregularities in the bond curve at the maturity points of the cheapest-to-deliver bonds. The dynamics of swap spreads are discussed in the curve modelling section in the next chapter.

4.0.5. The government bond-swap spread to be applied to the swap curve at maturity T to produce a risk-free curve discount factor at T is the spread to the swap curve that equates the net present value (NPV) of a liability projected for mortality and inflation to T and discounted off the government bond curve with the NPV using the swap curve plus this spread. The sponsor credit spread is applied to this risk-free rate, only in respect of unfunded liabilities, to give:

$$\begin{aligned}
 & \text{discount curve value for market value of liabilities at } T \\
 &= \text{swap curve at } T \\
 &\quad - \text{government bond-swap spread at } T \\
 &\quad + \text{sponsor credit spread at } T, \text{ in respect of unfunded liabilities}
 \end{aligned}$$

4.0.6. If the liability is subsequently traded, the fair-value credit spread to apply to the risk-free discount curve is determined using the principles set out in chapter 6 using the credit risk of the entity paying the pension and any collateral. The entity assuming the liability (such as an insurer) may

be regulated with capital requirements and possible recourse to an industry protection fund but no collateral posted against the liabilities.

4.1. Pension liabilities

4.1.1. The liability to the fund in respect of DB pensions in payment is the value of the pensions accrued at retirement and increased with past pension increases plus allowance for future pension increases in accordance with the fund's pension increase policy. In the case of a spouse's or dependents' pension, the liability is the value of this pension, including accrued pension increases and allowance for future pension increases.

4.1.2. Pension increases are awarded in terms of the fund rules or, in the absence of specific rules, in terms of the practice established by the fund trustees and the reasonable expectations of pensioners. These reasonable expectations form from factors such as the trustees' increase policy and investment returns (Institute and Faculty of Actuaries 1998).

4.1.3. In some jurisdictions the minimum pension increase is stipulated by pensions law. For example, UK pensions law specifies various forms of limited price indexation for minimum pension increases for pension benefits accrued over certain periods. These forms of limited price indexation are discussed in chapter 5.

4.1.4. The annual SANLAM survey states that 71% of SA retirement funds surveyed link increases in disability benefits to the CPI². Kendal & Franklin (1993) survey the pension increases recommended by valuers of SA pension funds and find increases are usually benchmarked against the annual increase in headline CPI. The majority of funds recommended a percentage of CPI, for example 75% of CPI. Affordability is a consideration since the Pension Funds Act of 1956 provides for a minimum pension increase of the lesser of the increase in CPI and the investment performance of the assets backing the pensions. It is generally not possible to reduce pensions in payment through negative pension increases. Pension increases are therefore often linked in some way to inflation, but with a 0% floor.

4.1.5. At present in South Africa the valuation of pension liabilities is highly contentious. Legislation regarding the equitable disposal of pension fund surplus requires that surplus be measured on a realistic basis. Labour and employers are far from agreement on what constitutes a realistic basis. The market valuation basis proposed in this research may be viewed as sufficiently realistic and objective by both parties.

² 'Annual Survey of Retirement Benefits in South Africa', SANLAM, 2007.

CHAPTER 4

Linear inflation liabilities

In this chapter the valuation of inflation-linked pension liabilities which are linear functions of zero coupon inflation bond prices are addressed. Market values of linear inflation liabilities can be produced directly off zero coupon curves if the market is sufficiently complete for the law of one price to be feasible.

It is demonstrated that the use of a zero coupon real yield curve or a combination of nominal- and inflation-bond zero coupon yield curves is equivalent since both approaches produce the same market value. The advantages of using swap rates as an alternative basis for valuation are discussed. We address the modelling of bond and swap curves, specifically for the SA inflation market, including the need to allow for inflation seasonality and projection of the long end of the curve.

Even the inflation-bond and swap market poses interest and inflation rates incompleteness issues and these are discussed. The actuarial hedging approach of immunisation and financial economic theories of hedging are then described and compared.

1. Notation

1.0.1. While cashflows indexed with inflation are paid at discrete intervals, it is convenient to define the following continuously compounded rates

$f(t, \tau)$ is the instantaneous nominal forward-rate at time t for term τ ;

$g(t, \tau)$ is the instantaneous real forward-rate at time t for term τ ; and

$Q(t)$ is the (spot) inflation index published for time t .

Then the time t price of the nominal zero-coupon bond with maturity time T is

$$P^n(t, T) = \exp \left[- \int_0^{T-t} f(t, \tau) d\tau \right],$$

and the price at time t of a real zero-coupon bond with maturity time T is

$$P^r(t, T) = \exp \left[- \int_0^{T-t} g(t, \tau) d\tau \right].$$

Hence the nominal price at time t , including inflation indexation, of the real zero-coupon bond with maturity time T is $\frac{Q(t)}{Q(0)} P^r(t, T)$. Further, let $R(t, T)$ be the spot nominal rate at time t for maturity time T , defined as

$$R(t, T) = -(\ln P^n(t, T)) / (T - t).$$

Similarly, let $S(t, T)$ be the spot real rate at time t for maturity time T be defined as

$$S(t, T) = -(\ln P^r(t, T)) / (T - t).$$

1.0.2. Inflation-linked asset or liability cashflows must necessarily be indexed to lagged inflation to accommodate the delay between the inflation-basket survey date (generally the first day of the month) and its publication. The CPI is typically published monthly. Indexation of cashflows is based on:

- for dates other than the first of the month, interpolation across the cashflow month using lagged indices at the start and end of the month (inflation-bond style); or
- the inflation index published for a previous month (as is common for pension increases).

1.0.3. The instantaneous forward breakeven inflation rate at time t is then

$$b(t, T) = f(t, T) - g(t, T) \tag{1}$$

in continuously compounded units with the same indexation lag as implicit in the real rate g . It is more convenient to define breakeven inflation with respect

to a base inflation index, $Q(0)$, and no indexation lag as

$$b(t, T) = \ln[Q(t)/Q(0)]/(T-t). \quad (2)$$

Here t and T are defined for monthly points. Theoretically $Q(t)$ should be rounded to one decimal place as published but it will make little difference to a typical pension valuation to use an unrounded $Q(t)$. There is a trend to publication of inflation with greater precision, e.g. Euro-harmonised inflation is now published to two decimals, giving a further reason for not rounding forward values of $Q(t)$.

2. Breakeven inflation and real yield valuation

2.0.1. Figure 2 shows a typical pension increase profile. If increases are granted in January, the index lag is 3 months for the January payment, increasing by one month for each subsequent month's payment. The index lag for December payments is effectively 14 months.

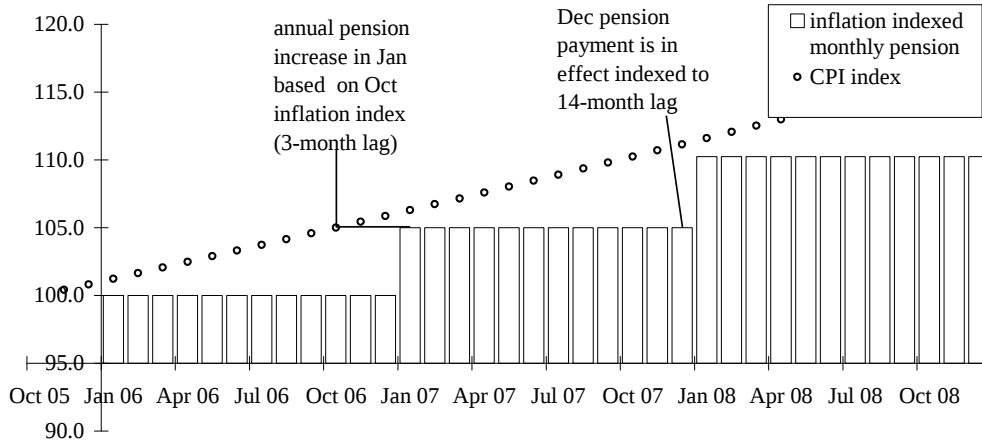


Figure 2: Illustration of variable indexation lag for monthly pension payments with annual increases

2.0.2. A conventional actuarial market valuation of inflation-linked liabilities would discount the liabilities at the real gross redemption yield of an inflation-linked bond of comparable duration (Griffiths, Imam-Sadeque, Ong & Smith 1997). If greater precision was required a valuation would use a real yield-curve point corresponding to the term of each cashflow, but a monthly pension indexed annually in arrears requires a different real yield-curve for each payment month index lag. These curves would each differ from the real yield-curve for the inflation-linked bond market, making the calibration of the curves a challenge.

2.0.3. It may be considered easier to price the liability by indexing up the real cashflows with breakeven inflation and discount these projected nominal cashflows with a nominal curve. The breakeven inflation identity (1) ensures this method produces the same value as discounting at the real yield. That is,

$$P^r(t, T) = \frac{Q(T)}{Q(t)} P^n(t, T).$$

2.0.4. Modelling breakeven inflation has several advantages:

- (1) In addition to the nominal curve already used to value nominal liabilities, only a single breakeven curve is required.
- (2) Modelling forward inflation directly using a breakeven curve facilitates inclusion of inflation forecasts and seasonality. The curve can be adapted to value related indices (such as the SA CPI excluding interest rates, CPIx).
- (3) In the more advanced inflation-swap markets the inflation tradable with the greatest market depth and price transparency is the zero-coupon inflation swap (Deacon, Derry & Mirfendereski (2004));
- (4) Risks of real and nominal liabilities can be evaluated and managed on a consistent basis through a common nominal curve.
- (5) Dynamic modelling of non-linear and forward starting pensions can include serial correlation between nominal and inflation curves.
- (6) It may be easier to set a nominal margin to the nominal swap curve to discount indexed liabilities than to set an equivalent real margin to the real swap rate.

2.0.5. Much of the curve-modelling literature dates from when the inflation-bond market was deeper than the inflation-swap market and is therefore directed at modelling real yield-curves. This research is directed at economists and central banks with the objective of estimating inflation expectations as the difference between government nominal and inflation-linked bond curves rather than to price inflation derivatives. Before considering the modelling of yield curves in detail, the salient features of the SA inflation market are examined and market disequilibrium and the implications for market values are discussed.

3. Inflation markets

3.1. Inflation-linked bonds

3.1.1. Figure 3 shows the size of the SA real and nominal debt markets using data from the Bond Exchange of SA at 30 June 2006. Inflation-bond notionals have been rebased to this date. The total outstanding notional of R685.5 billion falls well short of the R1098bn total assets of the SA retire-

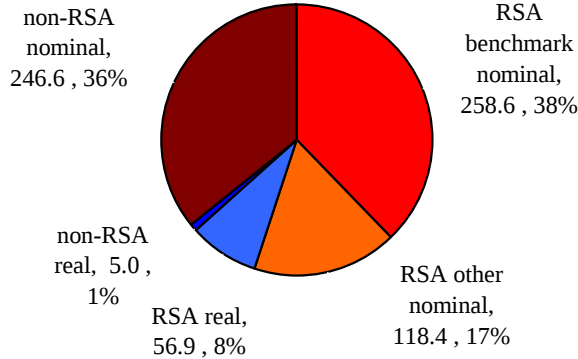


Figure 3: SA bond market, 30/06/2006

ment industry at 31 December 2004 given in the latest Financial Services Board report¹. Inflation-linked bonds constitute 9% of the market by notional but a higher proportion when bond issues are weighted by duration. The swaps market generates further hedging capacity as discussed in section 6.

3.1.2. The benchmark nominal SA bonds, in which issuance and liquidity is concentrated, are the R194 2008, R153 2010, R157 2015, R186 2026 and, subsequent to the above snapshot, the R209 2036 issued on 18 July 2006. Since government is able to issue longer than other (agency, bank and corporate) borrowers, these benchmark bonds account for a greater proportion of the total market weighted by duration than by outstanding notional.

3.1.3. The SA inflation-linked bonds are the R198 2008, R189 2013, R197 2023 and R202 2033, which collectively generate the real cashflows shown in figure 4. These bonds trade on real yield and are indexed to the CPI in a manner consistent with the major inflation-linked bond markets, but with a 4-month lag². SA inflation-linked bonds cannot be stripped. The SA inflation-linked bond market is incomplete along the dimensions of maturity and indexation lag.

3.2. Inflation swaps.

3.2.1. The SA inflation-swaps market is less advanced than the larger G7 markets. While the market for inflation-swaps between dealers and investment institutions has been active for the last five years, the limited interbank inflation-swap market restricts price transparency. The UK, US and Euro

¹ South Africa Financial Services Board Annual Report, 2006.

² The Bond Exchange of South Africa pricing methodology for inflation-linked bonds sets out the conventions.

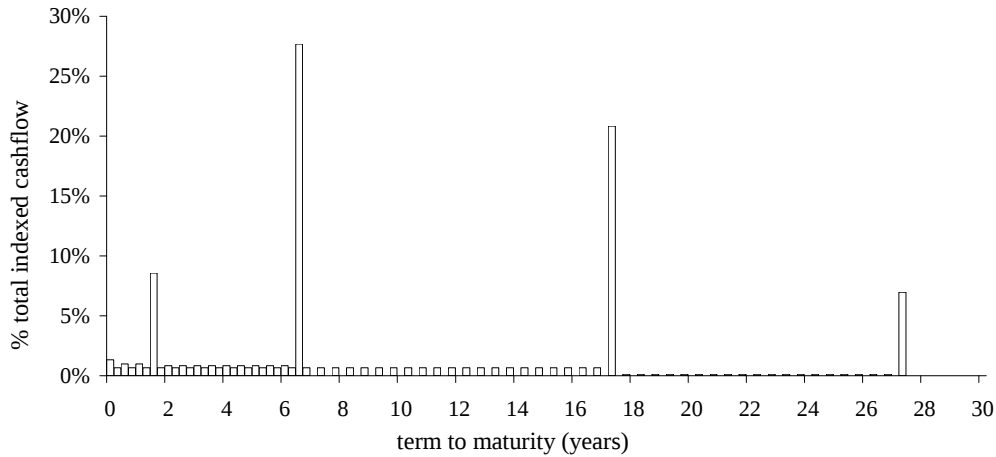


Figure 4: Real cashflows for 2008, 2013, 2023 and 2033 SA inflation-linked bonds, indexed to 30 June 2006

inflation markets have an active zero-coupon inflation-swap market where a compounded fixed rate is swapped for compounded inflation. This allows a breakeven curve to be bootstrapped to within 3 basis points out to 30 years, and frequently out to 50 years. The Euro market has monthly inflation futures contracts out to a year³. The SA interbank market is not actively brokered and daily broker closing levels for inflation swaps are not yet available, as they are in more developed markets. Trades are infrequent and usually structured as a swap of compounded JIBAR for a compounded real rate. JIBAR is the Johannesburg Interbank Agreed Rate, an average of bid and offer deposit rates compiled by the South African Futures Exchange.

3.2.2. It is possible to swap the cashflows of an SA inflation-linked bond to JIBAR but the bonds do not trade in conjunction with such a swap (termed an asset-swap). Inflation-linked bonds frequently trade as asset-swaps in the UK, US and Euro markets. The asset-swap levels of inflation-linked bonds determine the relationship between the cost of hedging inflation in the cash and swap markets. An active asset-swap market for inflation-linked bonds requires an efficient reverse repo market, to allow dealers to sell short inflation-linked bonds. SA National Treasury has decided recently to include inflation-bonds in its weekly reverse repo activities.

3.2.3. A pension fund can, in principle, hedge the inflation and nominal rates risk of its liabilities by transacting swaps with the required indexation lags (Waisberg et al. 2004). The swap will generally have credit support in

³ 'CME Eurozone HICP Futures', Chicago Mercantile Exchange, 2005.

the form of government bonds posted as security so credit risk is mitigated. SA pension funds will have the form, but not the economic substance, of the swap restricted by Regulation 28 (Waisberg et al. 2004). There were also tax implications from the Tax on Retirement Funds Act (act 38 of 1996) until it was recently repealed. Tax undermines the otherwise linear nature of a pension liability and will be discussed further in the next chapter.

3.2.4. Inflation swap dealers generate inflation supply through swaps and loans with corporates, for example the National Roads Agency paid R500m of 3-year inflation to a domestic bank in April 2002. This provides banks with hedging opportunities beyond the constraints of the government inflation-linked bond market.

4. Market disequilibrium

4.0.1. Thomson (2002) and Van Bezooeyen, Mehta & Smith (1999) address several reasons that have been advanced for not applying the law of one price to value inflation-linked pension liabilities with reference to inflation bonds, disregarding background risks, including:

- the inflation-linked bond market is not in equilibrium due to an imbalance between supply and demand;
- the size of the market is too small in relation to the size of the liabilities to allow the liabilities to be completely hedged.

4.0.2. Figure 5 shows the movement in real yields for the SA inflation-linked bonds from issue to 30 June 2006. The early years were characterised by

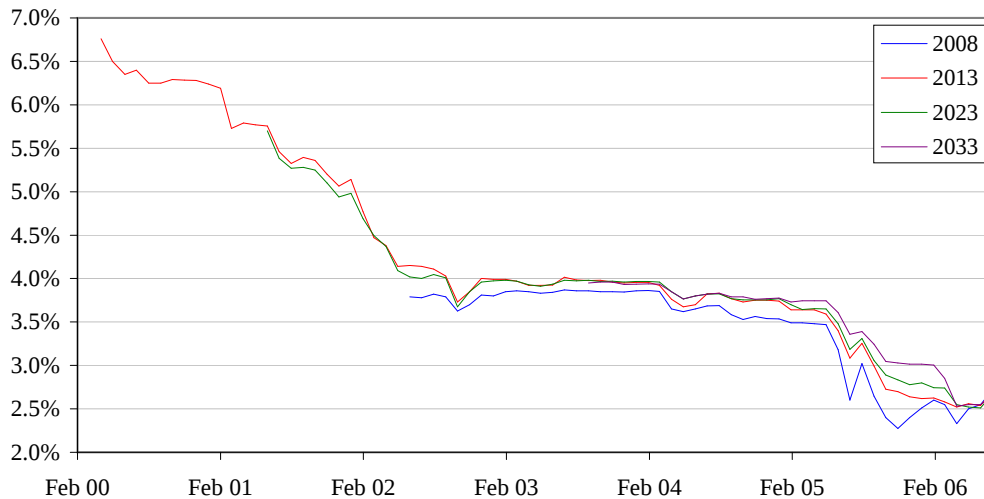


Figure 5: SA inflation-linked bond real yields

extreme illiquidity. Price discovery was limited to auction dates and restricted auction sizes generally forced the real yield down. Investors were reluctant to bid for bonds in the secondary market for fear this signal would drive real yields down further. In the past three or so years it has become much more of a two-way market as many investors have accumulated their benchmark positions and have become free to sell inflation-linked bonds when judged expensive. There is also more liquidity available to buy and sell inflation through the swaps market and through bank inflation-linked bond placements. We now have stronger evidence that supply and demand are in balance and the market is in equilibrium.

4.0.3. Even if the market is not in equilibrium, the existence of banks that are prepared to transact swaps which mimic the inflation and interest rates elements of a given pension liability allows the LOOP to be invoked to place bounds on the market price. As discussed in chapter 3, the market value of liabilities relates to a balance between marginal supply and demand rather than total supply and demand. This mid-market value will not be appropriate if the intention is to transfer the liabilities to a third party. The offer market value of a pension liability should then be based on the price which balances demand consistent with the size of the liability with supply consistent with the market impact of the hedging transaction. The limited pool of assets in relation to total pension liabilities has no further relevance for the value of the liability. Small rates-hedging deals are dealt in SA every month at market levels.

5. Curve modelling

5.1. Valuations

Section 4 of chapter 3 discussed the margin to LIBOR at which the liabilities are to be discounted. The liability discount curve should be set as a spread to the swap curve since a swap achieves a far better hedge of the interest-rates and inflation characteristics of the liability. The margin to the swap curve reflects the degree of security with which the liability is funded. In markets with visible interest and inflation swap rates, such as the UK, the nominal and breakeven curves can be simply bootstrapped from LIBOR and zero coupon inflation swap market rates and no curve modelling is required. However, in SA the absence of tight public inflation swap screens or marks from dealers or brokers forces the use of nominal bond and implied bond breakeven curves. The corresponding inflation swap curve can then be derived and the appropriate margin to JIBAR applied before discounting.

5.2. Rationale

5.2.1. Market instruments embody information about the prices of zero-coupon bonds but the number of instruments traded in the market is typically sparse compared with the range of future cashflow dates. This is particularly true for real instruments where there is the additional dimension of the indexation lag. Let m_i denote the market settlement price of market instrument $i = 1, \dots, I$ with maturity T_i and cashflow c_{ij} at future time $j > 0$. This describes an interest-rate market at time t as a system of I equations in n , the number of future cashflow dates, unknowns:

$$m_i = \sum_j^{T_i} c_{ij} P^n(t, j) \quad ; \quad i = 1, \dots, I. \quad (3)$$

The above set of equations relates to the nominal market where m_i is the dirty price in the case of a nominal bond and zero in the case of a nominal swap traded at the market level. An analogous system applies for the real market with the added complication of inflation indexation for a given index lag.

5.2.2. Curve models hypothesise that zero-coupon bond prices in between tradables are smooth functions of term to maturity. A lack of smoothness creates irregular forward-rates that are inefficient as in a perfect market these can be exploited to earn expected positive profits with low risk. Cashflows of arbitrary maturity can then be priced off the curve in a manner consistent with market rates.

5.2.3. The functional form imposed on the system will often model forward-rates or zero rates rather than bond prices, but these representations are equivalent. Since $I < n$, there are infinitely many curves which satisfy the system. Therefore the curve is constrained to positive forward-rates and additional objectives are set such as smooth forward-rates and a parsimonious functional form. These objectives may be prioritised at the expense of the fit to market rates so that resulting rates do not satisfy the system but the discrepancy between fitted and market rates is minimal.

5.3. Nominal curves

5.3.1. Much has been written on curve modelling, including research by Anderson et al. (1996), James & Webber (2000), Cairns (2004b). Stander (2005) addresses curve modelling in SA. There are two dominant approaches:

- (1) Parametric curve methods fit a smooth curve to forward-rates using least squares, maximum likelihood or Bayesian methods. Examples include Nelson & Siegel (1987) and the extension by Svensson (1995). Cairns (1998) demonstrates how shifts between local and global minima for the objective function can result in extreme jumps between

curve shapes despite no change in bond yields. He proposes a methodology similar to Svensson but fixing certain parameters to minimise the risk of such jumps.

- (2) Spline methods fit piecewise polynomials between knot points on the zero-coupon curve. Examples using cubic splines and a combined goodness-of-fit and roughness objective are Fisher, Nychka & Zervos (1995) and the variable roughness modification to this method by Waggoner (1997). The Bond Exchange Actuarial Society of South Africa (BEASSA) yield-curves use quartic splines and knots at the maturities of the BEASSA All Bond Index SA constituents on which the yield-curve is based⁴. Splines may be more stable across segments than parametric curves but may display exaggerated kinks within segments if knots are too close together. Anderson & Sleath (2001) suggest the subjectivity of knot placement is mitigated if roughness is penalised through the objective function, as in the methods above.

5.3.2. Figure 6 shows the published and perfect-fit BEASSA nominal zero-coupon bond yield curves on 26 June 2006. The published curve is smoothed using a roughness penalty in the objective function, creating the price discrepancy between the two curves for the benchmark SA bonds shown in Table 1.

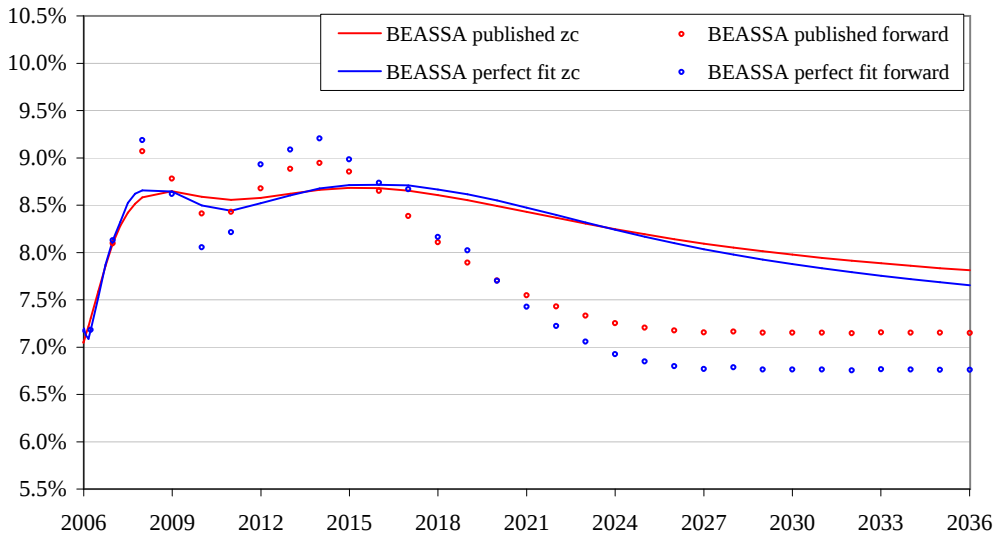


Figure 6: SA zero-coupon nominal bond-curves, 26 June 2006

⁴ An introduction to the Bond Exchange Actuarial Society of South Africa zero coupon bond yield curves, 2003.

Table 1: Pricing discrepancy between BEASSA perfect-fit and published nominal zero-coupon bond curves at 26 June 2006

RSA bond	2008	2010	2015	2026
yield discrepancy	0.05%	-0.07%	0.01%	-0.01%

5.3.3. The longest bond to which the above curves were fitted was the R186 2026. Section 10 addresses how the SA nominal curve was subsequently extended with the issue on 18 July 2006 of the R209 2036.

5.3.4. The emphasis of the methods surveyed above is curve smoothness and stability in order to provide a good description of the market or assist monetary policy. The trade-off between fidelity to market rates and curve smoothness quantified through the objective function often results in an imperfect fit. Cairns (1998) highlights the risk of a catastrophic jump that may occur in a curve fitted with a minimisation algorithm. Such a jump is caused by the global minimum shifting between two or more local minima corresponding to distinctly different curve shapes. The jump is considered to be catastrophic when the magnitude of the move in some segment of the curve is far greater than the move implied by the change in market rates alone.

5.3.5. Derivatives pricing will prioritise the fit to market data and maximise smoothness subject to recovering market prices of the benchmark underlying bonds. The pricing curve used by an interest-rate derivatives desk will recover the prices of the instruments used to hedge the book since the derivatives desk trades its hedge instruments with the market-making desk at market levels. Similarly, the market valuation of inflation derivatives embedded in pension liabilities should be based on a curve that recovers the market price of benchmark bonds. Smoothness should be viewed as a desirable but secondary objective.

5.4. Real curves

5.4.1. Real yield-curves are more challenging to model than nominal curves:

- There are few real instruments to calibrate the curve. Inflation-linked bonds are less liquid and have generally have wider bid–offer spreads than their comparator nominal bonds of the same maturity.
- In the absence of a liquid inflation futures market, there is no market real short-rate to anchor the short end of the curve. The notion of an instantaneous short-rate is defensible in nominal interest-rate space where the overnight deposit swap (the Rand Overnight Deposit Swap or RODS market in SA) and repo markets provide one-day rates for

the swap and bond curves respectively. However monthly inflation surveys rule out an analogue in real interest-rate space.

- There are no true real instruments as the necessity of a publication lag causes any inflation-linked instrument to yield a hybrid between a true (*i.e.* perfectly indexed) real rate and a nominal rate. The true real yield will not show a smooth progression due to inflation seasonality (see section 7) and the hybrid real yield will show additional variability from the published inflation index over the index lag.

5.4.2. Anderson & Sleath (2001) note the restricted Svensson (1995) real- and nominal-curves used previously at the Bank of England (BOE) were not reliable out to two years due to the sparseness of the curve and the curve form did not provide sufficient curvature to capture expectations of rate movements in the short-term. This suggests a parametric curve may be less suitable than the Waggoner (1997) spline method later adopted by the BOE or the Fisher, Nychka & Zervos (1995) spline method used by the US Federal reserve.

5.4.3. The current BOE method described in Anderson & Sleath (2001) draws on the approach of Evans (1998) to find the true real yield-curve allowing for the indexation lag. The lag for the original-design UK index-linked gilt is 8 months, but the UK DMO has also issued index-linked gilts with a new-design based on the standard 3-month lag used in other inflation-linked bond markets. Deflators are defined that use the inflation risk premium to translate between inflation-indexed and nominal cashflows to estimate the real yield curve for any lag. This approach is, however, designed for short-term inference of inflation expectations for monetary policy and is of limited use for pricing the annual uplift and forward start convexity features of inflation-linked pensions. These features are a cause of incompleteness and are discussed further in section 13 where it is argued that a dynamic model as described in chapter 5 is a more effective way to value these and non-linear effects.

5.5. Breakeven inflation curves

5.5.1. In section 2 it was argued that the optimal approach for pricing the inflation derivatives embedded in pensions is to project the inflation index using breakeven inflation curves and discount these projected cashflows using nominal curves. There are three sources of market prices for building breakeven inflation curves:

- (1) The inflation-swaps market. The breakeven curve is bootstrapped from market inflation swap rates.

- (2) The inflation-linked bond market. The swap curve is spread off the breakeven curve. The inflation swap curve will likely be linked directly to the price information from smaller and more frequent bond trades, the spread adjusted from time to time to reflect distinct swap market supply and demand dynamics.
- (3) Inflation-linked bond asset-swap levels. The swap curve is derived from inflation-linked bond prices and the margins to the swap curve at which inflation-linked bonds trade as asset-swap packages. If asset-swap investors are dominant, the price at which inflation-linked bonds trade as asset-swap packages will show good stability and price transparency and this price information may be used as the primary pricing source for the breakeven curve.

5.5.2. In SA, the nascent and opaque inflation swaps market rules out the first and third approaches. The absence of a basis between swap and bond breakevens, discussed in section 6, implies setting the swap curve equal to the bond breakeven curve. To be used for derivatives pricing, the inflation swap curve must incorporate seasonality and an inflation forecast or futures strip, the topics of sections 7 and 8.

6. The bond-swap inflation basis

6.0.1. The inflation-linked bond and swap markets may imply significantly different levels for the forward inflation index. This is termed the basis between the inflation-linked bond and swap markets. The basis can be quantified for a particular bond using its swap spread. The swap spread for an inflation-linked bond is the spread to the swap curve required to discount the outstanding bond cashflows to equal the current bond price. An inflation-linked bond swap spread can be decomposed as

$$iss = nss + b$$

where iss is the inflation-linked bond swap spread, nss is the swap spread on a nominal bond with equivalent indexed cashflows and b is the inflation-linked bond-swap basis.

6.0.2. A long position in a government bond that is financed in the repo market and receiving the fixed cashflows on an interest-rate swap are essentially the same transaction, so the difference in the bond and swap rates should reflect the difference between the bond's average future repo rate and the swap floating rate (termed LIBOR without loss of generality). This average future repo rate will generally be close to the general collateral (GC) repo

rate generally available on government bonds used as collateral for the repo transaction. However, there is the potential for an occasional windfall if the bond trades special on repo (when the repo rate is below the GC repo rate).

6.0.3. Other factors relating to the balance between supply and demand for swaps and government bonds may be more significant than this GC LIBOR spread. Cortes (2003) and Mussche (2002) identify the slope and level of the curve, equity-market-implied volatilities and the budget deficit among such factors. Swap spreads also reflect systemic banking sector credit risk as the LIBOR fixing leg is based on a poll of interbank deposit rates. This should not be confused with the counterparty credit risk of the swap itself, which is normally minimal since the mark-to-market value of the swap is fully collateralised.

6.0.4. The inflation-linked bond-swap basis will reflect the balance between supply and demand for inflation in the cash and swap markets. Nascent markets may be unbalanced, with government bonds the sole source of supply. In time, investors buying inflation-linked bonds on asset-swap and issuers of swapped inflation debt emerge, helping to balance the market. SA inflation swaps are at present priced to within the swap bid-offer margin of bond breakevens so there is no significant inflation-linked bond-swap basis. This implies that inflation-linked bond-swap spreads will be similar to swap spreads for nominal bonds with the same (indexed) cashflow profile. Figure 7 plots swap spreads for SA inflation and nominal bonds on 26 June 2006 against duration. The difference between inflation and nominal swap spreads is attributable to the different cashflow profiles, which duration standardises for only in a crude way. SA inflation-linked bonds can typically be financed at the GC repo rate which has been close to JIBAR, so there is no immediate cost to hedging short inflation swap positions with bonds. However, there is no SA inflation-linked bond asset-swap market and no resulting sellers of inflation to balance demand for inflation swaps. Dealers may begin to price an inflation-linked bond-swap basis into swaps as there is a limit to the amount of basis risk they can warehouse.

6.0.5. It should be noted that even an active inflation-linked bond asset-swap market will not in itself force swap breakevens in line with bond breakevens. Inflation-linked bonds are not fungible with nominal bonds so an asset-swap investor will not necessarily require the same swap spread on real and nominal bonds. It may be that inflation-linked bonds are expected to trade special on repo less frequently than nominal bonds, implying a less negative real swap

spread. Conversely, since a sovereign can debase nominal but not inflation-linked debt, the inflation-linked bond asset-swap spread may be more negative than for nominal bonds. In both the US and UK markets inflation-linked bond asset-swap spreads have at times traded over 5 basis points away from nominal bond asset-swap spreads for similar durations⁵.

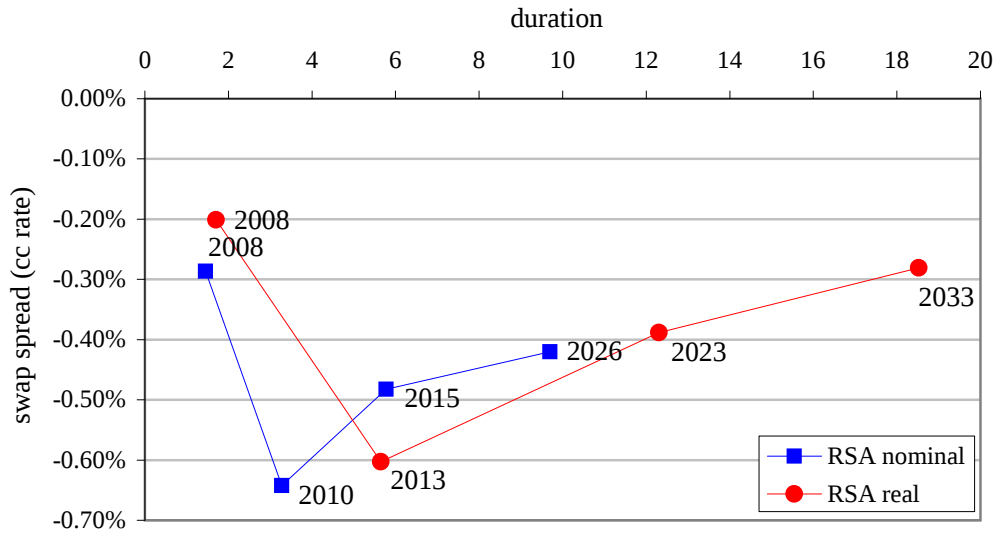


Figure 7: SA nominal and inflation-linked bond-swap spreads (continuously compounded rates) assuming no inflation-linked bond-swap basis, 26 June 2006

7. Inflation seasonality

7.1. Sources of seasonality

7.1.1. Inflation, like many economic time series, has a regular seasonality pattern. In South Africa this CPI index seasonality pattern stems from:

- irregular survey dates for certain basket items (e.g. education costs are surveyed annually in March); and
- seasonal price variations of certain basket items (e.g. food).

7.1.2. In many countries the statistical agency responsible for producing the index for consumer prices release publishes figures both with and without seasonal adjustment. From March 2005 Statistics South Africa stopped producing a seasonally adjusted CPI due to irregular behaviour of the series but continues to research whether it can “produce a reliable series for groups, such as food, which do exhibit seasonality”⁶. This suggests a bottom-up approach

⁵ Barclays Capital Research, Global Inflation-Linked Monthly, June 2006.

⁶ Statistics South Africa March 2005 Inflation Release, P0141.1.

where seasonality is separately modeled for each CPI basket group and then combined, although this is not addressed in the literature.

7.2. Seasonality modelling

7.2.1. There are two distinct approaches to modelling seasonality:

- (1) stochastic seasonality included with a stochastic inflation and interest-rates model; and
- (2) deterministic seasonality adjustments to forward inflation (which may be stochastic).

7.2.2. The impact of seasonality on the present value of an inflation-linked liability, relative to interest-rate risk factors, depends on:

- the term of the liability, longer liabilities being less sensitive;
- the spread of cashflows, an annuity-style liability being less sensitive than a bullet liability; and
- whether indexation applies from the current date or some future date, since forward-indexed liabilities are exposed to seasonality on both the start and end of indexation.

Liabilities in respect of inflation-linked DB pensions in payment are therefore fairly insensitive to the seasonality assumption. Furthermore, in view of the parameter risk of the seasonality factors estimated below, it may be spurious to model seasonality with a high degree sophistication. Chriqui & Tzucker dismiss stochastic seasonality as “complicated and redundant” in general⁷.

7.2.3. Deterministic models for CPI seasonality use econometric and harmonic analysis theory to decompose the series into an underlying trend and monthly seasonality factors. Two alternative approaches commonly proposed (Belgrade & Benhamou 2004a) are:

- (1) parametric methods, such as the linear regression; and
- (2) nonparametric methods, such as the X-11 and X-12 methodologies developed by the US Census Bureau⁸ based on moving average filters.

7.2.4. The X-11 methodology is not suitable for South African seasonality since the second stage uses a moving average to reduce noise in the seasonality factors under the assumption of a smooth seasonality profile over the year. This is inappropriate for the periodic survey dates effect. The 13-period moving average of the preliminary stage of standard X-11 may not precisely isolate the underlying trend of the volatile SA CPI series.

⁷ ‘Advanced Inflation Derivatives’, Presentation, Barclays Capital Global Annual Inflation Conference, Miami(2006)

⁸ ‘X-12-ARIMA Reference Manual, Version 0.2’, US Census Bureau, Washington, DC, 2002

7.2.5. Figure 8 shows the X-11 trend, the linear regression trend described below and year-on-year CPIx (the SA CPI excluding interest rates). CPIx is a better basis than CPI for stripping out seasonality since it excludes the confounding effect of interest rates.

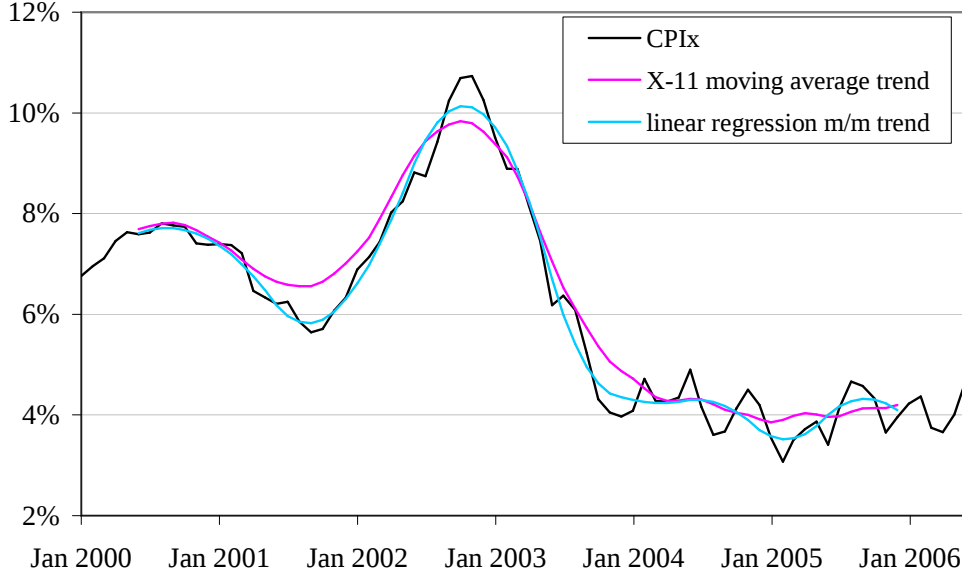


Figure 8: Underlying SA CPIx trend excluding seasonality produced using X-11 preliminary stage and linear regression on the month-on-month trend

7.2.6. A linear regression seasonality model may be defined as:

$$\ln\left(\frac{Q_t}{Q_0}\right) = m_t + s_{t'} + \varepsilon_t \quad (4)$$

where t is time in years at discrete monthly intervals, $t = \frac{1}{12}, \frac{2}{12}, \dots$;

$t' = \text{int}\left(12 * \left[\frac{t}{12} - \text{int}\left(\frac{t}{12}\right)\right] + 1\right)$ is the month of time t ; $t' = 1, \dots, 12$;

m_t is the underlying linear cumulative inflation trend to time t ; $t > 0$;

$s_{t'}$ is the month t' seasonality factor, such that $\sum_{i=1}^{12} s_i = 0$; and

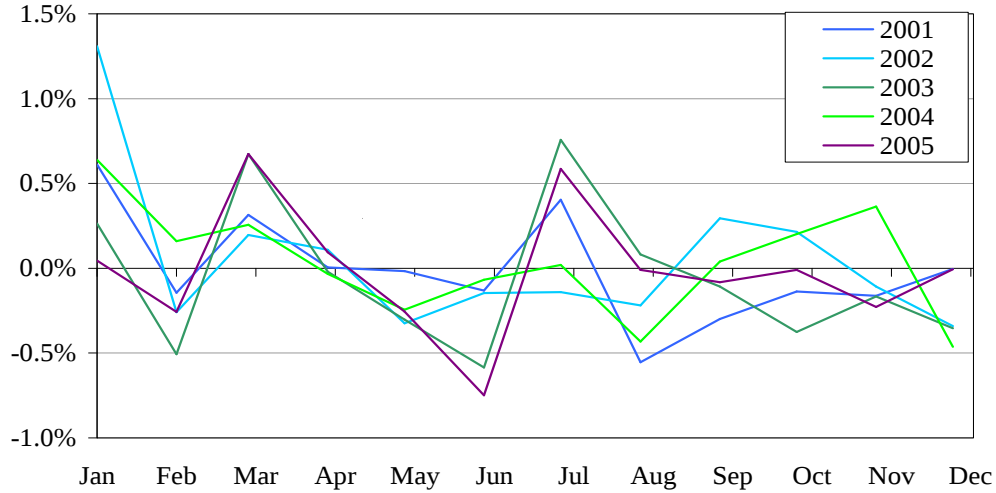
ε_t is a sequence of i.i.d. normal error terms.

7.2.7. Equation 4 has a multiplicative seasonality factor (additive in the log of the index). This is suitable for the SA CPI due to the periodic survey dates for some basket items. Table 2 shows estimates of month-on-month seasonality factors, s_t , for SA inflation based on June 2000 to June 2006 for SA. A 6-month piecewise linear m_t function was used to account for the volatility of the inflation series.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
s_t	0.58%	-0.19%	0.43%	0.04%	-0.22%	-0.33%	0.30%	-0.23%	-0.04%	-0.04%	-0.11%	-0.19%

Table 2: Estimated SA CPI seasonality factors, June 2000 to June 2006

7.2.8. Plots of the month to month seasonality pattern for each year from 2001 to 2005 (i.e. the deviation $\ln(Q_t/Q_{t-1}) - (m_t - m_{t-1})$ from the linear regression trend) are shown in figure 9. Noise masks the seasonality pattern.

**Figure 9:** Empirical seasonality profiles (deviation against trend in 4)

7.2.9. The parametric approach using regression has the advantage that confidence intervals can be produced for seasonality factor estimates. The estimates and 90% confidence intervals are shown in figure 10. These confidence intervals only show the parameter risk. There is additional risk that the choice of seasonality model may be incorrect. Only the March and May factors are (just) significant at the 90% level. Seasonality model risk is high, but seasonality may be considered significant enough to warrant inclusion in the valuation. Based on the December seasonality factor to January base of 0.994 in table 2, valuing a pension with January inflation increases as December increases is expected to result in about 5 real yield basis points undervaluation for the 60-year-old male specimen annuity of Appendix A.

7.3. Materiality

7.3.1. While seasonality may have a material effect on pension valuation, the variability within each year of the CPI from the seasonality process risk identified above is spread across the future real cashflows in a typical liability profile. Figure 11 shows the risk from variation in seasonality for two different

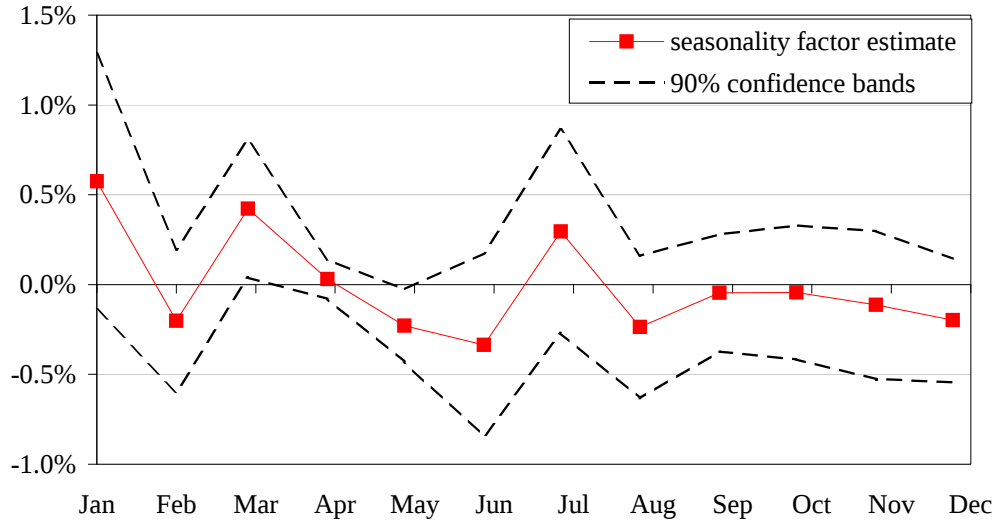


Figure 10: Estimated monthly seasonality factors and confidence intervals for SA CPI using linear regression

liabilities under the hypothesis of the linear regression model fitted above. The two liabilities compared are a 15-year zero-coupon inflation liability with full indexation starting in 12 months and no mortality risk and the 60-year-old male specimen real annuity of Appendix A with full indexation commencing at the valuation date. The risk is defined as the difference in PV (discounting at SA real rates for 26 June 2006) between the liability indexed to January and matching inflation assets with the same cashflow years but indexed to July. Standard deviations of the present value are 0.14% for the real annuity and 0.78% for the forward starting zero-coupon inflation liability.

7.3.2. Seasonality risk only affects the distribution between the months of the annual inflation increase, so if seasonality deviations are independent they are likely to offset. Note that this assumption is unlikely to be true in practice as there may be a systematic change in seasonality pattern (for example due to change in survey methodology). Forward stating liabilities are exposed to seasonality at both the indexation start and end dates. Seasonality risk for the typical pension liability profile hedged with assets indexed to inflation up to a year from the liability indexation dates will be limited and may be assumed to be independent of other risks.

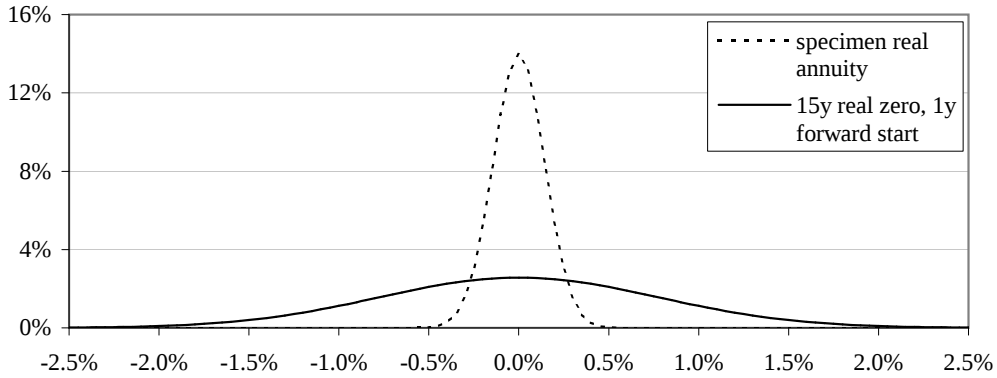


Figure 11: Hedging risk from seasonal variation in inflation for zero-coupon inflation liability with forward-starting indexation versus specimen real annuity of Appendix A under the hypotheses of the regression model 4 and seasonality factors of Table 2

8. Inflation forecasts

8.1. Flow seasonality

The market-implied forward inflation index at each future inflation indexation date will incorporate index seasonality patterns and possibly also flow seasonality, defined as price distortions from an imbalance of supply and demand for inflation on certain future dates (Belgrade, Benhamou & Koehler 2005). In the Euro Harmonised Index of Consumer Prices (HICP) inflation market where substantial short dated inflation exposures (or "reset risks") have accumulated, flow seasonality can be extracted from inflation futures and the spreads between inflation swaps maturing on different monthly date points. In South Africa the longer-dated nature of the market reduces the significance of flow seasonality. The limited number of SA inflation-linked bonds concentrates supply in certain indexation months but there is insufficient market information to calibrate seasonality factors.

8.2. Predictability of inflation

8.2.1. Ang, Bekaert & Wei (2005) found that inflation forecasts from surveys of economists were better out-of-sample predictors of US CPI than the models based on the nominal curve and economic or ARIMA models of the inflation time series. Elder et al. (2005) find strong evidence that the BOE RPIx (now CPI) fan-chart forecast is more accurate than the alternative of a regression on previous inflation and various economic indicators. In SA, Aron & Muellbauer (2006) find the South African Reserve Bank (SARB) one-year

absolute CPIx forecast errors to published inflation since targeting have averaged 1.3%. This is lower than the 2.3% standard deviation of the CPIx series itself.

8.2.2. So, it appears incorporation of forecast information into the inflation swap curve for the 12 months beyond the last published inflation figure can reduce swap pricing risk in the absence of a market for inflation-linked bonds or futures maturing in the next year. The breakeven curve is then anchored at the longest inflation forecast and extended to recover inflation-linked bond prices. Forecasts normally include the seasonality effect, implicitly if not explicitly, so the seasonality model is only applied beyond the forecast. The inflation swap curve results from applying the required inflation swap spread, if any, to the bond breakeven curve.

8.2.3. Forecasts introduce subjectivity and it is fair to question whether they are an improvement over the alternative of anchoring the breakeven curve at the latest published inflation figure. It is often the case that inflation-basket price changes are likely following recent movements in currency, oil prices, crop estimates, etc. Competitive pressures delay these price changes and their impact on the CPI but the eventual effect can be forecast with a degree of certainty that justifies using the forecast. The challenge of incorporating an inflation forecast in the curve is that there will be times, such as just after an inflation release, when updated forecasts are required but not available since it takes time to prepare a fresh forecast. However, the value of pension liabilities is generally calculated using curves generated from close of business market levels rather than at data-sensitive times.

8.3. Materiality

The inflation forecast will not have a material effect on the valuation of a pension annuity where a small proportion of future cashflows is due in the next year. For the 60-year-old male specimen annuity of Appendix A, the change in present value from increasing one-year forward inflation by 1% is equivalent to the change in present value from a decrease of 0.6 basis points in the valuation rate for the fully indexed liability. The effect of the forecast on the curve shape beyond one-year may be more significant for the valuation, especially if a parametric curve is employed.

9. Inflation risk premium

9.1. Fisher's relationship

9.1.1. Long before the first inflation-linked bond, Fisher (1930) theorised that the yield on a nominal bond, n , can be decomposed into expected inflation,

i , to compensate loss of purchasing power, an inflation risk premium, p , to reflect inflation uncertainty, and the real yield, r :

$$(1 + n) = (1 + r)(1 + i)(1 + p). \quad (5)$$

This can be extended to include premia in respect of, for example, liquidity and credit differentials. For small r and i , (5) can be approximated as

$$n = r + i + p. \quad (6)$$

9.1.2. While Fisher's relationship may seem a promising basis for building a breakeven inflation curve, the inflation risk premium is too nebulous to be useful. Section 8.2 cited evidence that inflation can be forecast with reasonable accuracy over a short horizon. This suggests that the inflation risk premium should be insignificant over the short term where forecasts are available, and not well defined beyond this. Surveys of long-term expectations that relate to consumers and employers are of more relevance to monetary policy than to derivatives pricing, while surveys of market forecasters are likely to produce expectations close to breakeven inflation for inflation bonds.

9.1.3. The inflation target is the only objective measure beyond the maximum term of forecasts, usually two years. The SA inflation target range is less helpful as it must be reduced to a point estimate using the mid-point or some other assumption. The inflation risk premium calculated in this way is very unstable and does not consistently show the characteristics expected intuitively, such as increasing monotonically with term and positive correlation with inflation volatility. Evidence from academic studies suggests the inflation risk premium varies with time, see for example the UK study of Evans (2003).

9.1.4. Consequently i and p may be combined into a breakeven inflation term, that we denote b . This is the discrete-time version of equation (1):

$$n = r + b. \quad (7)$$

The quantity b can be interpreted as a form of risk-adjusted inflation expectation and is usually modelled directly.

10. Curve extension

10.1. Necessity of extending valuation curves

10.1.1. Pension liabilities may extend well beyond the longest nominal and inflation-linked bonds in the market. In South Africa the longest nominal SA bond in the market at 26 June 2006 was the R186 maturing in December 2026, although swaps were available to 30 year term and government issued its R209

March 2036 bond on 18 July 2006. The longest inflation-linked bond is the R202 December 2033.

10.2. Yield-curve theory

10.2.1. Theories proposed to date for the shape of the yield-curve can at best explain only certain aspects of its behaviour (Hull 2003):

- Expectations theory holds that curve forward-rates equal the market's expectation of future spot rates. However, the price of a bond or swap is a weighted sum of zero-prices, which are a convex function of forward-rates. Hence, Jensen's inequality implies a positive convexity adjustment must be added to market forward-rates to estimate expected future spot rates. Expectations theory may therefore be considered to be consistent with inverted curves if forward rates are expected to be constant. This conflicts with the upward slope normally exhibited by real and nominal curves.
- Liquidity preference theory accounts for the upward curve slope by the presence of a risk premium. If the price volatility of bonds or swaps increases by term to maturity, this risk premium increases with maturity and an upward sloping curve results. However, in some markets the dominant investors in long-maturity bonds invest to match long-maturity liabilities and there may be less cause for a risk premium.
- According to the market segmentation theory, rates in each maturity band of the curve reflect the balance between investors and issuers in this band. The behaviour in each part of the curve could be somewhat distinct and a variety of curve shapes may result. In SA, the demand for short-maturity bonds by banks to meet liquid assets requirements may reduce the yields on these short-maturity bonds.

10.2.2. A combination of these theories can explain the inversion of long real and nominal rates seen in liability-driven markets such as the UK and SA, but such theories cannot necessarily be used to extrapolate the behaviour of the curve beyond the longest debt instruments. This is because it is unlikely market participants will have formed expectations about interest rates this far into the future and there will be a high degree of uncertainty about very long-term liabilities.

10.3. Convexity

10.3.1. Research by Brown & Schaefer (2000) and Ilmanen & Byrne (2003) has shown that convexity has a significant effect on the shape of the US Dollar, Euro and Sterling curves, but the scale of the convexity adjustment is very

sensitive to assumptions about the behaviour of the long end of the curve where there is limited implied volatility information from the options market. The convexity adjustment for the breakeven curve has the additional dimension of complexity of depending on the interaction between real and nominal rates. Gurkaynak et al. (2006) attribute the second hump in the long end of their parametric US Treasury curve to convexity.

10.3.2. Figure 12 shows three approaches to curve extension. The first method in panel (a) assumes the spot curve remains constant after the 30-year maturity of the longest traded instrument. This rather naive method produces a discontinuity in forward-rates.

10.3.3. An alternative approach is to extend the curve with a constant forward-rate as shown in panel (b). This ensures a continuous forward-rate. If the curve is built with splines, an analogous method is to assume the forward-rate is constant at the longest curve forward-rate, the approach used in the projection of the BEASSA zero-coupon yield curve.

10.3.4. Panel (c) shows a stylised example of the extrapolation of a parametric forward-rate curve of Svensson (1995) form.

10.3.5. Since there is limited market data to define the long end, the longest forward-rate projected using these methods may be unstable. This is particularly true for parametric curves, although giving smoothness a higher priority in the objective function will mitigate instability.

The prices of coupon strips can be used to give better definition to the long end, but Sack (2000) observes that extrapolation of the US STRIPS (Separate Trading of Registered Interest and Principal of Securities) curve does not produce a stable long forward-rate. In less liquid strips markets than the US

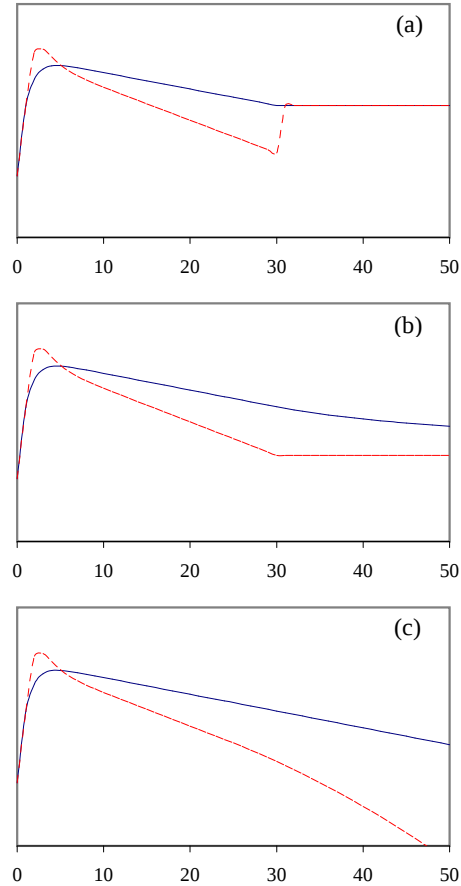


Figure 12: Stylised spot (solid) and forward (dashed) curves showing various extrapolation methods

there may not be sufficient liquidity to give reliable prices, e.g. the UK DMO (2000) do not include strips in their debt management curve.

10.3.6. The instability of the long forward-rates resulting from these curve extrapolation-approaches causes convexity gains or losses from revaluation of very long-dated liabilities. Since convexity increases dramatically with term to maturity, these losses can be high in relation to the present value of liabilities.

10.3.7. For method (a) the volatility of zero rates beyond the curve will be set equal to the volatility of the longest zero-coupon rate on the curve. For method (b) the volatility of long forward-rates will equal that of the longest point on the curve. The volatility of long forward-rates in method (c) is indeterminate. Only by coincidence will these methods probably accord with a declining term structure of market-implied forward volatility.

10.3.8. The forward curve extension should ideally account for the convexity effect. This will decrease forward-rates in the extended curve and increase the present value of liabilities with higher convexity than hedge assets available in the market. The correct convexity adjustment will balance convexity losses on rehedging duration with the unwinding of this valuation margin. However, the convexity impact cannot be purged easily from the extended forward curve as the convexity adjustment depends on the process for future forward-rates, as discussed in the next chapter. Even if the process is correctly specified and calibrated to market implied-volatilities, the volatility of the future evolution of the curve will likely deviate from implied volatility and convexity gains or losses will result. Here we can draw an analogy with delta hedging a short option position (Smith 2001*b*).

10.4. The Dybvig Ingersoll Ross Theorem (DIR)

10.4.1. In a perfect market, Dybvig et al. (1996) showed that the convexity of the limiting-term spot rate, $\lim_{T \rightarrow \infty} R(t, T)$, will produce arbitrage opportunities unless this limiting rate (which we term the “long rate”) is constrained to be nondecreasing with time t .

10.4.2. McCulloch (2000) noted that without the perfect market assumption, transaction costs will render the long spot rate undefined.

10.4.3. Empirical evidence has at times appeared to conflict with the DIR theory. McCulloch & Kochin (1998) estimate long real and nominal forward-rates for the US Treasury market using their spline-based method and conclude that “No evidence is found that the estimated forward-rate beyond 30 years is nondecreasing over time, or even has lessened variance.” Cairns (2004*b*) shows that the long forward-rate for the UK market estimated using the Cairns (1998)

restricted exponential curve has periodically decreased, although he notes that parametric estimates of long forward-rates may not be statistically significant.

10.4.4. Smith (2001*b*) argues that the long-rate should be fixed when pricing long dated liabilities, in accordance with the DIR result and to ensure a stable curve that minimises spurious convexity-hedging profits and losses.

10.4.5. Smith & Wilson (2000) describe a method for generating a spot curve with fixed long forward-rate and mean reversion parameters that maximises forward-curve smoothness and flatness (in the sense of the curve lying close to the long forward rate). The curve has a perfect fit to zero-coupon bond prices.

10.5. New bond issues

10.5.1. Swaps trade with constant maturity, while the term to maturity of bonds declines each day. A perfect-fit bond yield curve can only be extended when a new, longest maturity, bond is issued. Swap market liquidity will depend on bond-market liquidity and, in the absence of new debt issues, the swap curve will also gradually shorten.

10.5.2. Ideally the shape of the valuation curve will anticipate perfectly the effect of extension from new bond issues (Smith 2001*b*). If the new bond price in relation to existing bonds and swaps is that implied by the existing curve, there will be no hedge losses from revaluation at the new perfect-fit curve.

10.5.3. Valuation metrics the market uses to price a new bond issue include:

- (1) nominal curve (nominal bond) or real curve (inflation-linked bond);
- (2) asset-swap curve; and
- (3) breakeven inflation curve.

10.5.4. Different types of investor will use different valuation metrics, e.g. a bond fund manager benchmarked against the market bond index may overweight a bond if it is cheap on the curve while an asset-swap investor will judge the new bond on the basis of its swap spread. Issuers will have their own metrics to weigh financing cost against any mismatch between balance sheet assets and liabilities. The valuation metric dominant in the long end of the curve may change with time, but using the predominant metric to extend the curve will limit hedge losses after new long maturity debt issues. There is a strong theoretical case to fix the limiting long forward-rate and incorporate convexity considerations into the curve extension, but the acid test of the curve extension methodology is whether it correctly anticipates the price of new issues. We briefly examine the SA 30-year nominal and UK 50-year index-linked gilt bond issues in the last year.

10.5.5. The 18 July 2006 launch of the R209 2036 bond extended the SA nominal curve by 10 years. Bank research, e.g. RMB⁹, suggested fair value be calculated off the bond yield curve extended using constant forwards to give a yield 15 basis points below the 2026 bond. Projecting the breakeven curve and applying this to the 2033 inflation bond real yield gave a similar level, but projecting the asset-swap spread would give a nominal yield around 7 basis points lower than this. The bond was issued at a nominal yield 16 basis points below the 2026 yield, in line with the nominal curve projection. This was very close to the 14.5 basis points fair value that would be implied by the smoothed published BEASSA bond zero-coupon yield curve projected to 30 years.

10.5.6. The 22 September 2005 issue by syndication of the UK gilt 2055 extended the real curve by nearly 20 years. Bank research, such as that of RBS¹⁰ and UBS¹¹, identified swap market interest as potentially decisive since there was a significant bond-swap difference between inflation bond breakevens and inflation swap rates at the time. This resulted from strong interest in receiving long inflation swaps, counterbalanced by interest in asset-swap investors in buying long-dated index-linked gilts on asset-swap at attractive levels compared with nominal gilts. The bond was issued at a real rate below and breakeven inflation rate above that implied by the real and nominal curves and broadly in line with the projected asset-swap spread.

10.5.7. The above experiences confirmed the importance of identifying the dominant valuation metric when projecting the curve. Investor perceptions may be framed by any valuations implied by an industry-accepted curve, such as in SA. Researchers identified risk from projection off the longest bond, since there may be a realignment in value between this bond and the new issue as investors substitute holdings. Finally, despite the real yield at issue being lower than the curve extension suggested, convexity was not rewarded in the pricing of the UK 50y bond to the extent of the fair value convexity adjustment identified by researchers. This was no surprise as convexity does not appear to be fully priced into the long end of the nominal curve in this and similar developed markets.

10.6. Materiality

10.6.1. Figure 13 shows SA nominal, real and inflation curves for 26 June 2006. The nominal curve shown is the BEASSA perfect-fit curve. The breakeven

⁹ 'Interest Rate Weekly - Estimation of Issue Price for the new SA long bond (2036)', RMB Bond Analytics Research, 14 July 2006

¹⁰ 'UK 50-Year Index-Linked Gilt Launch Preview', RBS Markets Rates Strategy, 2005

¹¹ 'UBS Gilt Strategy Perspectives', UBS Research, 13 September 2005

curve uses the 4-month lag for SA inflation-linked bonds, the most recent one-year SARB CPI forecast in the short end and piecewise linear forwards between bond maturity points. Seasonality has been superimposed using the factors in Table 2, giving a jagged appearance to the curve, which becomes less prominent relative to the rate as maturities increase. The real curve is the 4-month lag curve derived from the nominal and breakeven curves.

10.6.2. A piecewise linear forward curve is in effect a perfect-fit bootstrapped curve that assumes linearly interpolated forward rates. This curve is a benchmark against which smoother perfect-fit curves can be judged. While the forward breakeven curve decreases sharply between the 2023 and 2033 inflation-linked bond maturity points, it is unlikely any other perfect-fit curve will be appreciably more smooth than this as the annual forward-rates to 2023 are smooth for this piecewise linear forward curve. The sharp decrease is the result of the 2033 inflation-linked bond being very cheap on the real and breakeven curves, although the extent of this discrepancy has been decreasing. Pension annuity liabilities will not be very sensitive to the precise curve shape of a perfect fit curve because undervaluation of one cashflow from a high curve point will most likely be counteracted by overvaluation of some other cashflow from a lower curve point required to recover dependent bond prices.

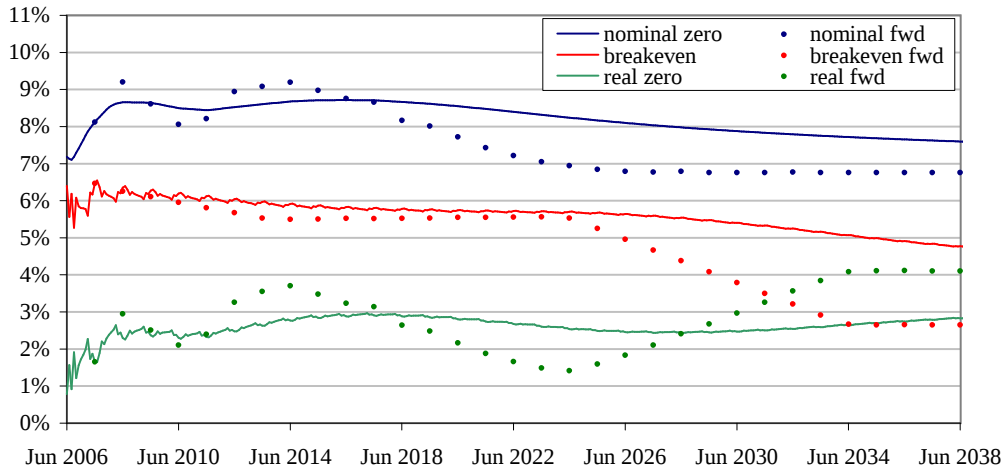


Figure 13: SA nominal, breakeven and implied 4-month lag real bond curves, 26 June 2006

10.6.3. The nominal and real curves have been extended using a constant forward-rate at the longest fitted curve point. This extension of the nominal curve was subsequently vindicated by the 2036 nominal bond issue in July

2006. Using these curves, 2.5% of the present value of the fully indexed specimen pension annuity of appendix A is in respect of cashflows with maturity greater than 30 years. The curve-extension assumptions are therefore of minor significance. Increasing the forward breakeven rate after 30 years by 1% from 2.65% to 3.65% (or equivalently decreasing the nominal curve forward rate by 1%) is equivalent to a 0.9 basis points decrease in the valuation rate. The assumptions would be more significant for a longer-term liability, such as a joint-and-survivor pension deferred for many years until retirement. Curve extension for such long liabilities should be convexity-neutral and use a fixed long rate, e.g. the method of Smith & Wilson (2000).

11. Hedge risk for linear liabilities

11.1. Risk measures

11.1.1. In this section the hedge instruments for a linear pension liability are identified, under simple assumptions about the evolution of breakeven and nominal rates over time. Let V denote the value of the portfolio of liabilities and assets held by the fund. The liabilities are projected with breakeven inflation and mortality, without allowance for background risks. Let $\underline{y} = [y_1, \dots, y_m]$ denote the curve used to value the assets and liabilities, where the y_i is the spot, forward or market rate to maturity i . Here m is the longest maturity of the liability cashflows. Over time step Δt , the change in nominal value of the portfolio as a result of time and changes in nominal rates, ΔV , has the Taylor expansion:

$$\Delta V = \frac{\partial V}{\partial t} \Delta t + \sum_{i=1}^m \frac{\partial V}{\partial y_i} \Delta y_i + \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \frac{\partial^2 V}{\partial y_i \partial y_j} \Delta y_i \Delta y_j + \dots \quad (8)$$

There are no stochastic differential terms as we assume the evolution of the curve is deterministic. We ignore higher order terms in Δy as they are likely to be small for small Δt . We also ignore cross terms in Δy and Δt but note that the risk measures below will change in time unless the portfolio is well hedged. The change in nominal value of the portfolio can then be approximated using nominal risk measures Θ^n , Δ^n and Γ^n (the ‘‘Greeks’’) as

$$\Delta V \approx \Theta^n \Delta t + \sum_{i=1}^m \Delta_i^n \Delta y_i + \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \Gamma_{ij}^n \Delta y_i \Delta y_j$$

where $\Theta^n = \frac{\partial V}{\partial t}$, $\Delta_i^n = \frac{\partial V}{\partial y_i}$ and $\Gamma_{ij}^n = \frac{\partial^2 V}{\partial y_i \partial y_j}$. (9)

11.1.2. The portfolio assets may be physical, derivative or both. Our approach is to value the portfolio at a margin to the swap curve, as discussed above in section 5.1. For consistent hedge risk measurement of pension liabilities and bond holdings in the portfolio, bonds should be valued using the bond curve which is defined in terms of the swap curve and an additional variable for the bond-swap basis.

11.1.3. Inflation-linked liabilities are also exposed to the risk of changes in the forward inflation indices $\underline{Q} = [Q_1, \dots, Q_k]$ used to index up the portfolio cashflows up from a common base index Q_0 . Here k is the number of forward CPI dates used to index the portfolio assets and liabilities. If these additional risks are, for now, assumed orthogonal to nominal rate risks:

$$\Delta V \approx \Theta^n \Delta t + \sum_{i=1}^m \Delta_i^n \Delta y_i + \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \Gamma_{ij}^n \Delta y_i \Delta y_j + \Theta^Q \Delta t + \sum_{i=1}^k \Delta_i^Q \Delta Q_i$$

where $\Theta^n = \frac{\partial V}{\partial t}$ and $\Delta_i^Q = \frac{\partial V}{\partial Q_i}$. (10)

Since risk is linear with respect to the forward inflation indices, there is no Γ^Q convexity term. However, if risk is measured with respect to a breakeven inflation spot rate curve, $\underline{b}$ with $b_i = \frac{\ln(Q_i/Q_0)}{i}$, then breakeven convexity would require a gamma term.

11.2. Theta

Theta is the change in nominal value of the portfolio for no change in rates. Theta is deterministic, so is not strictly a risk measure. The condition of no change in rates is ambiguous without specifying rates as forward or spot rates (Flavell 2002). As portfolio assets and liabilities can be financed and cashflows reinvested at the forward curve rates over Δt , the assumption that forward-rates between each fixed future maturity date are unchanged is consistent with no arbitrage, since no profit or loss is expected to emerge (we ignore the profit or loss on decay of convexity profit or loss discussed in section 10.3.8). This is true even if leveraged positions are held, such as swaps or bonds financed on repo. If the portfolio is perfectly hedged theta will be zero regardless of which of the two calculations is used.

11.3. Delta and duration

11.3.1. If there is no delta risk with respect to nominal rates and inflation at any cashflow date, the portfolio is hedged with respect to nominal rate risk. This is the strategy of ‘dedication’ in actuarial practice and ‘defeasance’ in financial theory.

11.3.2. Flavell (2002) shows how delta risk at each curve grid-point can be calculated with respect to spot, forward or market rates:

- Spot rate delta at grid-point i results from moving the spot curve at this point up by a small rate move (or ‘bump’). That is, the sensitivity to the spot rate for grid-point i corresponding to maturity T_i , $S(0, T_i)$, is calculated by increasing this spot rate to $S(0, T_i) + 0.01\%$ and revaluing the liability.
- Delta with respect to the forward-rate between $i-1$ and i is found by bumping up the forward curve between these points. For forward-rates after curve grid-point i to remain constant spot rates after i must be adjusted. Consider the forward-rate between the maturities T_{i-1} and T_i corresponding to the grid-points $(i-1)$ and i respectively. If this forward-rate is denoted by $F(T_{i-1}, T_i)$ then:

$$F(T_{i-1}, T_i) = \left[\frac{(1 + S(0, T_i))^{T_i}}{(1 + S(0, T_{i-1}))^{T_{i-1}}} \right]^{1/(T_i - T_{i-1})} - 1$$

A bump in the forward-rate curve is therefore consistent with a different movement in the spot curve than a spot curve bump. However, the sum of all sensitivities with respect to forward-rate bumps and spot rate bumps across all grid-points will be approximately equal.

- The delta risk with respect to a market rate for an instrument with maturity i used to build a valuation curve gives the hedge position holding for that instrument as a useful by-product. This delta is found by adjusting the instrument price up by an amount corresponding to a 0.01% increase in the market rate and rebuilding the curve to recalibrate to the adjusted price. The change in market rate will imply a certain change in spot and forward-rates, dictated by the form of the (perfect fit) curve used. Curve methodologies which give high priority to smoothness will tend to dissipate the curve bump as far as possible to retain smooth forwards. A sparsely calibrated parametric curve with a complex form may respond to up and down bumps with unpredictable, and possibly asymmetric, forward-rate changes in distant segments of the curve. Consider a curve built from swap rates. Denote the swap rate for maturity T_i by $S(0, T_i)$. Flavell (2002) shows that the swap rate can be expressed in terms of the spot rates, ignoring daycount adjustments, as:

$$S(0, T_i) = \frac{1 - P^n(0, T_i)}{\sum_{k=1}^i P^n(0, T_k)}.$$

Since

$$P^n(0, T_i) = \exp(-T_i S(0, T_i)),$$

a bump in the swap rate $S(0, T_i)$ is not equivalent to a bump to the spot rate $S(0, T_i)$. Again, the sum of sensitivities to all spot and market rates for all grid-points will be approximately equal.

11.3.3. These three methods will give very similar delta sensitivities if the bump is small enough for convexity effects to be negligible. To reduce the convexity or gamma effect, a bump of less than one basis point may be used and the sensitivities calculated by bumping rates up and down and averaging the effects. The number of grid-points may be reduced by mapping cashflows to buckets using an algorithm to preserve present value and risk characteristics or the portfolio may be expressed in terms of equivalent holdings of key hedge instruments (e.g. 2, 5, 10, 15, 20 and 30-year swaps). This is helpful for dynamic risk management of complex derivative portfolios where risk may change dramatically across maturities, but less critical for a regular series of cashflows (such as a pension liability with indexation accruing from valuation date) as the risk profile of the hedge portfolio will be more regular.

11.3.4. Delta risk may be summed across grid-points to give a total delta, Δ^n , often expressed for a one basis point move in rates and termed the ‘pvbp’ or ‘dv01’. If the curve is expressed in continuously compounded rates, an asset A with delta $\Delta^n(A)$ and value V^A has duration

$$-\frac{\Delta^n(A)}{V^A}. \quad (11)$$

Duration is unitless and positive signed. For liability L , where the delta is already positive,

$$\frac{\Delta^n(L)}{V^L}. \quad (12)$$

Duration as defined above is not meaningful for a swap without decomposing the swap asset and liability legs into equivalent notional flows and adding these to the other assets and liabilities. The duration of a portfolio can be found via its additive deltas or as the discounted mean term of future cashflows. Note that the above measures all implicitly assume equal rate moves with perfect positive correlation (i.e. a parallel curve shift) since they sum across the partial durations for each curve point.

11.4. Gamma and convexity

Gamma will be material for long-dated liabilities, such as pensions, but there are $\frac{n(n+1)}{2}$ distinct partial gammas with respect to a curve with n maturity points (Hull 2003). Again, the assumption of a parallel curve shift

allows the partial gamma measures to be summed to give the total gamma, Γ^n . Gamma is positive for an asset or liability, due to the convex shape of the discount function, allowing the unitless convexity measure for an asset to be defined for a continuously compounded rate curve as

$$\frac{\Gamma^n(A)}{V^A}. \quad (13)$$

The convexity measure for the liability is

$$\frac{\Gamma^n(L)}{V^L}. \quad (14)$$

11.5. Immunisation

11.5.1. The actuarial theory of immunisation was developed by Redington (1952) for the hedging of life insurance liabilities exposed to nominal rate risk. Redington showed that a portfolio of assets and liabilities is ‘immunised’ against losses under the paradigm of sufficiently small parallel curve shifts, if the following conditions hold:

- (1) present value of the assets equals present value of the liabilities;
- (2) duration of the assets equals duration of the liabilities; and
- (3) convexity of the assets is greater than the convexity of the liabilities.

This result follows directly from equation (8) with the assumption of zero portfolio theta over the period of the curve shift, since higher-order terms will be negligible for sufficiently small parallel yield-curve shifts $\varepsilon (= \Delta y_i \forall i)$. The theory originally assumed a flat yield-curve, but this was later generalised by Fisher & Weil (1971), see Maitland (2001). The theory may be extended to inflation-linked liabilities if the breakeven inflation rate used to project the liabilities is suitably well behaved, e.g. if assets and liabilities have the same inflation delta and parallel shifts in the breakeven curve are of magnitude less than ε .

11.5.2. Maitland (2001) cautions that the assumptions underpinning the theory are too strong in practice. Immunisation theory assumes a sufficiently small parallel yield-curve shift but figure 14 shows that a larger shift may result in a hedge loss. This, in itself, does not invalidate the theory since the assets could be structured to satisfy enough higher-order risk measure conditions of the form

$$\frac{\partial^m A}{\partial \Delta y^m} > \frac{\partial^m L}{\partial \Delta y^m}; \quad m \geq 2.$$

This ensures these higher-order terms in equation (8) do not cause losses for curve shifts of larger magnitude. It is assumed that sufficiently convex assets are available to the fund. Interest- and inflation-rate derivatives can provide

this convexity, however there will be a cost imposed by this increased convexity if the yields are lower on assets with high convexity. This cost is evident in the SA bond market where the yield-curve is inverted for long-term maturities.

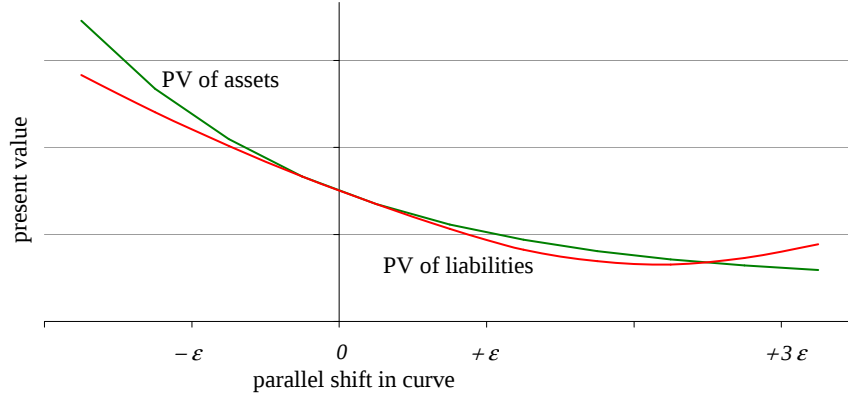


Figure 14: Example of immunisation failing for a $+3\varepsilon$ parallel curve shift

11.5.3. Alternatively, the hedge can be rebalanced after shifts of size ε , so the assumption of a perfect market can safeguard the assumption of sufficiently small curve shifts.

11.5.4. If the curve is steeply inverted and the cost of convexity is high, it is not obvious that the benefits of convexity outweigh the costs. Moreover, if there is no curve volatility there may be negative theta in equation (8) as convexity decays with time without a commensurate gamma benefit. The market may also reappraise the price of convexity – for instance if an inverted curve normalises.

11.5.5. This points to the real weakness of immunisation theory, that it is “theoretically deficient, being riddled with arbitrage” (Smith in the discussion of Feldman et al. (1998)). To illustrate how immunisation is incompatible with the absence of arbitrage we consider the simple model for the real-world evolution of the forward-rate, $f(t, T)$:

$$df(t, T) = \sigma dW_t + \alpha(t, T)dt \quad (15)$$

This model is a single factor model of the form considered by Heath, Jarrow & Morton (1992) (HJM).

$$\alpha(t, T) = \sigma^2(T-t) + \sigma\gamma_t$$

where γ_t is the market price of risk, which is only a function of t .

Since γ_t is independent of T , it is not possible for $\alpha(t, T)$ to be always constant. Therefore exclusively parallel curve shifts cannot exist in the arbitrage-free, risk-neutral world.

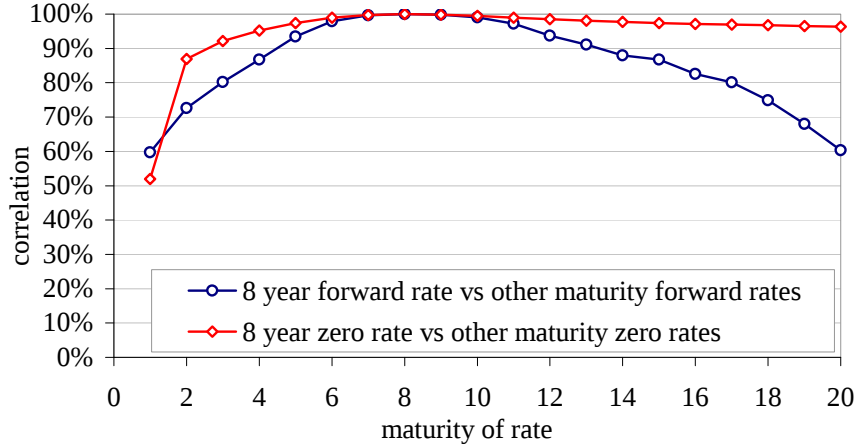


Figure 15: Correlation of SA nominal bond curve 8-year rate with other maturity rates for June 2004 to June 2006

11.5.6. There is also substantial empirical evidence against parallel curve shifts. Figure 15 shows imperfect correlation between monthly absolute changes in the 8y and other maturity rates for BEASSA published bond curve zero and (annual) forward-rates between 26 June 2004 and 26 June 2006.

11.5.7. Flavell (2002) describes measures of curve risk which account for imperfect correlation between rate moves by constructing hedges against pre-defined curve movements, such as changes in curve slope. Golub & Tilman (2000) describe how ‘key rate durations’ can be defined for each instrument in a portfolio which capture the sensitivity of the instrument to a limited number of key maturity points on the yield-curve.

11.6. Principal Components Analysis

11.6.1. Principal Components Analysis (PCA) can be used to define the characteristic curve shifts used by yield-curve risk-management methods. Figure 15 can be extended to all annual curve points to produce the correlation matrix of rate movements. Intermediate points can be interpolated to form a correlation surface to depict this correlation matrix, as shown in Figure 17. Let $\mathbf{R}$ denote this empirical correlation matrix. The singular value decomposition of a correlation matrix will always exist (McNeil, Frey & Embrechts 2006) and can be denoted by $\mathbf{R} = \mathbf{E}\mathbf{D}\mathbf{E}^\top$ where $\mathbf{E}$ is the orthonormal matrix with the eigenvectors (the ‘factors’) of $\mathbf{R}$ as columns and $\mathbf{D}$ is the diagonal matrix of eigenvalues corresponding to these eigenvectors, ranked in decreasing order

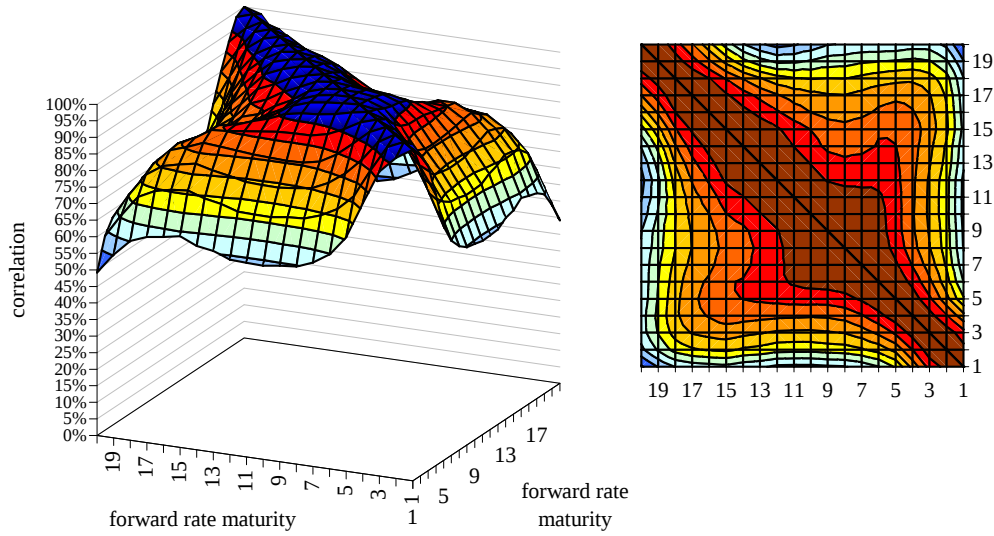


Figure 16: SA nominal bond forward rate correlation matrix for monthly absolute changes in rates from June 2004 to June 2006

of magnitude. PCA reduces the dimension of the curve movements from the number of curve points to the minimum number of components required to give an adequate representation of curve movements (Rebonato 1998).

11.6.2. Table 3 shows the results of the PCA applied to the correlation matrix of Figure 16. The first factor accounts for 83.6% of the total variance of forward curve movements. The second and third factors explain a further 7.3% and 5.8% of the total variance, so 96.6% of total variance is captured by the first three components.

factor number, i	1	2	3	4	5	6	7
eigenvalue, λ_i	16.71	1.46	1.16	0.47	0.16	0.020	0.015
variability explained, $\lambda_i/\Sigma\lambda_i$	83.6%	7.3%	5.8%	2.3%	0.8%	0.1%	0.1%
cumulative variability explained	83.6%	90.8%	96.6%	98.9%	99.8%	99.9%	99.94%

Table 3: First seven eigenvalues for the PCA of the correlation matrix of SA nominal bond forward-rates depicted in Figure 16

11.6.3. Figure 17 shows plots of the factor loadings on each curve point for the first three component factors of the PCA in Table 3. PCA of forward-rates is useful for calibrating many interest rate models Rebonato (2002) but PCA may also be applied to changes in spot (i.e. zero-coupon) rates. Table 4 and Figure 18 show the results of this alternative analysis, which are very similar to those of Maitland(2001, 2002) for the covariance matrix of par rate moves in the JSE-Actuaries curve for monthly data between 1 February 1986 and 1 May

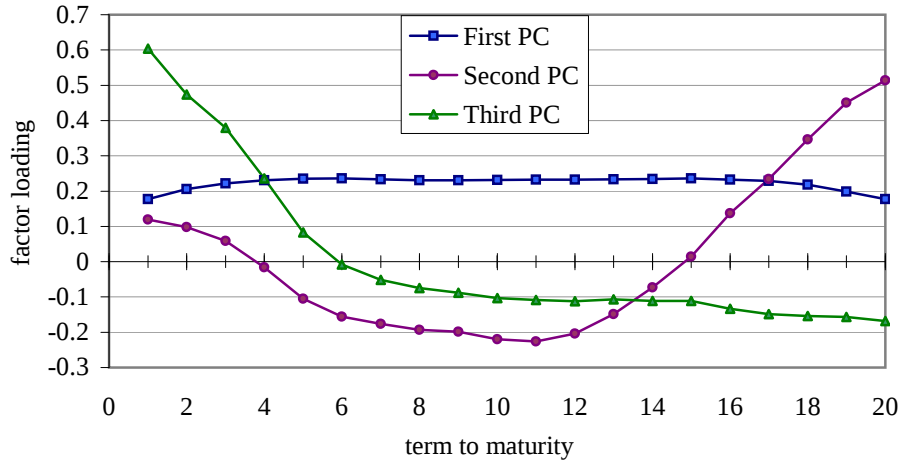


Figure 17: Factor loadings for first three principal components for the PCA using SA nominal bond forward-rates in Table 3

2000. We see that the first, second and third spot rate factors account for 99.6% of total variance and the loading profiles correspond to the parallel shift, slope and shape moves characteristic of many interest rate markets (Rebonato 2002).

factor number, i	1	2	3	4	5	6	7
eigenvalue, λ_i	18.27	1.45	0.20	0.05	0.02	0.005	0.0004
variability explained, $\lambda_i/\Sigma\lambda_i$	91.4%	7.3%	1.0%	0.2%	0.1%	0.03%	0.002%
cumulative variability explained	91.4%	98.6%	99.6%	99.9%	100.0%	100.0%	100.0%

Table 4: Factor loadings for first three principal components of the correlation matrix of spot rates for SA nominal bonds, June 2004 to June 2006 monthly rate moves

11.6.4. Forward-rates in the long end of the curve are less stable than spot rates (which are an average of forward-rates up to the spot rate maturity). Therefore it is not surprising that a PCA of spot rates explains more of the curve variation, even with the smoothing of forward-rates imposed by the BE-ASSA methodology and the sparse SA bond curve. The second and third eigenvectors of the forward-rate PCA have eigenvalues of similar magnitude, so one can reconcile the components of the two analyses.

11.6.5. Table 5 and Figure 19 show a PCA of SA bond breakeven inflation rate changes. The analysis uses spot rates since breakeven curve shapes may be less stable than nominal curves for the reasons discussed in section 5.4. The SA real curve has also shown significant realignment over the period of the data and the PCA is dominated by one outlier shift. Parallel shifts dominate, as

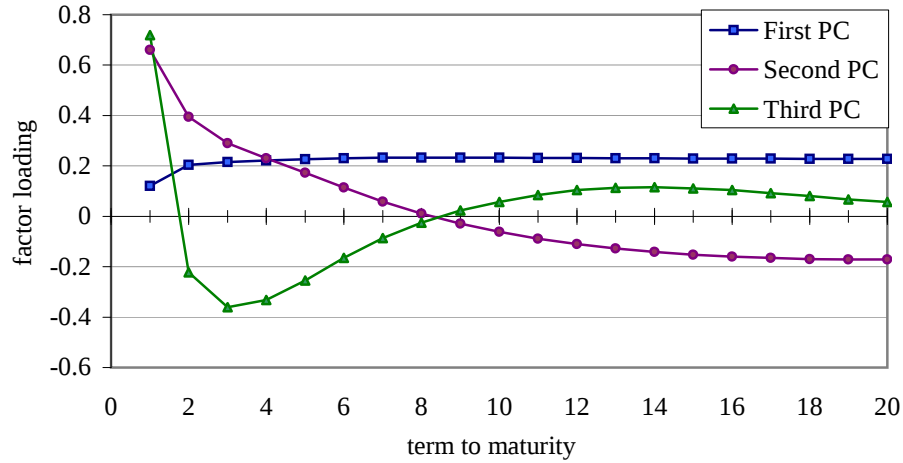


Figure 18: Factor loadings for first three principal components for SA nominal bond rate PCA using spot rates of Table 4

they do for nominal rates, but the first component has less explanatory power.

<i>factor number, i</i>	1	2	3	4	5	6	7
eigenvalue, λ_i	14.35	4.26	0.78	0.32	0.23	0.022	0.010
variability explained, $\lambda_i/\Sigma\lambda_i$	71.7%	21.3%	3.9%	1.6%	1.2%	0.1%	0.1%
cum. variability explained	71.7%	93.1%	97.0%	98.6%	99.7%	99.9%	99.91%

Table 5: Factor loadings for first three principal components of the correlation matrix of spot rate changes for the SA bond breakeven curve, June 2004 to June 2006 monthly rate moves

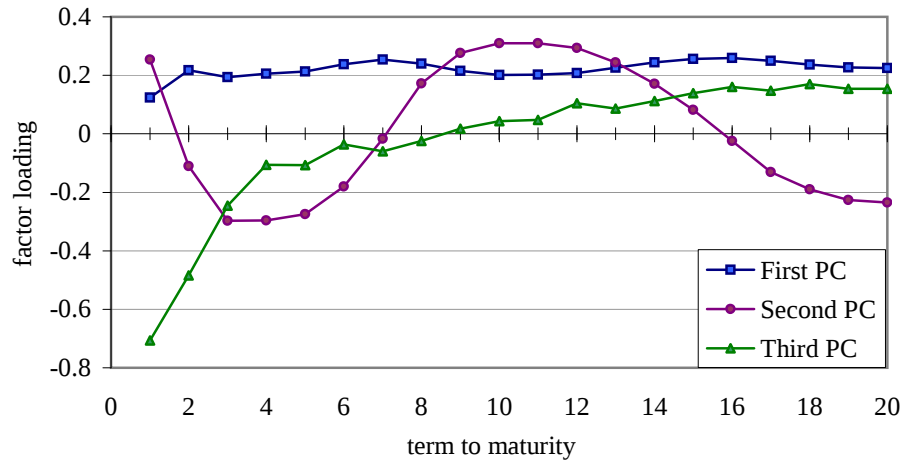


Figure 19: Factor loadings for first three principal components for the PCA using SA breakeven spot rates of Table 5

11.6.6. Risk management with respect to moves in nominal and breakeven rates can use PCA to identify the representative curve shift against which the portfolio is to be hedged. Since the factors are orthogonal by construction, the total portfolio delta for a bump of size ε (e.g. 1bp) is

$$\Delta V = \sum_{i=1}^m \sum_{j=1}^m \frac{\partial V}{\partial y_i} \lambda_i P_i(j) \varepsilon$$

where $P_i(j)$ denotes the loading of the i -th factor on the j -th curve point.

11.6.7. PCA can be also performed on the covariance matrix of rate moves. Then, if it captures the majority of the variation, the profile of the first factor's loadings will reflect the historical term structure of volatility of the rates for the time period between rate moves in the data. However, this scales the curve shifts at each point to its historical standard deviation rather than the small ε bump size required for estimating portfolio delta risk with respect to yield-curve moves.

11.6.8. Maitland (2001) analysed the hedge portfolio that immunised an SA interest rate liability against PCA-derived yield-curve shifts. The hedge portfolio was optimised to maximise expected immunisation profits. It was shown in section 11.5.5 that a model designed in this way is not arbitrage-free, so the market value of the hedge cannot be used as a market value for the liability if the market is assumed to be complete.

11.6.9. Tilley & Mueller (1991) also consider immunisation free of the assumption of a parallel shift in rates. They use orthogonal polynomials rather than principal components to simulate curve shifts, again without constraining the curve dynamics to be free of arbitrage. The first three orthogonal polynomials correspond to parallel shift, slope and shape moves.

11.6.10. PCA-based risk methods for interest rate risk management are dependent on past curve data for calibration. The correlation matrix may be distorted by outliers, the period of the time series may not be representative of the curve dynamics or the dynamics may change in future. PCA on a the correlation matrix of absolute rate moves ignores the correlation between the level of rates and changes in curve shape. For example, there is evidence of 'bull steepening' where increases in rates are associated with a steepening of the curve slope (Golub & Tilman 2000). In the next section it is shown that optimal hedging strategies require a dynamic interest and inflation rate model that is free of arbitrage, calibrated to market rates and incorporates the key statistical features of the processes.

12. CPI revisions

12.0.1. There are two forms of revisions to the CPI index by the statistical agency responsible for calculation. The first type of revision, e.g. for Eurozone inflation, is a regular update from the ‘flash’ to the ‘final’ CPI estimate to reflect more accurate price-level information obtained since the time of the survey. Inflation linked assets and liabilities will reference the final index and the index lag will be sufficient to accommodate final index publication.

12.0.2. The second type of revision is restatement of past indices due to a material calculation error. Several inflation markets have experienced revisions to the CPI following calculation errors. The US Bureau of Labor Statistics revised January 2000 to August 2000 CPI-U data by up to 0.2 index points, or 0.12% in year-on-year terms¹². In May 2003 Statistics South Africa revised the CPI published from February 2001 to March 2002. The cumulative effect of the revision was a 2.3% reduction in headline CPI. In the same year a revision took place in the Netherlands with an effect on Eurozone inflation.

12.0.3. It is generally not practical for the national debt manager to neutralise the past effect of index restatement since this would entail adjusting the price level of all affected inflation-linked bond transactions. Future transactions and cashflows will be based on the corrected, published CPI index. Unless asset and liability indexation and cashflows are coincident, revisions may create mismatches between inflation-linked assets and inflation-linked pension liabilities. The market has taken steps to eliminate mismatches in indexation between inflation-linked bonds and derivatives through the International Swaps Derivatives Association (ISDA) definitions used in inflation derivatives documentation¹³. The ISDA definitions align the indexation of inflation derivatives with inflation-linked bonds.

12.0.4. Revisions are discontinuities in the inflation process and may result in hedging losses. This has not been explicitly addressed in the literature. Inflation revisions are similar to jumps in asset prices, which have been addressed in some models, but the irregular and legal effects of revisions suggests that they are closer in nature to operational risks.

¹² US Department of Labor, Bureau of Labor Statistics, ‘Revisions in January to August 2000 CPI data’, www.bls.gov/cpi/cpirev01.htm

¹³ ISDA Inflation Derivatives Definitions, 2006. See <http://www.isda.org/publications/pdf/2006inflationderdefs.pdf>

13. Market completeness

13.0.1. Even linear liabilities assumed free of background risks are an incomplete market as a result of:

- an indexation lag not aligned with the traded instruments
- inflation seasonality
- liabilities extending beyond the term of the markets

13.0.2. However, this incompleteness is not a significant factor for the market price of inflation-linked pensions in payment. For such liabilities, it has been shown that seasonality and the method used to model and extend the yield curve do not have a material effect on the price of a typical pension profile. Inflation swaps dealers in effect complete the market for these liabilities by offering hedges which precisely match the interest and inflation rate risk of these liabilities.

CHAPTER 5

Non-Linear inflation liabilities

In this chapter the market valuation of pension liabilities is extended to include pensions imperfectly indexed with inflation. The various forms of limited inflation indexation of pensions are classified and analysed as inflation-derivatives embedded within the pension liabilities.

An overview of the theory of derivatives pricing is provided and the assumptions underpinning the Black–Scholes model are considered. The Black–Scholes formula and the currency analogy for the Consumer Price Index are used to derive closed-form valuation formulae for the various forms of inflation indexation proposed in the actuarial literature.

The Black–Scholes-world assumptions are dissected in order to critically appraise whether these pricing methods can be applied to value pension liabilities as inflation derivatives. The assumptions are tested against data for the South African inflation market and evidence of mean reversion and fat tails for the inflation process is considered. The published inflation derivatives pricing models are then analysed in the light of these findings.

It is motivated why the Black–Scholes valuation approach using volatilities adjusted to account for unrealistic assumptions is a useful basis for market valuations. However, the non linear nature of most inflation-linked pension liabilities necessitates a stochastic model for interest rates and inflation to capture subtle convexity and path-dependent effects.

1. Forms of pension indexation

1.1. Introduction

1.1.1. Chapter 3 discussed the nature of pension indexation to inflation. In many countries the level of the indexed pension is usually not directly proportional to the level of the inflation index. This limited price indexation (LPI) gives rise to a non-linear market valuation since it is not linear in real and nominal zero-coupon bond prices.

1.1.2. Pension increases can be limited in relation to the Consumer Price Index (CPI) in a number of ways. The limit may apply to the increase in the index or to the level of the index. Building on the taxonomy of Wilkie (1984) and Van Bezooyen et al. (1997), five types of CPI indexation encountered in practice can be identified:

- type-1. Unlimited indexation. The pension increase is the percentage increase or decrease in the CPI since the previous indexation date, usually 12 months prior. This produces linear pension liabilities.
- type-2. Inflation index caps and floors. The pension increases in line with the inflation index, but the pension is subject to a minimum floor level of average annual increase in the index and maximum cap level of average annual increase in the index. The legislated minimum increase for deferred pensions in the UK is of this form.
- type-3. Inflation rate caps and floors with claw back. The pension increases with type-1 or type-2 indexation, but the pension is also floored at its current level. This increase is in the trust deeds of some UK occupational pension funds (Van Bezooyen et al. 1997).
- type-4. Inflation rate caps and floors. The pension is adjusted by the inflation rate since the last increase, but the rate of pension increase is subject to a minimum floor level and a maximum cap level. An example is 0% floor and 2.5% cap LPI for UK pensions in payment accrued post 6 April 2005.
- type-5. Fractional indexation. The pension increase is a fixed fraction of the inflation rate since the last adjustment.

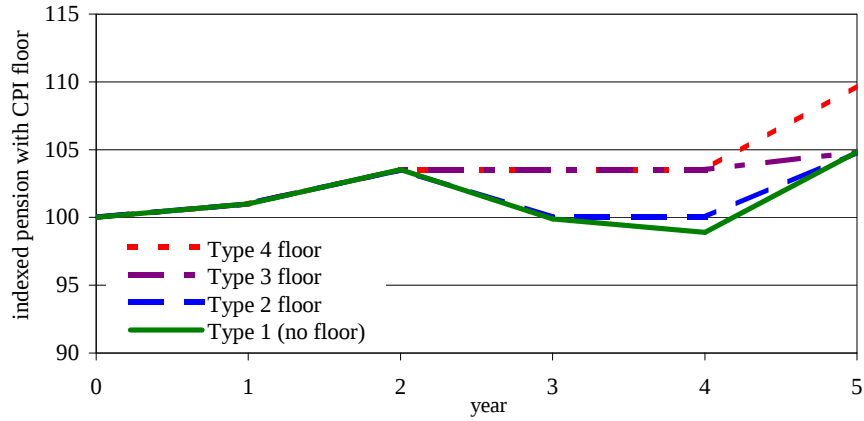
Table 6 and Figure 20 use the case of a floor at 0% p.a. to illustrate the difference between the floor types for a contrived CPI time series.

1.2. Relationship between the different forms of indexation

1.2.1. The relationship between the level of the pension benefit under each type of indexation floor struck at the same level is:

$$\text{type-1} < \text{type-2} < \text{type-3} < \text{type-4}.$$

t	Type-1		Type-2, 0% floor		Type-3, 0% floor		Type-4, 0% floor		Type-5, 75% increase	
	CPI _t	increase	pension	increase	pension	increase	pension	increase	pension	increase
0	100.00		100.00		100.00		100.00		100.00	
1	101.00	1.0%	101.00	1.0%	101.00	1.0%	101.00	1.0%	100.75	0.75%
2	103.53	2.5%	103.53	2.5%	103.53	2.5%	103.53	2.5%	102.64	1.88%
3	99.90	-3.5%	100.00	-3.4%	103.53	0.0%	103.53	0.0%	99.94	-2.63%
4	98.90	-1.0%	100.00	0.0%	103.53	0.0%	103.53	0.0%	99.20	-0.75%
5	104.84	6.0%	104.84	4.8%	104.84	1.3%	109.74	6.0%	103.66	4.5%

Table 6: Calculation of pension increases for different forms of indexation**Figure 20:** Indexed pension payments for the different forms of indexation

1.2.2. These types of limited CPI indexation may occur in combination. In SA it is common to have a type-4 floor at 0% and no cap, combined with a type-5 fractional increase (of say 75% of CPI).

2. Inflation derivatives and pension indexation

2.0.1. Advanced inflation markets trade derivatives with features similar to some forms of pension liabilities.

- Zero-coupon inflation caps (floors) are options on the CPI. The buyer of a floor struck at k pays a premium upfront and receives a payoff per unit of notional at time T of

$$\max[CPI_T/CPI_0 - (1+k), 0].$$

- Year-on-year inflation caps (floors) are options on the annual inflation rate. The buyer of a T -year floorlet struck at $k\%$ pays a premium upfront and receives a payoff per unit of notional at time T of

$$\max[CPI_T/CPI_{T-1} - (1+k), 0].$$

The notional is fixed in nominal terms for each caplet or floorlet. One-year zero-coupon and year-on-year inflation options are equivalent.

- LPI swaps exchange a compounded fixed rate swap leg for a type-4 LPI swap leg. On maturity at time T , for a swap with unit notional the payer of the fixed leg dealt at rate f pays $(1+f)^T$ and receives the growth in the type-4 LPI index from time 0 to time T . UK LPI swaps collared between 0% and 5% are reasonably liquid, trading several times a week in multiples of the £25 million standard notional size.

2.0.2. A type-2 collared LPI liability is equivalent to a combination of type-1 LPI liability, a long position in a zero-coupon inflation floor and a short position in zero-coupon inflation cap. A type-4 collared LPI liability is the compounding of type-1 inflation increase with a one-year zero coupon collar. A type-3 LPI liability is a lookback option on the inflation index (Van Bezooen et al. 1997).

2.0.3. Inflation-linked bonds in many markets (e.g. SA, USA and Euro zone) have a deflation floor on the principal. A zero-coupon inflation floor struck at the bond's base index is embedded in the bond. Accrued inflation will cause the strike to move further from the money. For example, if the SA inflation-linked bond issued on 2 May 2002 with a base index of 109.03871 and maturing on 31 March 2008 is purchased for settlement on 31 March 2006 when the reference inflation index is 129.5, then the floor has an effective annual inflation rate strike of -8.98% (being $(109.03871/129.5)^{1/2} - 1$ and ignoring any subsequent published inflation figures).

3. The Black–Scholes world

3.1. Notation

In this section we consider the basic Black–Scholes model for derivatives on the inflation index. Q_t follows exponential Brownian motion with real-world probability measure $\mathbb{P}$ on the σ -algebra $\mathcal{F}$ of subsets of state space $\mathbb{R}^+$,

$$\begin{aligned} dQ_t &= Q_t(\sigma_t dW_t + \mu_t dt) \quad \text{where } W_s \text{ is a Brownian motion under } \mathbb{P} \\ \Leftrightarrow \ln Q_t &= \ln Q_0 + \int_0^t \sigma_s dW_s + \int_0^t \mu_s ds \end{aligned} \tag{16}$$

3.2. Conditions for claim replication

3.2.1. Black & Scholes (1973) and Merton (1973) showed how certain claims on an underlying, such as European options, can be hedged under certain conditions using a dynamic portfolio of tradable assets. This replicating portfolio

is self-financing and has the same payoff as the claim. For there to be no arbitrage, the LOOP dictates that the (unique) price of the claim is the price of the replicating portfolio. The Black–Scholes-world assumptions are:

- (1) perfect markets, where positions of unlimited size can be bought or sold instantaneously without incurring transaction costs, crossing bid–offer spreads or market impact;
- (2) the stochastic price process for the underlying is exponential Brownian motion, resulting in normal returns and lognormal prices;
- (3) interest rates are deterministic and so unrelated to the price process for the underlying; and
- (4) the process for the underlying has deterministic volatility.

3.2.2. The assumptions are patently unrealistic, but are nonetheless the basis for pricing in inflation and other options markets. The assumptions have been gradually weakened in the literature and experience has shown the Black–Scholes model, duly modified, to be robust to departures from the assumptions.

4. Derivatives pricing theory

4.1. Black–Scholes partial differential equation

4.1.1. Consider a derivative V_t which has a payoff only at T which is a sufficiently regular function p of Q_t . In the Black–Scholes world the price of V satisfies the following partial differential equation (PDE)

$$\frac{1}{2}\sigma^2Q^2\frac{\partial^2V}{\partial Q^2} + (f - g)Q\frac{\partial V}{\partial Q} + \frac{\partial V}{\partial t} = fV$$

where f and g are the deterministic instantaneous nominal and real rates.

4.1.2. This PDE can be expressed in terms of the Greeks defined in section 11 of chapter 4 as

$$\frac{1}{2}\sigma^2Q^2\Gamma + (f - g)Q\Delta + \Theta = fV. \quad (17)$$

The boundary condition $V_T = p(Q_T)$ produces the process for V_t from this PDE, allowing the derivation of an analytic derivative price (Wilmott 1999). The PDE may be adapted for path-dependent options (such as type-3 LPI).

4.2. Martingale Pricing Theory

4.2.1. Harrison & Kreps (1979) showed that the price of a derivative in the Black–Scholes world is its expectation using the probability measure $\mathbb{Q}$ for which the discounted price process of the underlying tradable is a martingale. In this idealised world, the Girsanov Theorem implies that the change of measure is equivalent to a change in drift of the driving Brownian motion

(see Musiela & Rutkowski (1997)). For the discounted price of the underlying to be a martingale it must be driftless and therefore the derivative payoff is independent of the real-world drift μ_t . The term market price of risk of the underlying under $\mathbb{Q}$ is zero and so $\mathbb{Q}$ is called the risk-neutral measure.

4.2.2. The term risk-neutral can be misleading as martingale pricing theory does not draw on any assumptions about the utility functions of market participants other than to assume they prefer more wealth to less so that the markets they invest in become arbitrage free. The resulting derivative price is preference independent. The risk-neutral world is a parallel economic universe which only exists to facilitate the valuation of derivatives.

4.3. Deflators

The separation of process and measure is a key principle differentiating the financial approach from the traditional actuarial approach of the best-estimate valuation as the real-world expected value. Modern actuarial theory (see, for example, Jarvis et al. (2001)) often uses deflator processes to produce market-consistent valuations as the real-world expected value of the payoff adjusted by the deflator.

5. Black-Scholes currency model for inflation

5.1. Background

Van Bezooyen, Exley & Smith (1997) used a simple Black-Scholes currency model of Garman & Kohlhagen (1983) to price type-2 and type-4 liabilities. They showed that the model performed well in an empirical investigation. The authors found that the theoretical price of an LPI liability (with a collar between 0% and 5%) was within 5% of the cost of the replicating portfolio using monthly rebalancing of the portfolio at mid-market rates for nominal and inflation-linked bonds. This hedge performance improved to a 3% replication error when the model was generalised to allow for interest rate variation. The investigation used UK market data for the period 1986 to 1995. The volatility assumption used to price the derivative was the actual volatility over the period, since there was no LPI derivative market-implied volatility at the time.

5.2. Type-2 LPI

5.2.1. We first consider the case where the indexation is coincident with the payoff date (i.e. there is no inflation indexation lag). The payoff at time T for a 100k% type-2 cap that incepted at time 0 when the base index $Q_0 = 1$ is

$$\max[0, Q_T - (1 + k)^T].$$

If Q_t follows exponential Brownian motion and the assumptions of section 3.2 hold, Van Bezooyen, Exley & Smith (1997) show the value at time t of this type-2 cap with maturity T is,

$$c_2(t, T, k) = Q_t P^r(t, T) \Phi(d_1) - (1+k)^T P^n(t, T) \Phi(d_2), \quad (18)$$

where Φ denotes the standard normal cumulative distribution function,

$$d_1 = \frac{\ln(Q_t P^r(t, T)) - \ln((1+k)^T P^n(t, T)) + \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}},$$

$$d_2 = d_1 - \sigma\sqrt{T-t}$$

and $\sigma = \frac{1}{\sqrt{T-t}} \left(\int_0^t \sigma_s^2 ds \right)^{1/2}$ is the root-mean-square volatility of $\ln Q_t$.

Type-2 floors at 100k%, denoted $f_2(t, T, k)$, can be valued via put-call parity:

$$\begin{aligned} Q_t P^r(t, T) + f_2(t, T, k) &= c_2(t, T, k) + (1+k)^T P^n(t, T) \\ \Rightarrow f_2(t, T, k) &= (1+k)^T P^n(t, T) \Phi(-d_2) - Q_t P^r(t, T) \Phi(-d_1). \end{aligned} \quad (19)$$

5.2.2. The notional amounts of the fixed and inflation-linked zero bonds required to hedge the options are the coefficients of $P^n(t, T)$ and $P^r(t, T)$ in (18) and (19). For example, at time t the hedge for a short position in an LPI type-2 cap with unit notional and maturity T is a long position in zero-coupon inflation-linked bond with notional $Q_t \Phi(d_1)$ and a short position in a zero-coupon nominal bond with notional $(1+k)^T \Phi(d_2)$.

5.2.3. The delta and gamma Greeks for type-2 LPI options with respect to the nominal and inflation-linked zero-coupon bond prices are readily obtained by differentiating equations 18 and 19 to give:

$$\begin{aligned} \text{For caps:} \quad \Delta_{P^r} &= \Phi(d_1) & \Delta_{P^n} &= -(1+k)^T \Phi(d_2) \\ \Gamma_{P^r} &= \frac{\Phi'(d_1)}{P^r(t, T) \sigma \sqrt{T-t}} & \Gamma_{P^n} &= -\frac{\Phi'(d_2)}{P^n(t, T) \sigma \sqrt{T-t}} \\ \text{For floors:} \quad \Delta_{P^r} &= -\Phi(-d_1) & \Delta_{P^n} &= (1+k)^T \Phi(-d_2) \\ \Gamma_{P^r} &= \frac{\Phi'(-d_1)}{P^r(t, T) \sigma \sqrt{T-t}} & \Gamma_{P^n} &= -\frac{\Phi'(-d_2)}{P^n(t, T) \sigma \sqrt{T-t}} \end{aligned}$$

5.2.4. Type-2 caps and floors are path independent as the payoffs depend only on the CPI at the payoff date and not on the path of the CPI over the intervening period. Cap and floor values can therefore be calculated independently. This is not true for type-3, -4 or -5 derivatives as these payoffs depend on the CPI adjusted by any caps or floors which have bitten to date. The value

at time t of the type-2 collared LPI cashflow due at time T , $LPI_2(T)$, is then

$$\begin{aligned} V_2(t, T, k, l) &= \mathbb{E}_{\mathbb{Q}}[P^n(t, T)LPI_2(T) | Q_t] \\ &= Q_t P^r(t, T) - c_2(t, T, k) + f_2(t, T, l) \end{aligned} \quad (20)$$

5.2.5. The inflation cap and floor level applies from the date the benefit commenced. Therefore cashflows with the same payment dates but different commencement dates cannot be aggregated for valuation purposes.

5.2.6. To generalise to any indexation lag, consider the two effects introduced by an indexation lag:

- (1) Figure 21 illustrates the period between t , the valuation date, and t' , the latest date in the life of the derivative for which the CPI has been published. The option valuation formula needs to be adjusted to account for this period where there is no inflation uncertainty, or the cumulative volatility factored into the price will be too high.

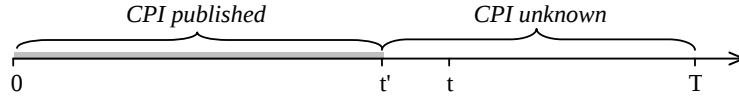


Figure 21: Published and unknown inflation at the valuation date t

- (2) The CPI must be projected to the indexation date corresponding to the payout at time T , as depicted in Figure 22. Section 2 of chapter 4 discusses how the indexation lag for monthly pension payments indexed annually may be up to 14 months.

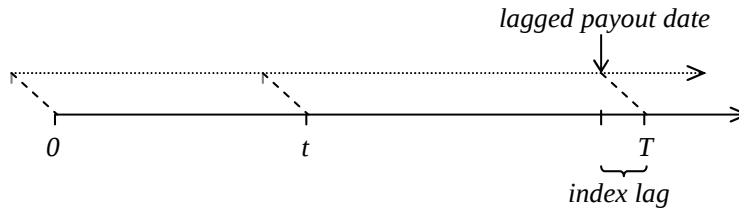


Figure 22: Lag between the date of the payout and indexation date.

5.2.7. The type-2 LPI valuation formulae (18) and (19) are readily adjusted for these effects:

- (1) To ensure the volatility of the inflation index over the period with no inflation uncertainty is zero, scale down the volatility assumption by

the square root of the ratio of the inflation uncertainty period to the discounting period, i.e. use volatility $\sigma' = \sigma \sqrt{\frac{T-t'}{T-t}}$ in place of σ .

- (2) To allow for the indexation lag of the payout, value the real zero $P^r(t, T)$ using a breakeven curve or real yield curve with the same indexation lag.

5.3. Type-3 LPI

5.3.1. The time T payoff for a type-3 LPI liability floored at 0% is defined recursively as

$$LPI_3(T) = \max(Q_T, LPI_3(T-1)).$$

Repeated application of this recurrence relation gives:

$$LPI_3(T) = Q_0 + \max_{\{s=0,1,\dots,T\}} (Q_s - Q_0). \quad (21)$$

This shows that a type-3 payoff is equivalent a T -year discrete time lookback option on the maximum of the CPI index ratio with a fixed strike at 1.

5.3.2. The path dependency of type-3 floors complicates the valuation. The distribution of the maximal path of a Brownian motion has been studied and analytic formulae exist for the case where the floor is set continuously at the maximum (see Goldman, Sosin & Gatto (1979)). Broadie, Glasserman & Kou (1999) have shown how lookback option values can be adjusted when the strike is set at discrete intervals. This is the case when the type-3 floors are applied annually (as is typically the case). Where a type-3 floor occurs in combination with any type of cap, the path dependency becomes complex and valuation requires Monte Carlo simulation.

5.4. Type-4 LPI

5.4.1. The time T payoff for a type-4 benefit capped at 100k% is

$$\begin{aligned} LPI_4(T) &= LPI_4(T-1) \min \left[1+k, \frac{Q_T}{Q_{T-1}} \right] \\ &= LPI_4(T-1) \left(\frac{Q_T}{Q_{T-1}} + \underbrace{\min \left[(1+k) - \frac{Q_T}{Q_{T-1}}, 0 \right]}_{\text{type-2 cap payoff}} \right). \end{aligned} \quad (22)$$

5.4.2. Type-4 caps and floors are therefore path dependent. A one year type-4 cap or floor is equivalent to a 1 year type-2 cap or floor since there is no compounding. Then from equation (20) the value at time $T-1$ of this type-4 cap is

$$\begin{aligned} c_4(T-1, T, k) &= \mathbb{E}_{\mathbb{Q}} \left[P^n(T-1, T) LPI_4(T-1) \max \left[\frac{Q_T}{Q_{T-1}} - (1+k), 0 \right] \mid Q_{T-1} \right] \\ &= LPI_4(T-1) c_2(T-1, T, k). \end{aligned}$$

Similarly, the value at time $T-1$ of the corresponding type-4 floor is

$$\begin{aligned} f_4(T-1, T, k) &= \mathbb{E}_{\mathbb{Q}} \left[P^n(T-1, T) LPI_4(T-1) \max \left[(1+k) - \frac{Q_T}{Q_{T-1}}, 0 \right] \mid Q_{T-1} \right] \\ &= LPI_4(T-1) f_2(T-1, T, k). \end{aligned}$$

The value at time $T-1$ of a type-1 payment at time T , collared by a type-4 cap at $100k\%$ and a type-4 floor at $100l\%$ is

$$\begin{aligned} &\mathbb{E}_{\mathbb{Q}} [P^n(T-1, T) LPI_4(T) \mid Q_{T-1}] \\ &= LPI_4(T-1) \{ P^r(T-1, T) - c_2(T-1, T, k) + f_2(T-1, T, l) \}. \end{aligned} \quad (23)$$

Therefore, the value at time t of the payment at time T and subject to type-4 caps and floors is

$$\begin{aligned} &\mathbb{E}_{\mathbb{Q}} [P^n(t, T) LPI_4(T) \mid LPI_4(t)] \\ &= \mathbb{E}_{\mathbb{Q}} \left[\mathbb{E}_{\mathbb{Q}} [P^n(t, T) LPI_4(T) \mid LPI_4(T-1), LPI_4(t)] \mid LPI_4(t) \right] \\ &\quad \text{using the Tower Law} \\ &= \mathbb{E}_{\mathbb{Q}} \left[P^n(t, T-1) LPI_4(T-1) \{ P^r(T-1, T) - c_2(T-1, T, k) + f_2(T-1, T, l) \} \mid LPI_4(t) \right] \\ &\quad \text{from (23)} \\ &= \mathbb{E}_{\mathbb{Q}} \left[P^n(t, T-1) LPI_4(T-1) \left\{ P^r(T-1, T) [1 - \Phi(d_{1,k}) - \Phi(-d_{1,l})] \right. \right. \\ &\quad \left. \left. + P^n(T-1, T) [(1+k)\Phi(d_{2,k}) + (1+l)\Phi(-d_{2,l})] \right\} \mid LPI_4(t) \right] \end{aligned}$$

where
$$d_{1,k} = \frac{\ln(Q_t P^r(t, T)) - \ln((1+k)^T P^n(t, T)) + \frac{1}{2}\sigma^2(T-t)}{\sigma\sqrt{T-t}},$$

$d_{2,k} = d_1 - \sigma\sqrt{T-t}$ and $d_{1,l}$ and $d_{2,l}$ are defined in a similar manner.

5.4.3. Inflation is a Markov process in the basic Black-Scholes world, so applying the recurrence relation gives

$$LPI(t) = LPI(t-1) \left\{ \left(\frac{Q_t}{Q_{t-1}} \right) + \max \left[(1+l) - \frac{Q_t}{Q_{t-1}}, 0 \right] - \max \left[\frac{Q_t}{Q_{t-1}} - (1+k), 0 \right] \right\}.$$

5.4.4. We then reconcile to the value given in Van Bezooyen et al. (1997) for the present value at time t of a type-4 LPI payment collared between cap

k and floor l until payment time T , $V_4(t, T, k, l)$

$$\begin{aligned}
&= \mathbb{E}_{\mathbb{Q}} [LPI_4(T) \mid LPI(t)] \\
&= LPI(t) \prod_{u=t}^{T-1} \left(P^r(u, u+1) \{ 1 - \Phi(d_{1,k}) - \Phi(-d_{1,l}) \} \right. \\
&\quad \left. + P^n(u, u+1) \{ (1+k)\Phi(d_{2,k}) + (1+l)\Phi(-d_{2,l}) \} \right). \tag{24}
\end{aligned}$$

5.4.5. Type-4 liabilities cannot be expressed more compactly in terms of real and nominal zero-coupon rates. The risk with respect to these zero-coupon rates is therefore not readily derived as it was for type-2 LPI in section 5.2.2. Greeks with respect to the forward real rates, $g(t, 1) = \ln P^r(t, t+1)$, and forward nominal rates, $f(t, 1) = \ln P^n(t, t+1)$, can be derived analytically from (24). The curve bumping procedure section of 11.3 of chapter 4 above can be used to derive the delta hedge in terms of real and nominal zero-coupon yields.

5.5. Type-5 LPI

5.5.1. Let π denote the fraction of CPI by which the benefit is increased each year. The time T payoff can be expressed using the recurrence relation:

$$LPI_5(T) = LPI_5(T-1) \left(1 + \pi \left(\frac{Q_T}{Q_{T-1}} - 1 \right) \right). \tag{25}$$

The value at time t of a type-5 payment at time T , $V_5(t, T, \pi)$

$$\begin{aligned}
&= \mathbb{E}_{\mathbb{Q}} \left[P^n(t, T) LPI_5(T) \mid LPI_5(t) \right] \\
&= \mathbb{E}_{\mathbb{Q}} \left[\mathbb{E}_{\mathbb{Q}} [P^n(t, T) LPI_5(T) \mid LPI_5(T-1), LPI_5(t)] \mid LPI_5(t) \right] \text{ using the Tower Law} \\
&= \mathbb{E}_{\mathbb{Q}} \left[P^n(t, T-1) LPI_5(T-1) \left\{ P^n(T-1, T) \left[1 + \pi \left(\frac{Q_T}{Q_{T-1}} - 1 \right) \right] \right\} \mid LPI_5(t) \right] \text{ by (25)} \\
&= \mathbb{E}_{\mathbb{Q}} \left[P^n(t, T-1) LPI_5(T-1) \left\{ (1-\pi)P^n(T-1, T) + \pi P^r(T-1, T) \right\} \mid LPI_5(t) \right].
\end{aligned}$$

Repeated application of (25) then gives:

$$V_5(t, T, \pi) = LPI_5(t) \prod_{u=t}^{T-1} \left((1-\pi)P^n(u, u+1) + \pi P^r(u, u+1) \right). \tag{26}$$

5.5.2. If a type-5 LPI liability is collared between a $100k\%$ cap and $100l\%$ floor, the present value at t is

$$\begin{aligned}
V_5(t, T, \pi, k, l) = LPI(t) \prod_{u=t}^{T-1} \left(\pi P^r(u, u+1) \{ 1 - \Phi(d_{1,k}) - \Phi(-d_{1,l}) \} \right. \\
\left. + (1-\pi)P^n(u, u+1) \{ (1+k)\Phi(d_{2,k}) + (1+l)\Phi(-d_{2,l}) \} \right). \tag{27}
\end{aligned}$$

5.5.3. Type-5 LPI, like type-3 and type-4, is path-dependent. The Greeks with respect to shift in the real and nominal forward rates between payment

increases can be derived analytically.

The delta with respect to the real zero-coupon bond price is,

$$\Delta_{Pr(u,u+1)} = \frac{\partial V_5(t, T, \pi)}{\partial Pr(u, u+1)} = \frac{\pi V_5(t, T, \pi)}{(1-\pi)P^n(u, u+1) + \pi Pr(u, u+1)}. \quad (28)$$

The delta with respect to the nominal zero-coupon bond price is,

$$\Delta_{P^n(u,u+1)} = \frac{\partial V_5(t, T, \pi)}{\partial P^n(u, u+1)} = \frac{(1-\pi)V_5(t, T, \pi)}{(1-\pi)P^n(u, u+1) + \pi Pr(u, u+1)}. \quad (29)$$

5.6. Effectiveness of the Black–Scholes currency model

5.6.1. The currency option analogy for inflation allows analytic formulae to be produced for the value of many inflation liabilities. The assumptions underpinning the model are unrealistic, but the model can be used as a heuristic base case for further model development.

5.6.2. Van Bezooyen et al. (1997) note that the premise of no interest rate variability may be defensible for short-dated options where the underlying is not closely related to interest rates (such as equities) but is not realistic for long-dated options on inflation. We return to this discussion in section 9.

6. Condition 1: Perfect Markets

6.1. Interdependence between conditions

In the next four sections we consider the Black–Scholes-world conditions of section 3.2. We consider the assumptions in isolation, like the ‘ceteris paribus’ assumption frequently used in economics. The interaction between conditions may be significant. For example, the perfect markets assumption degrades significantly when the market exhibits periods of high volatility, which violates the deterministic volatility assumption.

6.2. Liquidity of inflation-linked bond and derivative markets

Inflation-linked bond and derivative markets are considerably less liquid than their nominal counterparts (Deacon et al. 2004). A high proportion of investors hold inflation-linked assets for hedging, such as matching inflation-linked pension liabilities, and therefore transact less frequently. The average monthly turnover of ZAR 3.95bn for SA inflation-linked bonds is less than ten percent of the turnover of their nominal comparator bonds as a percentage of

bonds in issue¹. This average monthly turnover is 0.4% of total SA pension liabilities but is still significant in relation to the liabilities in respect of pensions in payment for many funds. Pension funds will usually be in a position to delay transactions until there is adequate liquidity, such as the monthly SA inflation-linked bond auctions. The inflation swap market further increases the liquidity available to investors.

6.3. Transaction costs

6.3.1. Various costs are associated with inflation-linked bonds and swaps:

- bid-offer spreads;
- agent's commission, trading exchange charges; and
- taxes such as stamp duty.

6.3.2. SA inflation-linked bonds are not subject to stamp duty, in South Africa termed Marketable Securities Tax. Bid-offer spreads for bonds are of the order of 2bp or less on standard sizes (R5m face value). They may be lower if an order is left with a dealer, but at the risk of market movements. Exchange charges are significant only for very small trades.

6.4. Short positions

The perfect-markets condition assumes it is possible to take a short position of unlimited size. A reverse repo transaction is required to cover a short bond position. The SA inflation-linked bond reverse repo market is very limited and National Treasury at present places a limit on the volume of inflation-linked bonds made available as part of its reverse repo market operations. In more advanced markets a reliable reverse repo market has developed. Short squeezes in inflation-linked bond, when it becomes difficult to borrow bonds to maintain a short position, occur infrequently. There are no such constraints to maintaining a short position by paying inflation in an inflation swap.

6.5. Discrete hedging

6.5.1. In practice, derivatives can be hedged only in discrete time and it is inevitable a hedge error will result. This section probes the significance of this hedge error.

6.5.2. Consider the indexed price of a T -year inflation-linked zero-coupon bond when real rates g are constant and the inflation process is given by (16)

¹ Bond Exchange of SA Monthly. Data from the Bond Data Report for the three months to June 2006, excluding repo trades.

with constant drift and volatility. Denote this indexed inflation-linked zero-coupon bond price by P^r and the log of this price by X . Then:

$$dX = d \ln P^r = P^r ((\mu + g)dt + \sigma dW_t), \quad \text{where } W_t \text{ is a } \mathbb{P}\text{-Brownian motion.}$$

The change in the log of the indexed inflation-linked zero-coupon bond price over a discrete but small time interval Δt is denoted by ΔX , where:

$$\Delta X = (\mu + g - \frac{1}{2}\sigma^2)\Delta t + \epsilon\sigma\sqrt{\Delta t} \quad \text{where } \epsilon \sim N(0, 1) \quad (30)$$

6.5.3. Following Wilmott (1994), we investigate the optimal hedge and the hedging error when rehedging at discrete intervals. Let Π_t denote the value of the portfolio consisting of an inflation derivative with value V_t and, as the hedge instrument, real zero bonds with values P^r . In the simple Black–Scholes world there is no nominal or real interest rate volatility, so the sole source of variability for V and P^r is the inflation index. Let β denote the optimal inflation-linked zero-coupon bond hedge ratio. The portfolio hedged at time t is worth

$$\Pi_t = V_t - \beta P^r(t, T) = V_t - \beta \exp(X_t).$$

A Taylor series expansion of Π with respect to X and t gives:

$$\Delta \Pi \approx \Delta t \left(\frac{\partial V}{\partial t} \Delta t - \beta P^r \frac{\partial X}{\partial t} \right) + \Delta X P^r \left(\frac{\partial V}{\partial P^r} - \beta \right) + \frac{1}{2} (P^r \Delta X)^2 \frac{\partial^2 V}{\partial (P^r)^2}$$

This expansion is to first-order in Δt and second-order in ΔX to ensure that, on substitution for ΔX using (30), the result is to order Δt :

$$\begin{aligned} \Delta \Pi \approx & \sqrt{\Delta t} \epsilon \sigma P^r \left(\frac{\partial V}{\partial P^r} - \beta \right) \\ & + \Delta t \left\{ \frac{\partial V}{\partial t} + P^r \left(\frac{\partial V}{\partial P^r} - \beta \right) (\mu + g + \frac{1}{2}\sigma^2(\epsilon^2 - 1)) + \frac{1}{2}\sigma^2\epsilon^2 (P^r)^2 \frac{\partial^2 V}{\partial (P^r)^2} \right\} \end{aligned} \quad (31)$$

Using the Black–Scholes hedge ratio, $\beta = \Delta_{P^r} = \frac{\partial V}{\partial P^r}$, gives hedge error over Δt :

$$\Delta \Pi \approx \Delta t \left\{ \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2\epsilon^2 (P^r)^2 \frac{\partial^2 V}{\partial (P^r)^2} \right\} \quad (32)$$

6.5.4. By comparison, in the idealised world of continuous rehedging the Black–Scholes partial differential equation (4.1) implies that a delta hedged portfolio experiences the following changes due to theta and gamma only in an infinitesimal time step:

$$d\Pi = \frac{\partial V}{\partial t} dt + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} dt \quad (33)$$

For small Δt , the difference between (32) and (33) is the discrete time hedge error:

$$\Delta t \frac{1}{2} \sigma^2 (P^r)^2 \frac{\partial^2 V}{\partial (P^r)^2} (\epsilon^2 - 1) \quad (34)$$

The discrete reheding error is therefore proportional to the gamma of the derivative, $\Gamma_{P^r} = \frac{\partial^2 V}{\partial (P^r)^2}$.

6.5.5. Since $\epsilon^2 \sim \chi_1^2$, the hedge error over Δt is a Gamma random variable with shape parameter $\frac{1}{2}$, scale parameter $\Delta t \frac{1}{2} \sigma^2 (P^r)^2 \Gamma_{P^r}$ and a shifted location. The expected value of the hedge error is always zero, but its standard deviation depends on $P^r(t, T)$. The cumulative hedging error is a sum of Gamma variables with different scale parameters. These scale parameters are highly path-dependent.

6.5.6. Figure 23 shows Δ_{P^r} , the inflation-linked zero-coupon bond hedge ratio, for a 3-year type-2 inflation cap, struck at the money forward, when the hedging takes place daily and monthly. The increase in gamma, and consequent hedge error, is apparent as maturity approaches. As discussed in section 8.2 of chapter 4, the risk attributable to inflation indexation for the liabilities becomes more predictable with a time horizon of a year. Smith (1999) shows how the information structure can be incorporated in a model for pricing inflation derivatives. However, when there is under a month to the indexed payoff, inflation volatility actually increases relative to time outstanding since the index moves in discrete monthly steps.

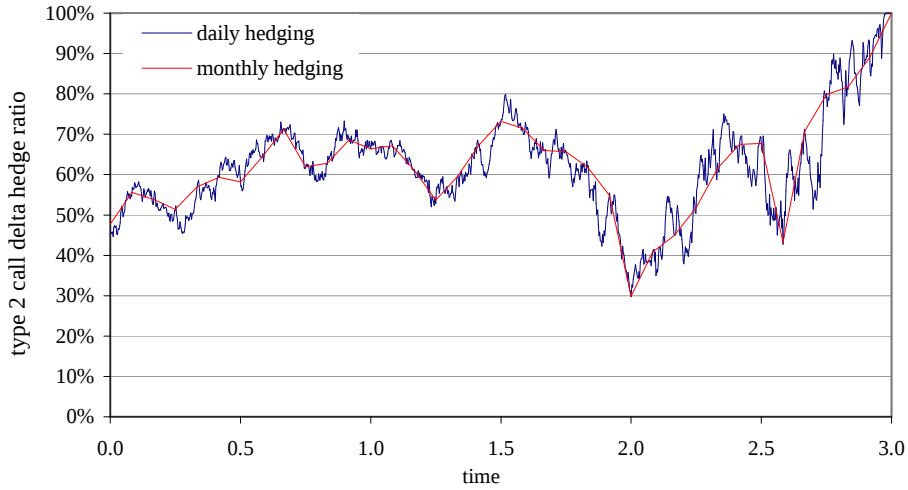


Figure 23: Monthly and daily hedge ratios for a 3-year type-2 inflation cap

6.5.7. Consider the hedging error for a derivative with present value $V(t, T)$ at t and payoff at T that is hedged at N time intervals of length $T/\Delta t$. Let σ_N denote the standard deviation of the cumulative hedge error. (Kamal

& Derman 1999) give the following approximation for σ_N for large N (*i.e.* frequent rebalancing):

$$\sigma_N \approx \sigma \sqrt{\frac{\pi}{P^r N}} \left. \frac{\partial V(t, T)}{\partial \sigma} \right|_{t=0}. \quad (35)$$

The cumulative hedging error therefore depends on the vega, $\frac{\partial V}{\partial \sigma}$, at inception of the option. Kamal & Derman note their approximation will tend to understate the true σ_N for options that are deeply out-of-the-money (as is often the case for inflation options embedded in pension benefits).

6.5.8. If the perfect markets assumption is violated, riskless replication is not possible and derivative prices are not preference-independent. The optimal hedge will depend on the utility function of the hedger. Let $\mathbb{P}$ be the real-world measure (in the sense of the being a representative agent's measure). If β is chosen to minimise the variance of the hedge error under $\mathbb{P}$, Wilmott (1994) shows that:

$$\beta = \frac{\partial V}{\partial P^r} + \Delta t (\mu - (f - g) + \sigma^2) P^r \frac{\partial^2 V}{\partial (P^r)^2}. \quad (36)$$

In contrast to the Black-Scholes continuous hedge ratio, $\frac{\partial V}{\partial P^r}$, the minimum-variance hedge ratio depends on the real-world drift μ of the inflation process. If μ is high relative to breakeven inflation, $f - g$, the minimum-variance hedge anticipates the movement in the CPI to a greater degree, adjusting the hedge ratio accordingly. A higher gamma with respect to $P^r(t, T)$ increases the required degree of anticipation.

6.5.9. In figure 24, the minimum-variance delta (36) is compared with the continuous-hedge delta, Δ_{P^r} , for a long position in a 5-year type-2 capped liability. The discrete hedging period is between $t = 0$ and $t = \Delta t = 1/12$. Over this first hedging period the real-world drift is $\mu = 9\%$, the market forward-rate $f - g = 6.68\%$ and volatility $\sigma = 3\%$. The probability distribution of the hedging error in Figure 24 is depicted by marking the percentiles from 10% to 90% in bands of 10%. Since the hedging error is unbounded, 1% and 99% percentiles are used for the outer two bands. The graph therefore marks a 98% confidence interval for the hedging error. For the Black-Scholes hedge ratio this interval is $[0.088\%, 0.183\%]$ and for the discrete minimum-variance hedge ratio this interval is $[0.143\%, 0.07\%]$, which is slightly narrower. The graph can be viewed as analogous to the marginal distribution (for the first slice of time) of a CPI fanchart of the hedging error. The significance of hedge errors further diminishes when the benefit is collared rather than just capped or floored. Figure 25 shows the hedge error for $\Delta t = 1/12$ and $\Delta t = 1/100$ for the

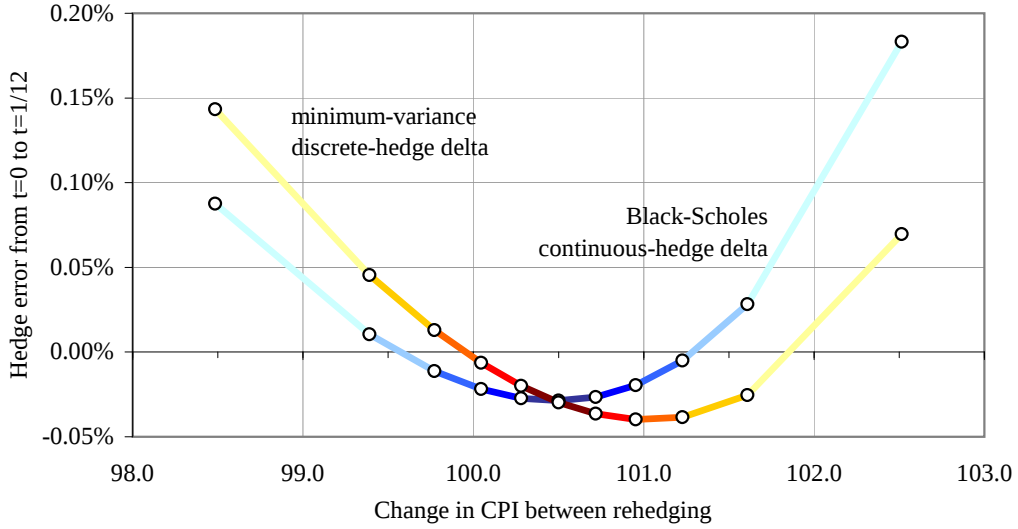


Figure 24: Distribution of hedge error over $\Delta t = 1/12$ for Black-Scholes continuous hedge and minimum-variance discrete hedge

short cap of Figure 24, a long floor with the same (but opposite sign) delta and the combined collar.

6.5.10. Although the optimal discrete hedge ratio will depend on the subjective drift estimate and utility function of the hedger, Wilmott argues that market prices are nevertheless preference-free because the market uses models based on the assumption of an idealised perfect market, despite hedging in discrete intervals. Then the hedge portfolio is expected to earn the (compounded) risk-free return over Δt :

$$\mathbb{E}[\Delta\Pi] = (r\Delta t + \frac{1}{2}r^2\Delta t^2 + \dots) \quad (37)$$

Substituting $\Pi_t = V_t - \beta P^r(t, T)$ and (31) into (37) gives a discrete hedging partial differential equation analogous to the Black-Scholes partial differential equation in equation (17):

$$\frac{1}{2}\sigma^2(P^r)^2\Gamma_{P^r} + (f-g)P^r\Delta_{P^r} + \frac{1}{2}\delta t(\mu-f-g)(f-g-\mu-\sigma^2)(P^r)^2\Gamma_{P^r} + \Theta_{P^r} = fV.$$

6.5.11. For type-2 LPI caps and floors, where there is no path dependency, this PDE can be solved to give the same formulae as the continuous hedging Black-Scholes formulae in (18) and (19), but with an adjusted volatility parameter:

$$\sigma^* = \sigma\left(1 + \frac{\delta t}{2\sigma^2}(\mu-f-g)(f-g-\mu-\sigma^2)\right)$$

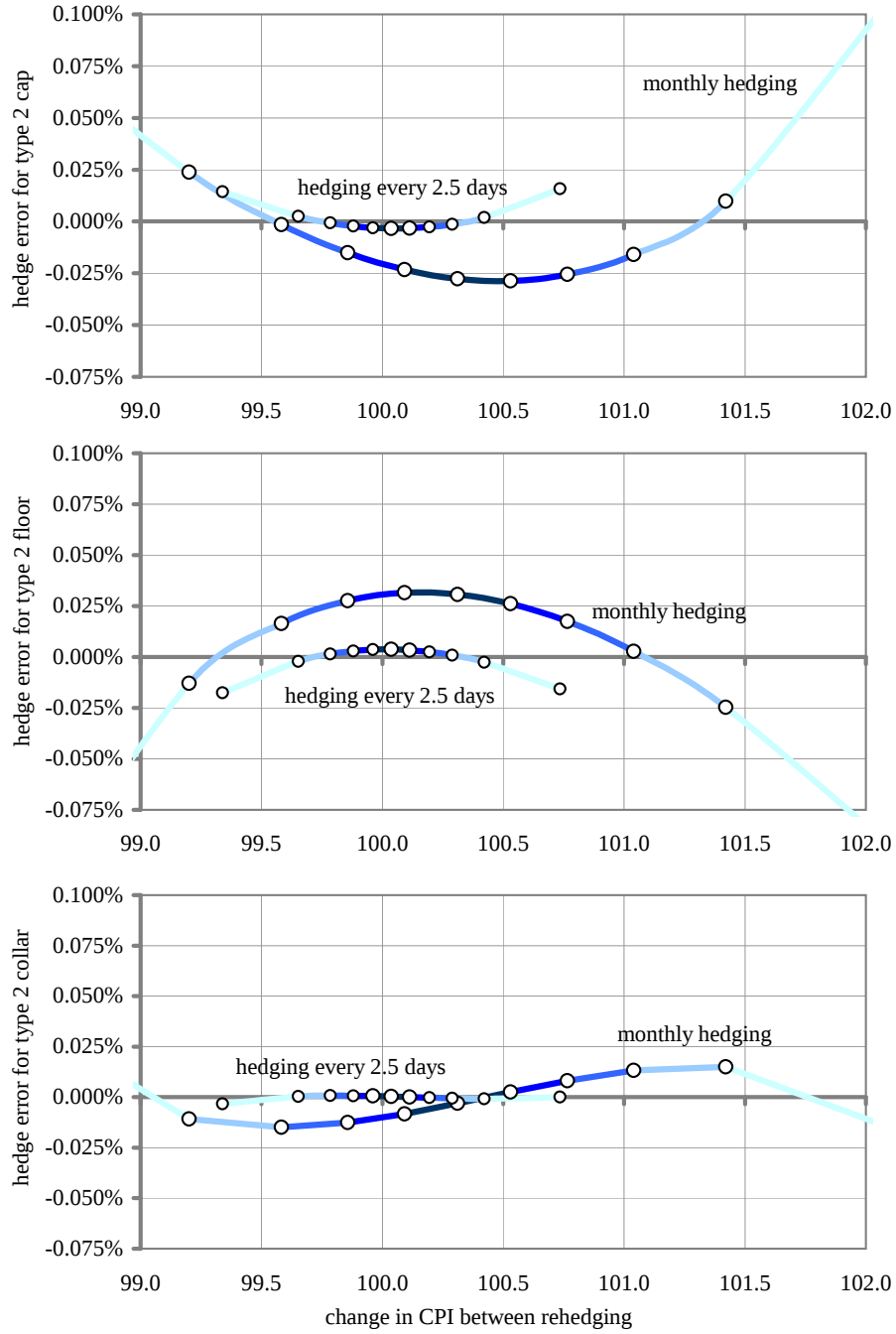


Figure 25: Distribution of hedge errors over $\Delta t = 1/12$ and $\Delta t = 1/100$ for 5-year type-2 cap, equivalent-delta floor and collar.

6.5.12. Since $\sigma^* < \sigma$, discrete hedging leads to lower inflation option prices than continuous hedging, despite hedging error risk being introduced. To an

extent this can be explained by the hedge being anticipatory, but the assumption made in the derivation of σ^* that the risky portfolio earns the risk-free rate may be dubious. Also, market implied volatilities will already reflect discrete hedging. The volatility parameter adjustment to account for discrete hedging is minimal if the estimated real-world drift μ and breakeven inflation $n - r$ are close. This difference is likely to be small for long dated pension liabilities. For very short maturities the hedger's inflation forecast may differ substantially from market forward inflation, for example if there is significant flow seasonality in market rates as discussed in section 8.1 of chapter 4.

6.6. Transaction costs

6.6.1. Leland (1985) showed that if:

- (1) the derivative is delta reheded every δt ;
- (2) transaction costs a proportion κ of the real zero price; and
- (3) the Black–Scholes conditions are otherwise satisfied,

then long option positions should be valued using the adjusted volatility:

$$\sigma^* = \sigma \left(1 - \frac{\kappa}{\sigma} \sqrt{\left(\frac{8}{\pi \delta t} \right)} \right)^{1/2}.$$

6.6.2. The volatility for short positions should be scaled up in a similar way. Transaction costs are therefore proportional to vega for individual options, but a portfolio of derivatives will often have offsetting hedging transactions and the portfolio gamma may change sign so the loading for costs is not additive. The costs of trading real and nominal bonds and swaps is also proportional to the market rate rather than price.

6.7. Materiality

6.7.1. Table 7 shows σ_N , the Kamal & Derman (1999) standard deviation of the approximate present value of cumulative hedging errors over the life of a 10-year type-2 inflation cap, given by equation (35). The true hedge error, estimated from 10 000 trials of a Monte Carlo simulation, is shown alongside this. The

cap has a strike at-the-money, where the standard deviation of the hedging error is maximised (since vega is maximised for at-the-money options). The real-world process, again viewed as the probability measure of the representative agent, is simulated with a volatility of 3%. The Kamal & Derman

δt	N= T/ δt	Hedge error std dev.	Kamal Derman approximation
1	10	24.80%	28.84%
1/12	120	7.28%	8.33%
1/100	1000	2.59%	2.88%

Table 7: 10-year atm type-2 inflation cap cumulative hedge error for various δt

approximation slightly overstates the true value. The cumulative hedge error, which is equally likely to be profit or loss, is reasonably low and drops rapidly beyond the range shown in the table, to 0.25% when $N=2500$ and the portfolio is reheded every day.

6.7.2. The hedge risk for a portfolio of inflation derivatives will depend on the portfolio gamma and vega, which a dealer will seek to mitigate through offsetting positions. There will also be some positions which have natural offset at inception, e.g. long caps and short floors in the collar of type-4 LPI, although this offset may decline as liability matures. Offsetting positions will also decrease hedging transaction costs.

6.7.3. The transaction costs incurred when hedging options will depend on the structure of the costs, frequency of reheding and the market presence of the hedger. There will be a natural trade-off between trade costs (including market impact) and the frequency with which the portfolio is reheded. The primary determinant of transaction costs will be the degree to which the hedger is a market-maker or a price-taker for the underlying.

6.7.4. Market-makers may be more tolerant of hedge errors, which will offset between diverse business lines. Market makers can also use the synergy between the cash and derivative inflation markets to decrease their transaction costs. There will be further synergies between related inflation and interest rates books (for example a Euro-zone inflation dealer may find it more cost-effective to hedge HICP inflation with French inflation until the position can be neutralised in the HICP market.) These factors make it difficult to gauge the materiality of transaction costs, but for dealers these costs are likely to be a secondary factor.

6.8. Discrete hedging and traditional actuarial valuations

6.8.1. The above discussion on discrete hedging strategy sheds further light on the distinction between the traditional actuarial approach and the financial approach to valuing inflation liabilities. The traditional actuarial approach is to discount the expected payoff under the real-world measure to give a best estimate for the inflation-linked derivative pension,

$$V_t^a = \mathbb{E}_{\mathbb{P}}[P^n(t, T)LPI(T) | \mathcal{F}_t]$$

6.8.2. The financial approach values the liability as the cost of its self-financing replicating portfolio in a perfect market. Consider the limiting case where $\delta t = T-t$, so that there is no portfolio rebalancing before maturity. The hedge portfolio, denoted by H , has value $H(t)$ at time t . Suppose H is chosen

to minimise the variance of the hedge error at maturity:

$$\min_{\mathbf{H}} \left[\text{var}_{\mathbb{P}} \left[P^n(t, T) (LPI(T) - \mathbf{H}(T)) \mid \mathcal{F}_t \right] \right].$$

$$\Rightarrow \mathbf{H}(T) = \mathbb{E}_{\mathbb{P}} [P^n(t, T) LPI(T) \mid \mathcal{F}_t].$$

6.8.3. The financial approach values the benefit as $V_t^f = \mathbf{H}(t)$, the current value of the hedge portfolio. Furthermore if $\mathbb{P}$ corresponds to the risk-neutral measure $\mathbb{Q}$ and interest rates are deterministic as in the Black–Scholes world of this section, then $\mathbf{H}(t) = \mathbf{H}(T) P^n(t, T)$. The traditional actuarial approach can therefore be viewed as the limiting case of the financial approach when there is no opportunity for portfolio reheding before maturity.

6.8.4. Note that the actuarial approach may alternatively define V_t^a as the present value of the LPI derived using the expected LPI rate under $\mathbb{P}$, $LPI\%(t)$ in year t . That is,

$$V_t^a = \exp \left\{ \sum_{u=t}^{T-1} \mathbb{E}_{\mathbb{P}} [LPI\%(u) \mid \mathcal{F}_t] \right\} P^n(t, T)$$

In this case, Jensen’s inequality implies that the actuarial approach will produce a lower value than the financial approach. This follows from:

$$\begin{aligned} & \exp \left\{ \sum_{u=t}^{T-1} \mathbb{E}_{\mathbb{P}} [LPI\%(u) \mid \mathcal{F}_t] \right\} P^n(t, T) \\ & \leq \mathbb{E}_{\mathbb{P}} \left[\exp \left(\ln P^n(t, T) \sum_{u=t}^{T-1} LPI\%(u) \right) \mid \mathcal{F}_t \right] = \mathbf{H}(T). \end{aligned}$$

7. Condition 2: Inflation driven by Brownian motion

7.1. Generalisation

The basic Black–Scholes assumption is that the stochastic price process for the underlying is exponential Brownian motion, resulting in normal returns and lognormal prices. This assumption can be extended to any Itô process:

$$dQ_t = Q_t (\sigma(t, Q_t) dW_t + \mu(t, Q_t) dt); \quad (38)$$

where W_t is a Brownian motion under $\mathbb{P}$ and σ and μ are adapted to the filtration of Q_t . This is the most general form for which, by Itô’s Lemma, any process which is a twice continuously differentiable function of Q_t is still an Itô process driven by Brownian motion and Girsanov’s Theorem can be applied (subject to its regularity conditions). The Martingale representation theorem

can then be invoked, subject to its regularity conditions, to prove the existence of a self-financing replicating trading strategy for derivatives on this tradable.

7.2. Tests for normality

7.2.1. It is natural to model the spot inflation rate as a normal variable, implying a lognormal inflation index. This corresponds to $\sigma(t, Q_t) = \sigma_t$ and $\mu(t, Q_t) = \mu_t$ in (38). Figure 26 shows a histogram of the absolute month-on-month change in the SA headline inflation rate, for the period since inflation targeting was introduced, after stripping out the underlying drift and seasonality of the process using the method for deriving the seasonality factors. The histogram is broadly consistent with a normal distribution for increments. There is a single outlier of +0.7% in January 2002, after sudden currency weakness where it is possible the detrending method failed to capture the sudden change in underlying trend.

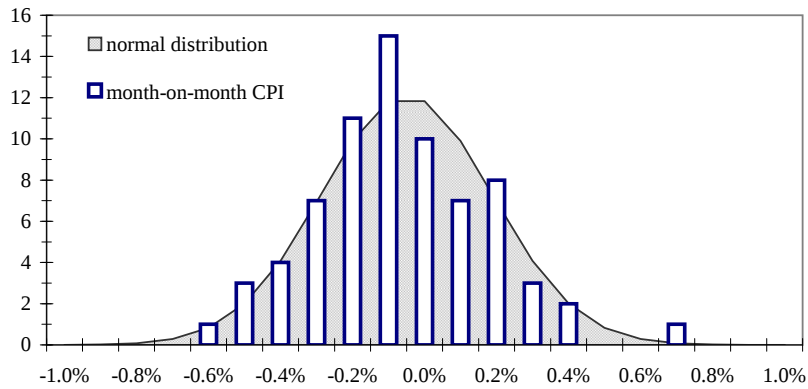


Figure 26: Histogram of month-on-month SA headline CPI, June 2000 to June 2006, standardised by the seasonality factors in table 2 of of chapter 4, section 7.

7.2.2. Rounding of the inflation index implies a discrete distribution for absolute changes. All tradable instruments have discrete price movements (or ‘ticks’) that are approximated with continuous price models. The rounding of the inflation index is particularly significant compared with its volatility, but this will be mitigated by the trend to the publication of the CPI to two decimals, as noted section 1.0.3 of chapter 4.

7.2.3. Another objection to modelling with Brownian motion is that the inflation index is a discrete-valued process in discrete (monthly) time periods. This need not be an issue since whether we model in continuous time and evaluate at discrete time points or model directly in discrete time, this feature of the inflation process is accounted for.

7.2.4. A more informative graphical test of a normal fit for $\Delta \ln(Q_t)$ is a quantile or probability plot against a normal model as shown in Figure 27. These plots are generated by ordering the k data points in ascending order $y_1 \leq y_2 \leq \dots \leq y_k$. Let $\hat{F}$ denote the cumulative distribution function of the fitted normal distribution. Then the Q-Q plot is a plot of the points $\{y_i, \hat{F}^{-1}(\frac{i}{k+1})\}$ in data space. The P-P plot is a plot of the points $\{\frac{i}{k+1}, \hat{F}(y_i)\}$ in probability space. Both plots show a good fit with a normal model.

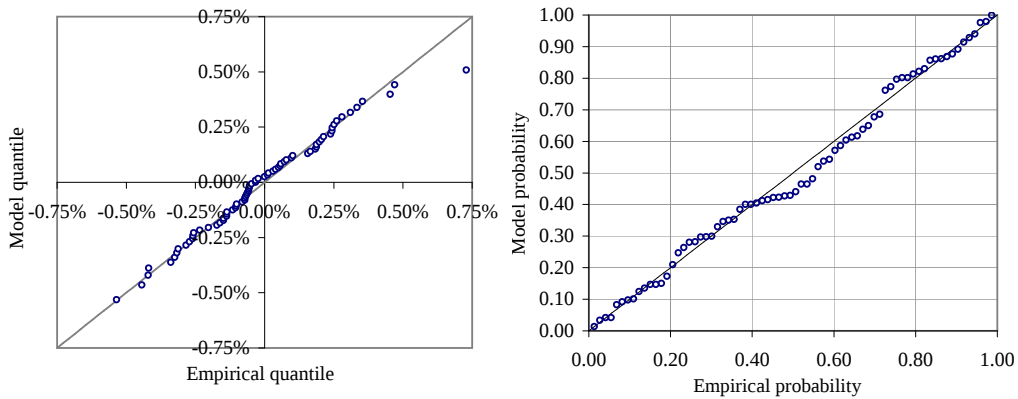


Figure 27: Q-Q and P-P plot for the absolute change in the SA headline inflation rates, June 2000 to June 2006.

7.2.5. There are also a number of statistical tests for the hypothesis of a normal model, such as the Kolmogorov–Smirnov test, Jarque–Bera test etc. These tests do not reject the hypothesis of a normal model for the monthly change in detrended SA headline inflation (from July 2000 to June 2006) at the 95% significance level.

7.2.6. For inflation-index increments to be driven by Brownian motion, the increments must also be independent. The next section investigates mean reversion of the spot inflation rate. As a prelude to weakening the Black–Scholes assumption of deterministic interest rates, we also investigate mean reversion of breakeven inflation and nominal interest rates.

8. Inflation targeting and mean reversion

8.1. Inflation targeting in South Africa

8.1.1. South Africa introduced inflation targeting in February 2000. The target was initially set at 3 to 6% for CPIX, the consumer price index for the metropolitan and other urban areas excluding interest on home loans. CPIX

differs from the headline CPI, for all items and for the metropolitan areas only, used to index South African inflation-linked bonds, but the interest rates element excluded accounts for only 10.3% of the survey basket so there is a close correspondence between CPIX and CPI. These indices have been 89% correlated since inflation targeting was introduced, see Figure 28.

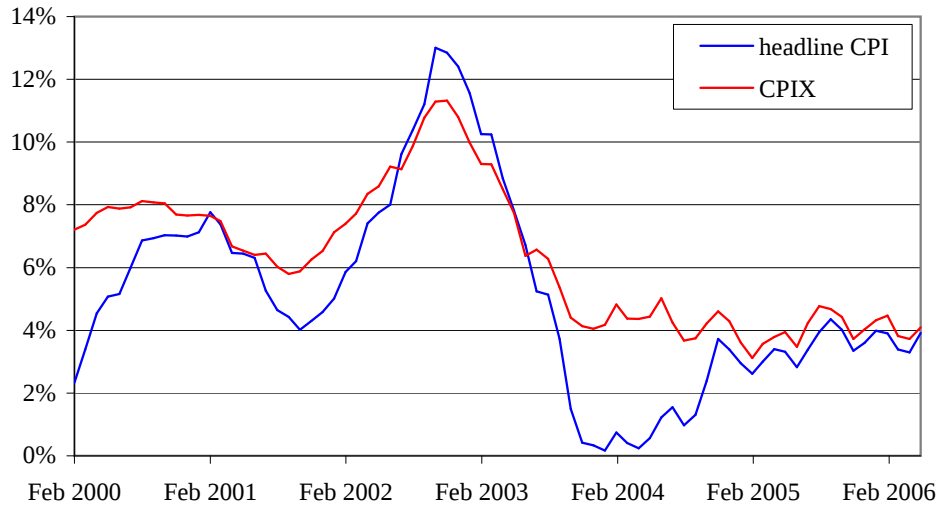


Figure 28: SA CPI and CPIX since introduction of inflation targeting

8.1.2. The SA government, through National Treasury, sets the target. The South African Reserve Bank (SARB), as independent central bank, is responsible for achieving the target through monetary policy. The CPIX target range for 2003 and 2004 was set lower at 3 to 5% but this target was missed due to the extreme currency weakness of late 2001². The 3 to 6% target was then reinstated and has remained at this level to date.

8.1.3. The adoption of an inflation target may alter the dynamics of inflation and interest rates. The target range will be set after the normal fluctuation of inflation with the business cycle is taken into account. An external shock, such as the currency weakness in 2001, may cause inflation to deviate far from the target, but a successful monetary policy response will channel inflation back into the target range in the medium to long-term. The inflation process can therefore be expected to show mean reversion with a broad cycle consistent in phase with the business cycle. Furthermore, short-interest rates, being the primary tool of monetary policy, also fluctuate within a range and exhibit mean reversion. Longer-term rates will too show mean reversion

² 'The objectives and importance of inflation targeting', T.T. Mboweni, South African Reserve Bank press release, 2002.

to the extent that parallel curve shifts are dominant, but the lower volatility of longer-dated forward-rates and decorrelation between forward-rates will undermine and dampen the cycle.

8.2. Correlation structure of SA interest rates and inflation

8.2.1. Table 8 shows continuously compounded real, nominal and breakeven rates using the medium-term benchmark R189 2013 inflation bond and its R157 2015 nominal comparator. The CPI reference for each 26 June date is taken as the prior 1 March for simplicity - the true reference CPI based on a 4-month lagged interpolation between February and March would give very similar results. The period is taken from the start of inflation targeting. The annual change in the nominal, breakeven and headline CPI rate is shown under columns ΔL_t , ΔY_t and ΔC_t and plotted in Figure 29. Mean reversion for nominal and breakeven rates and spot inflation is apparent and there is a negative correlation between nominal rates and spot inflation as expected.

date	reference CPI	headline CPI y/y C_t	R189 2013 real yield R_t	R157 2015 nominal yield L_t	breakeven inflation $Y_t = L_t - R_t$	ΔL_t	ΔY_t	ΔC_t
26-Jun-00	01-Mar-00	3.33%	6.38%	13.76%	7.38%			
26-Jun-01	01-Mar-01	7.12%	5.51%	10.68%	5.17%	-3.08%	-2.21%	3.78%
26-Jun-02	01-Mar-02	6.02%	4.14%	11.57%	7.43%	0.89%	2.26%	-1.10%
26-Jun-03	01-Mar-03	9.75%	4.01%	8.80%	4.79%	-2.77%	-2.64%	3.73%
25-Jun-04	01-Mar-04	0.41%	3.82%	10.04%	6.22%	1.23%	1.42%	-9.34%
27-Jun-05	01-Mar-05	2.96%	3.09%	7.82%	4.74%	-2.22%	-1.48%	2.55%
26-Jun-06	01-Mar-06	3.33%	2.70%	8.08%	5.38%	0.26%	0.65%	0.37%

Continuously compounded rates. Source: Statistics SA, Bond Exchange of SA

Table 8: Annual absolute change for nominal, breakeven and CPI inflation rates

8.2.2. Table 9 shows first-order serial-correlations between ΔL_t , ΔY_t and ΔC_t . The time-series absolute annual rate change data, ΔL_t , ΔY_t and ΔC_t , in Table 8 are assumed to be stationary. For example, ΔL_t is the correlation between series $\{\Delta L_t\}_{t=1,\dots,5}$ and $\{\Delta L_t\}_{t=2,\dots,6}$. The correlation structure accords with economic intuition for mean reversion but the limited size of the sample, with just 6 yearly increases, raises the question of statistical significance. The small size of the sample also creates a material bias in the correlations between the overlapping time series marked with an asterisk. This bias in the covariance is analogous to the bias of the biased sample variance used as an estimator of the population variance.

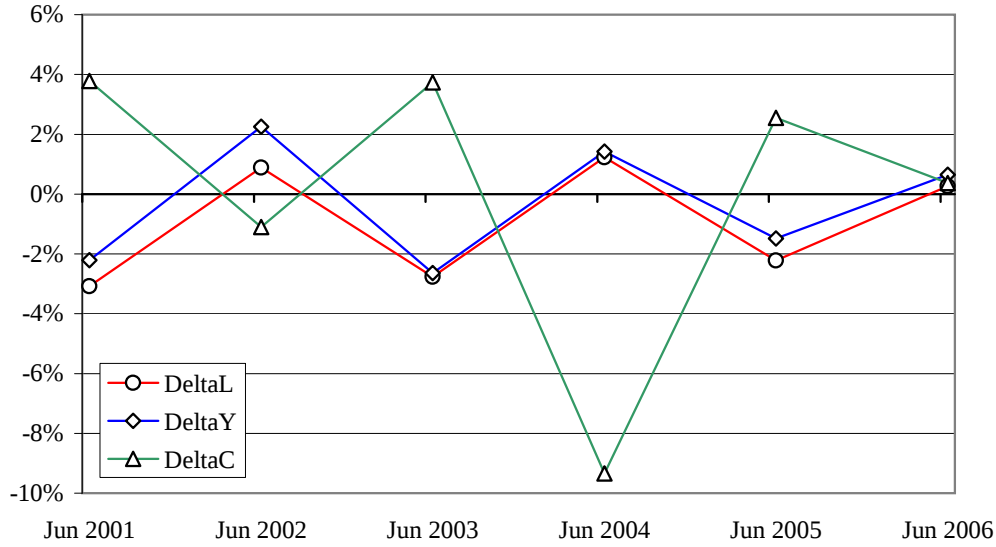


Figure 29: Annual absolute change for nominal, breakeven and CPI inflation rates

	ΔL_t	ΔY_t	ΔC_t	ΔL_{t-1}	ΔY_{t-1}	ΔC_{t-1}	ΔL_{t-2}	ΔY_{t-2}	ΔC_{t-2}
ΔL_t	1	0.97	-0.83	-0.98*	-0.96	0.71	0.96*	0.95	-0.70
ΔY_t		1	-0.74	-1.00	-0.97*	0.80	0.95	0.95*	-0.82
ΔC_t			1	0.79	0.71	-0.58*	-0.70	-0.71	0.26*

* indicates biased estimate

Table 9: Serial-correlations for the absolute change for nominal, breakeven and CPI inflation in rates, ΔL_t , ΔY_t and ΔC_t , shown in Table 8

8.3. Correlation significance tests

8.3.1. Table 10 shows the (two-side) p -value significance levels for the correlations in table 9 using two methods. The first assumes the ΔL_t , ΔY_t and ΔC_t time series are independent and identically distributed normal random variables, generates 50 000 realisations from this joint distribution and calculates the correlation coefficients for each realisation as for Table 9. The significance level of each correlation then corresponds to its percentile in the sampled distribution. The second method follows the same procedure but with 50 000 realisations bootstrapped from the sample in Table 9 with replacement. These bootstrapped significance levels are similar to those for the Ljung-Box statistic often advocated for testing serial-correlation (McNeil et al. 2006). The

serial-correlation in ΔL_t and ΔY_t is highly significant while the lower significance for ΔC_t can be largely ascribed to the outlying value for 2004 of -9.34% .

Resampled significance (i.i.d. normal assumption)									
	ΔL_t	ΔY_t	ΔC_t	ΔL_{t-1}	ΔY_{t-1}	ΔC_{t-1}	ΔL_{t-2}	ΔY_{t-2}	ΔC_{t-2}
ΔL_t		100.0%	2.2%	0.1%	0.5%	91.0%	99.1%	97.6%	15.0%
ΔY_t			4.6%	0.0%	0.2%	94.7%	97.6%	98.7%	9.0%
ΔC_t				89.3%	91.2%	18.2%	15.0%	14.4%	75.5%

Resampled significance (using sample with replacement)									
	ΔL_t	ΔY_t	ΔC_t	ΔL_{t-1}	ΔY_{t-1}	ΔC_{t-1}	ΔL_{t-2}	ΔY_{t-2}	ΔC_{t-2}
ΔL_t		99.7%	2.0%	1.4%	1.6%	92.9%	97.3%	95.6%	12.3%
ΔY_t			4.1%	0.0%	1.0%	95.6%	95.6%	97.6%	8.5%
ΔC_t				90.7%	92.6%	14.4%	12.3%	13.3%	73.9%

Table 10: Significance p -values for correlations in Table 9

8.3.2. A theory of mean reversion based on inflation targetting has been formulated before testing this theory. This theory-directed approach is safer than the pure data-directed approach where the data is used to generate theories (Huber 1997). Data-snooping, the weakening of statistical significance from producing and testing a theory using the same data or running multiple tests on the same data, is inevitable when analysing time-series data such as the dataset in Table 9.

8.3.3. At the further risk of confounding the statistical results due to inflation seasonality, the monthly inflation data are examined for stronger evidence of mean reversion. The correlogram for monthly lags is plotted out to 60 months for ΔC_t from March 2000 to March 2006 in Figure 30 with 95% confidence intervals. Here mean reversion is more evident.

8.4. Correction of serial-correlation bias

8.4.1. The biased serial-correlations marked with an asterisk in Table 10 stem from the use of the sample mean in the sample correlation coefficient estimate. Campbell, Lo & MacKinlay (1997) give a bias correction derived by Fuller (1976) for the k -th order serial-correlation coefficient from a sample of size T as

$$\tilde{\rho}(k) = \hat{\rho}(k) + \frac{T-k}{(T-1)^2} (1 - \hat{\rho}(k)^2)$$

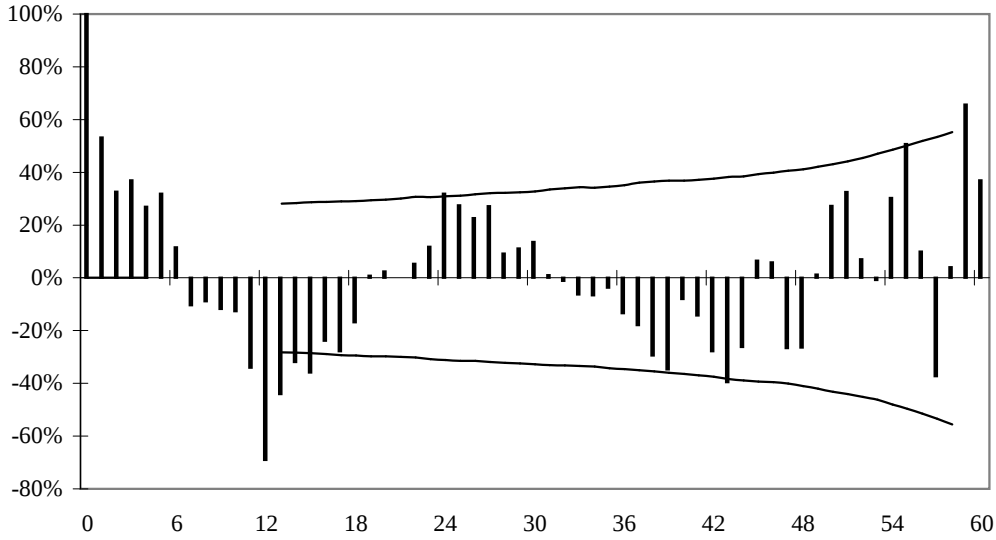


Figure 30: Correlogram for ΔC_t , by lag in months, showing 95% confidence intervals. Intervals for lags less than 13 months are biased from overlapping annual rates and so are not shown.

where $\hat{\rho}$ is the sample correlation coefficient. The bias in $\tilde{\rho}$ is then of $O(1/T^2)$. So, for example, in the sample of Table 8 we find $\tilde{\rho}(\Delta C_t, \Delta C_{t-1}) = -0.58 + \frac{1}{5}(1 - 0.58^2) = -0.45$. As a check, for a hypothetical $\hat{\rho} = 0$ estimate this adjustment is very close to the bias estimated by the resampling done above assuming independent and identically distributed normal increments. Table 11 shows Fuller bias-corrected serial-correlations for the biased estimates identified in Table 10.

	Resampled bias-corrected correlations (i.i.d. normal assumption)								
	ΔL_t	ΔY_t	ΔC_t	ΔL_{t-1}	ΔY_{t-1}	ΔC_{t-1}	ΔL_{t-2}	ΔY_{t-2}	ΔC_{t-2}
ΔL_t				-0.97			0.97		
ΔY_t					-0.96			0.97	
ΔC_t						-0.45			0.41

Table 11: Serial-correlations in Table 10 corrected for bias

8.5. Persistence of mean reversion

8.5.1. Until the 1970s the Phillips curve showed an inverse relationship between unemployment and earnings inflation (Phillips 1960). Since the critique by Lucas & Sargent (1978) of this trade-off, the prevailing economic view has been that price stability is favourable for economic growth. Money-supply targeting has been a dominant focus of monetary policy since then (Lipsey,

Steiner & Purvis 1987). New Zealand moved to inflation targeting in 1990 and by July 2004 21 countries had inflation targets (IMF 2005). The European Central Bank (ECB) has an inflation target while retaining a reference value for broad money supply as one of its pillars of monetary policy.

8.5.2. Price stability has been an economic goal of increasing importance in the past three decades, but there is no certainty the macroeconomic consensus will not change in future. Some commentators argue for a shift to broader targets for the stability of asset prices³. The South African Reserve Bank emphasises it will not respond mechanically to inflation shocks, but will exercise discretion “to avoid costly losses to output and jobs”⁴. For many central banks this discretion is formalised in an escape clause. Furthermore, government fiscal policy may become unsympathetic to inflation targeting.

8.5.3. To date central banks have generally been successful inflation targeters and no country has ever abandoned inflation targeting⁵. However it may take a long time, even relative to the term of long-term pension liabilities, for recovery from deflation or hyperinflation following extremely severe inflation shocks. It is therefore important to test a model for the effect of a break down in mean reversion.

8.6. Actuarial ALMs and mean reversion

8.6.1. Actuarial ALMs model relationships between the variables driving the values of assets and liabilities. They can be used to generate the assumed real-world ($\mathbb{P}$ -measure) behaviour of the prices of assets such as cash and real and nominal bonds. The primary aim of these models is realistic long-term behaviour of the market variables and output is not necessarily consistent with market forward-rates and traded non-linear instruments.

8.6.2. Of the ALMs surveyed in the literature review, the ARIMA framework models of Wilkie (1986), Wilkie (1995), Yakoubov et al. (1999) and Whitten & Thomas (1999) for the U.K. market and Thomson (1994) for the SA market exhibit serial-correlation for inflation and interest rate variables.

8.6.3. The Wilkie (1986) model has the following AR(1) process for annual UK price inflation:

$$I(t) = \ln(Q_t/Q_{t-1}) = QMU + QA[I(t-1) - QMU] + Z_t QSD. \quad (39)$$

³ The Economist, Economics Focus, 24 February 2005.

⁴ SARB Occasional paper 19, Inflation targeting in South Africa”, 2004.

⁵ Ibid.

where $\{Z_t\}$ is a sequence of independent and identically distributed standard normal random variables. Wilkie (1995) estimates strong positive autocorrelation of $QA=0.58$ for the model based on the period 1923-1994.

8.6.4. Thomson (1994) uses a similar approach for SA price inflation, but with a higher order AR(4) model. The first order autocorrelation parameter is 0.899 based on the period 1960 to 1993. However Maitland (1996) in his review of the model states that the oil shock of the 1970s “invalidates the assumption of weak stationarity and results in an AR(1) parameter which is too high”. A similar argument could be levelled at the Wilkie model, although this model was based on data from a longer period.

8.6.5. The highly positive AR(1) parameter and resulting mean aversion of the Thomson model inflation process cannot be reconciled with the statistically significant mean reversion of section 8.3 based on the 6-year period since inception of inflation targeting. Similar results might be obtained for the U.K. where inflation targeting has been successful since it began in 1992. The Huber (1997) review of Wilkie’s model rejects the hypothesis of a constant QA parameter at the 5% level. Beletski (2006) states that there is evidence against a constant mean reversion level in Euro inflation over the long-term.

8.6.6. These observations highlight the challenges long-term models face with lack of stationarity when fitting parameters to long historical periods with diverse economic circumstances. While mean reversion seems more defensible in the current environment of inflation targeting, any pricing model should address the possibility of positive or zero autocorrelation consistent with failure of inflation targeting. The results of De Gooijer & Vidiella-i-Angeura (2003) suggest that a self-exciting threshold autoregressive (SETAR) model that allows for switching between AR models is one way to reconcile the mean reversion of an inflation targeting regime with the persistence expected if targeting fails. Chan, Ng & Tong (2006) reach similar conclusions for UK price-inflation history spanning several centuries.

8.7. Mean reversion in other inflation models

8.7.1. Cairns (2004a) uses mean-reverting Ornstein-Uhlenbeck (O-U) processes as random drivers for his inflation and interest rates model set up within the Flesaker & Hughston (1996) framework. The financial economic models of Hull & White (1993) for interest rate derivatives and Jarrow & Yildirim (2003) for inflation derivatives also use O-U processes for mean reversion.

8.7.2. The O-U process has an SDE of the form

$$dr_t = \alpha(\mu - r_t)dt + \sigma dW_t;$$

where W_t is a Brownian motion and the parameters μ , σ and α can be dependent on time. An OU process with $\alpha > 0$ reverts to a long-term mean level of μ at a rate α such that the expected time until the distance between r_t and μ is halved is $\ln(2)/\alpha$. The OU process is discretised as

$$\Delta r_t = r_t - r_{t-1} = \alpha(\mu - r_{t-1}) + \sigma Z_t;$$

where $\{Z_t\}$ is a sequence of independent and identically distributed standard normal random variables. Hence the OU process is an AR(1) process with AR(1) parameter α .

8.7.3. The correlation structure between the increments Δr_t of an AR(1) process is

$$\rho(\Delta r_t, \Delta r_{t+h}) = \alpha^h \quad ; h = 0, 1, \dots \quad (40)$$

8.7.4. Cairns (2004a) uses several OU processes to model long-term nominal interest rates consistent with the variable patterns of mean reversion observed in markets over very long periods. The model is derived as an equilibrium model, but it may be possible to use a sufficient number of risk drivers that ensure the model recovers prices that are close enough to market levels to rule out arbitrage. The model includes real interest rates and inflation, although these are not mean-reverting.

9. Condition 3: Deterministic interest rates

9.0.1. The third assumption of the basic Black–Scholes world is deterministic real and nominal interest rates. While this may be a reasonable assumption for valuing short-dated options on an underlying which is independent of interest rates, this assumption is indefensible for inflation options. Furthermore, the link between nominal interest rates and inflation through inflation targeting is complex and may be unstable over time.

9.0.2. In continuous time, the drift of the spot inflation index Q_t depends on the instantaneous forward breakeven inflation rate $b(t, t)$ of equation (2), through:

$$Q_T = Q_t \exp \left[\int_0^{T-t} b(t, \tau) d\tau \right].$$

It is therefore unrealistic to expect breakeven inflation and the corresponding real rate to be static while spot inflation is volatile.

9.0.3. Models for pricing inflation derivatives can be extended to stochastic nominal and breakeven inflation rates, with correlation between the processes, provided these rates are also Itô processes. Van Bezooyen, Exley & Smith

(1997) (BES) derived the prototype model of HJM form for nominal and real rates and the spot inflation index. They consider short-rate models for real and nominal rates, such as the Ho–Lee and Hull–White extended Vasicek models as special cases of this model if there is no inflation index variability, and show they are compatible with the Black–Scholes currency model for inflation using volatilities scaled with time in a manner consistent with the model dynamics.

9.0.4. BES consider a Ho–Lee-type model with the following stochastic differential equations for the nominal short-rate r_f and real short-rate r_g :

$$\begin{aligned} dr_f(t) &= \alpha_f dt + \sigma_b dW^b(t) + \sigma_g(t) dW^g(t) \\ dr_g(t) &= \alpha_g dt + \sigma_g dW^g(t) \end{aligned}$$

Here $d\langle W^b, W^g \rangle = 0$. This model implies that short breakeven inflation and real rates are independent and the short breakeven inflation rate volatility is σ_b . BES show that this model leads to the value of a type-2 LPI liability equivalent to (18), but with volatility scaled to the power $\frac{3}{2}$ rather than the square-root of time:

$$\sigma \sqrt{(T-t)} = \sigma_b (T-t)^{3/2}. \quad (41)$$

9.0.5. Exley (2006) extends the BES version of the Hull–White extended Vasicek model further, giving valuation formulae for type-4 collared LPI liabilities and the inflation caps or floors used to calibrate the volatilities for the model. Exley notes that the zero-volatility inflation index assumption is equivalent to assuming that the actual inflation rate over the annual period between LPI increases is the same as forward breakeven inflation for the period. This assumption reduces the three-factor HJM model inflation model to two factors for real and nominal rates and simplifies calibration. However, this assumption is unrealistic for the same reason as currency models recognise random variation of the exchange rate is not fully accounted for by random variation in interest rate differentials. That is, under the risk-neutral measure,

$$dQ_t/Q_t = (f(t, t) - g(t, t))dt + \sigma dW_t^Q \quad ; \quad \sigma > 0, f \text{ and } g \text{ stochastic.}$$

9.0.6. Rogers (1997) argues that complete-markets models do not require an additional stochastic driver, W_t^Q , for the exchange rate. This apparent contradiction is resolved by noting that realistic models for breakeven inflation require at least two factors, and with the short breakeven inflation rate over infinitesimal time dt to some degree orthogonal to changes in longer dated breakeven inflation rates.

9.0.7. Figure 31 is a histogram of the daily absolute changes (in basis points) for the SA R189 2013 inflation-linked bond breakeven rate over the

two years to 30 June 2006. There is evidence of kurtosis in excess of a normal distribution, unlike the histogram for changes in the inflation index in Figure 26 (which was based on a much smaller sample of monthly changes).

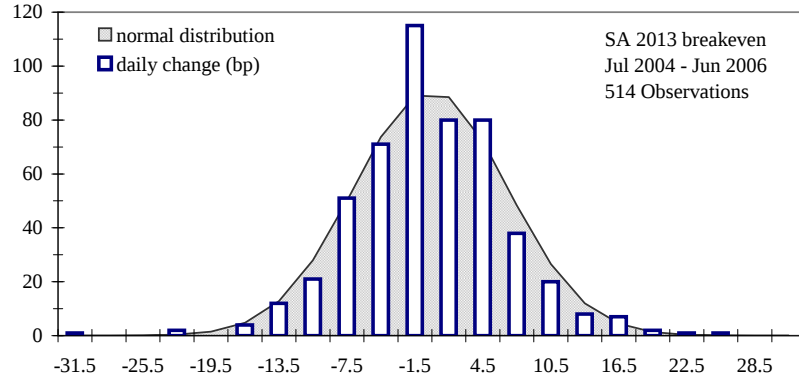


Figure 31: Histogram of daily absolute changes in SA 2013 bond breakeven

9.0.8. The systematic departure from normality is more evident in the normal Q-Q and P-P plots in Figure 32. Data points lying below and to the right of the model in the positive quadrant and above and to the left in the negative quadrant of the Q-Q plot suggest fatter tails than a normal distribution. This phenomenon of leptokurtosis is a feature of many financial markets. A cluster of zero daily moves is seen clearly in the P-P plot. The Jarque–Bera test emphatically rejects the hypothesis of a normal model for breakeven inflation rate changes at the 95% level, but not for changes in the CPI.

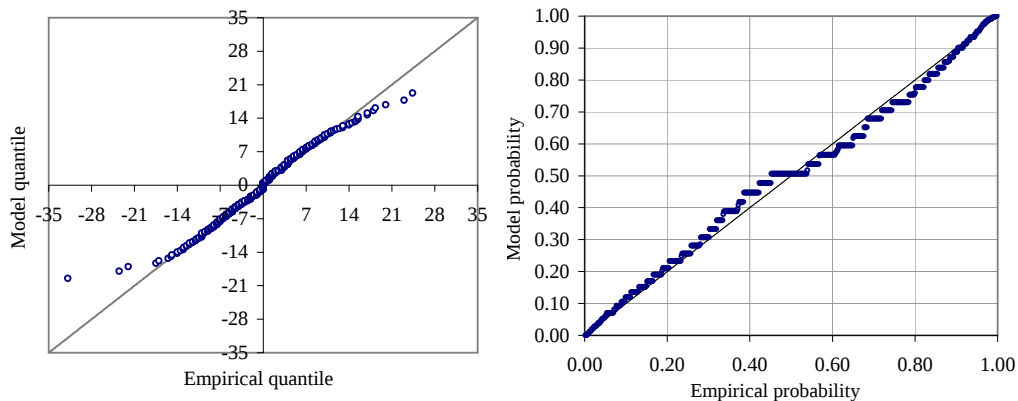


Figure 32: Q-Q and P-P plot of daily absolute changes in the SA 2013 bond breakeven rates shown in Figure 31

9.0.9. The general condition that nominal and breakeven rates are Itô processes admits a variety of terminal distributions for these rates. A model frequently used for nominal rates is the CEV model, so named for its constant elasticity of variance:

$$df(t, T) = \mu(f, t, T)dt + f(t, T)^\beta \sigma(t, T)dW_t. \quad (42)$$

This class includes normal ($\beta=0$) and lognormal ($\beta=1$) and the Cox Ingersoll Ross ($\beta=1/2$) processes as special cases.

9.0.10. Figure 33 plots the 3-month normal SA 2013 inflation-linked bond breakeven volatility in basis points per day against the breakeven rate level. This normal volatility is calculated using the absolute change in rate. The breakeven inflation rate plotted corresponds to the date in the middle of the 3-month volatility calculation period. The slope of the volatility as a function of breakeven level is an indicator of the CEV β coefficient. The chart and a regression of these variables shows no clear relationship, so a normal model may be the best model for breakeven inflation. This accords with the conclusions of Barclays Capital⁶ following analysis of other inflation-linked bond markets.

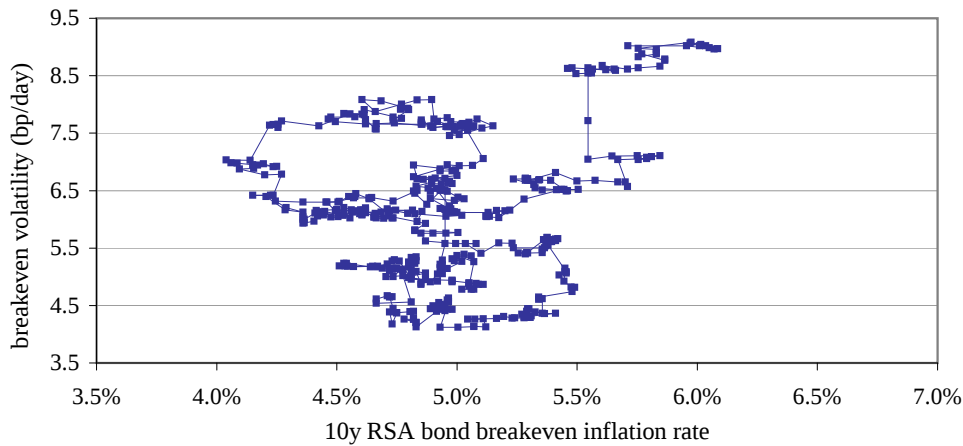


Figure 33: 3-month SA 2013 bond breakeven normal volatility versus breakeven inflation rate, two-year period to 30 June 2006.

⁶ Inflation Derivatives - A User Guide, Barclays Capital Research, January 2005

10. Condition 4: Deterministic volatility

10.0.1. In the Black–Scholes world the process for the underlying has deterministic volatility. This can be extended to an Itô process where the volatility is a function of the underlying. If interest rates are assumed to be stochastic, then these additional processes are also assumed to have volatilities and correlations that are deterministic or functions of the Brownian motions driving the processes.

10.0.2. Martingale pricing theory breaks down for processes with stochastic volatility or correlation because there is not a unique risk-neutral measure and the market is incomplete. Replication of derivative payoffs is not riskless and a method of selecting the pricing measure, such as utility theory, is required.

10.0.3. The fat tails for breakeven inflation in Figure 32 are evidence of stochastic volatility. In markets for nominal rates options, the implied volatility smile will reflect this dynamic and models can be calibrated accordingly. This has also become feasible for the UK and Euro inflation derivatives markets since dealers publish screens with prices of inflation caps and floors for a range of strikes. Where inflation options markets are not sufficiently developed to imply a volatility smile for calibration, such as in SA, it may be necessary to draw on historical data to estimate the effect of stochastic breakeven inflation volatility.

10.0.4. Figure 34 shows SA 2015 nominal and 2013 breakeven normal volatilities in bp per day, for the two years to 30 June 2006. This shows that breakeven and nominal rates volatilities are closely related and vary with time, as they do in other inflation markets. This is not surprising as breakeven inflation rates are the difference between nominal and real rates.

11. Financial market models for inflation derivatives

11.0.1. The first financial models for inflation derivatives were extensions to existing nominal rates and foreign exchange models. This makes it easier for dealers to accelerate product development of inflation products and allows for integrated risk management. These models focussed on the spot CPI, analogous to the exchange rate in a foreign exchange model. As the inflation swap and options market developed, forward CPI models evolved to incorporate specific aspects of the inflation process (such as correlation) that are essential to pricing and calibration.

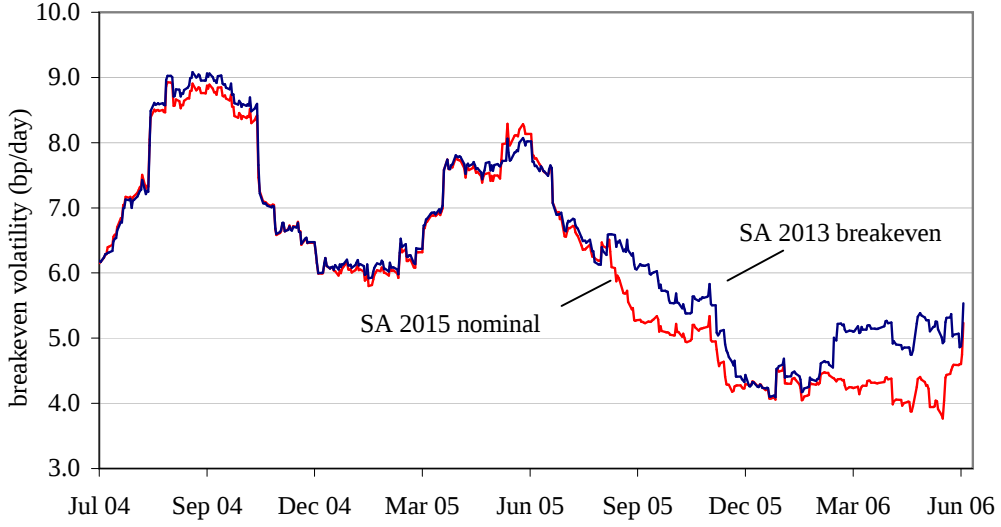


Figure 34: 3-month SA 2015 nominal and 2013 breakeven rates normal volatilities, two-year period to 30 June 2006.

11.1. Spot inflation index models

11.1.1. Jarrow & Yildirim (2003) (JY) published the first inflation and interest rates model in the finance literature, although Van Bezooyen et al. (1997) had previously produced a similar model. The JY model is a single factor model in the spirit of the Amin & Jarrow dual currency HJM model. The model describes the evolution of the spot inflation index and forward nominal and real rates with drifts adapted to the process and deterministic volatilities. The real-world dynamics of the model are:

$$\begin{aligned} df(t, T) &= \alpha_f(t, T)dt + \sigma_f(t, T)dW^f(t), \\ dg(t, T) &= \alpha_g(t, T)dt + \sigma_g(t, T)dW^g(t), \\ dQ(t)/Q(t) &= \mu_Q(t, T)dt + \sigma_Q(t)dW^Q(t). \end{aligned} \quad (43)$$

The three Brownian motions driving the processes are correlated.

11.1.2. Jarrow & Yildirim propose implementing the model with a constant inflation index volatility,

$$\sigma(t) = \sigma,$$

and time-homogeneous volatilities for nominal and real forward-rates of the form:

$$\begin{aligned} \sigma_f(t, T) &= \sigma_n e^{-a_n(T-t)}, \\ \sigma_g(t, T) &= \sigma_r e^{-a_r(T-t)}. \end{aligned}$$

11.1.3. This Hull White extended Vasicek implementation implies short nominal and real rates are mean reverting OU processes under $\mathbb{Q}$ with mean reversion rates a_n and a_r . The risk-neutral processes are then:

$$\begin{aligned} dn(t) &= (\mu_n(t) - a_n n(t))dt + \sigma_n d\tilde{W}_f(t) \\ dr(t) &= (\mu_r(t) - \rho(r, Q)\sigma_r\sigma_Q - a_r r(t))dt + \sigma_r d\tilde{W}_g(t) \\ dQ(t)/Q(t) &= (n(t) - r(t))dt + \sigma_Q d\tilde{W}_Q(t) \end{aligned}$$

where: $n(t) \equiv f(t, t)$, $r(t) \equiv g(t, t)$

$$\begin{aligned} \mu_n(t) &= \frac{\partial f(0, T)}{\partial T} \Big|_{T=t} + a_n f(0, t) + \frac{\sigma_n^2}{2a_n} (1 - e^{-2a_n t}) \\ \mu_r(t) &= \frac{\partial g(0, T)}{\partial T} \Big|_{T=t} + a_r g(0, t) + \frac{\sigma_r^2}{2a_r} (1 - e^{-2a_r t}) \end{aligned} \quad (44)$$

11.1.4. The drift for inflation under $\mathbb{Q}$ is therefore mean-reverting as it is the difference between the mean-reverting nominal and real short-rates. Jarrow & Yildirim developed the model for the bond market. Correlations and volatilities are calibrated to historical bond yield and inflation series and may be volatile. The model could be applied to the swap market, but inflation swaps trade on the breakeven inflation rate and calibration would not be as natural as for bonds that trade on real yields.

11.1.5. The Gaussian processes of the JY model make it possible to derive simple analytic values for a variety of inflation derivatives. Jarrow & Yildirim give closed-form values for zero-coupon inflation options (and hence type-2 LPI). Mercurio (2005a) gives formulae for year-on-year inflation swaps and options, Henrard (2005) for real rate European options and Malvaez (2005) discusses numerical procedures for valuation of Bermudan real rate swaptions.

11.1.6. A shortcoming of the JY model is that nominal rates can become negative. Mercurio (2005a) suggests adapting a lognormal LIBOR market model to include real rates. Nominal rates and real rates are then lognormal and guaranteed to be positive. The drift under $\mathbb{Q}$ for inflation is the difference between nominal and real rates, as in equation (44), but is not mean-reverting. The restriction of the real rate to positive levels and the skew of the lognormal process may distort the value of long-dated inflation liabilities more than negative nominal rates.

11.2. Forward inflation index models

11.2.1. Belgrade & Benhamou (2004b) (BBK) model the forward inflation index at each maturity T_i directly, rather than as functions of spot inflation

and real and nominal forward-rates,

$$dQ(t, T_i)/Q(t, T_i) = \sigma(t, T_i)d\tilde{W}_{T_i}^i(t).$$

where each $Q(t, T_i)$ is a martingale under its own T_i -year measure, $\mathbb{Q}_{T_i}$. A nominal rate model is integrated with this inflation model, its single Brownian driver having the same correlation with the $\{W^i\}$. This model is suited to developed inflation derivatives markets where the model can be calibrated directly to zero-coupon and year-on-year inflation swaps. Furthermore, for a time-homogeneous volatility specification the model can bound the zero coupon and forward year-on-year volatilities and correlations that determine the zero-coupon and year-on-year swap rates.

11.2.2. The BBK model uses the market model approach of modelling each rate maturity and draws an analogy between forward-rate and swaption volatilities in nominal space and year-on-year and zero-coupon index volatilities in breakeven space. The modelling of each discrete maturity gives the model a high dimension, although in practice the dimension is reduced to only a few factors and a limited number of parameters for the volatility and correlation functions. Zero-coupon and year-on-year inflation options are readily priced in the BBK model. Belgrade et al. (2004) give the prices of real-yield swaptions. It is straightforward to include inflation seasonality in model valuations.

11.2.3. The mean reversion characteristics of the inflation process in the risk-neutral world will depend on the volatility and correlation structure for the year-on-year and zero-coupon inflation rates. BBK show how the model can be calibrated consistently with a Hull–White mean-reverting process for inflation.

11.2.4. Even in the active Euro inflation market year-on-year swaps do not trade with anything like the frequency of zero-coupon inflation swaps. Volume is concentrated in maturities under 10 years, since year-on-year swap flow relates to structured notes purchased by individuals and smaller institutions and such notes are rarely issued with maturities longer than 10 years. It is questionable whether year-on-year swap rates can be used as a reliable calibration source, especially for long-dated inflation models such as those used to price pension liabilities. The main use of the model is by dealers pricing year-on-year inflation swaps in a market-consistent manner.

11.2.5. Mercurio (2005*a*) developed a model for the forward inflation index in a similar vein to the BBK model. Mercurio (2005*b*) compares the prices of 0% floors on the year-on-year inflation rate for the JY model, his LIBOR and real rate market model of section 11.1.6 and this forward CPI market

model calibrated to the same set of zero-coupon inflation swap rates and at-the-money nominal cap volatilities. Unsurprisingly, the different dynamics of the models lead to significantly different floor prices and these deterministic volatility models cannot reproduce the market-implied volatility smile.

11.2.6. Mercurio & Moreni (2006) describe a stochastic volatility extension to the forward CPI market model in the form first proposed by Heston (1993). Under the assumptions of constant nominal volatility and a common mean-reverting volatility for all forward inflation processes with a constant mean reversion rate and level, they derive an approximation for the prices of options on year-on-year inflation that exhibit a volatility smile. The assumption of a common volatility process for all forward inflation indices is very restrictive as it implies the increase in volatility of zero-coupon inflation swap prices due to increased duration is precisely offset by the decrease in the term structure of volatility of the inflation swap rate. The authors acknowledge that the assumption of a common volatility process “seems too restrictive if we aim at calibrating many maturities simultaneously”.

11.3. Risk-neutral and real-world inflation dynamics

11.3.1. Martingale pricing theory dictates that the price of an inflation derivative is the value of its expected payoff under the measure $\mathbb{Q}$ corresponding to the risk-neutral world. The drift under $\mathbb{Q}$ of all instruments with prices dependent on the inflation model variables is the instantaneous risk-free rate and so the real-world drift has no bearing on the value of inflation derivatives. The analysis of the real-world dynamics of the inflation process in section 8 is therefore relevant only as far as it illuminates the volatility of the process. This volatility is measure-invariant for the Black-Scholes world models considered earlier.

11.3.2. The term structure of market-implied volatilities for nominal rates will imply the rate of mean reversion in the risk neutral process of a time-homogeneous model calibrated to the nominal options market. For example, a volatility term structure in the Hull–White extended Vasicek form, $\sigma(t, T) = \sigma e^{-a(T-t)}$, for nominal forward-rates implies that rates follow an OU short-rate risk-neutral process with mean reversion rate a .

11.3.3. The HJM completeness and no-arbitrage conditions restrict the real-world drift for a given volatility term structure. This connects the real-world mean reversion with the risk-neutral world, even although the real-world dynamics are not required for risk-neutral derivatives pricing. Where there is no unique risk-neutral measure it may be considered desirable to limit the implied market inefficiency of the model. Smith (2001*a*) discusses how inefficiency

increases with the difference between the real-world and risk-neutral densities for the asset price process. This suggests real and risk-neutral measure mean reversion dynamics should be aligned to some extent.

CHAPTER 6

Background risks

This chapter addresses the impact on the market valuation of pensions of risks that cannot be hedged using market instruments. The relevant background risks for pensions in payment are longevity, credit risk from the sponsor in respect of unfunded liabilities, regulatory risks and tax risks.

The market valuation of pensions subject to background risk is, in essence, the problem of pricing in an incomplete market. A number of valuation methods have been proposed for incomplete markets liabilities. These all require that the background risks are modelled with a level of accuracy commensurate with their significance and their relationship with the tradable risks.

The primary background risk for pensions in payment is longevity risk. Market developments, such as mortality derivatives, may in time allow longevity risk to be transferred to the optimal counterpart and reveal the market cost of transferring the risk. Stochastic models for longevity risk that have been proposed in the literature are analysed. It is considered how a market valuation might be assigned to the liabilities using a risk-adjusted valuation approach.

The chapter concludes with a summary of the key considerations behind the market valuation of pension risks. There are a number of interesting parallels between inflation and longevity that aid overall understanding of these risks, although the idiosyncratic nature of pensions mortality risk makes it difficult to incorporate in a general valuation model. The key challenge for a risk-adjusted valuation is the estimation of the market price of background risks. Fortunately, these valuation variables are becoming more transparent as the capital and insurance markets develop and converge.

1. Background risks for a DB pension fund

1.0.1. A pension fund that uses market instruments to hedge its projected liabilities will still be faced with inevitable residual risk which cannot be hedged or diversified away. There may be a price for which another entity will take over the liabilities, but this does not necessary allow us to infer a market valuation for the liabilities (see chapter 3).

1.0.2. These residual risks have been termed “background risks” in the economics literature (Cardinale et al. 2006). The actuarial literature recognises that longevity risk is key risk for pension funds. However, since the market for longevity is risk is nascent and incomplete, it will be considered as a background risk. For a DB pension fund, background risks then include:

- (1) longevity risk;
- (2) credit risk from the sponsor;
- (3) regulatory and tax effects;
- (4) decrements such as early retirement, withdrawal, etc; and
- (5) basis risk between salary and CPI inflation for liabilities in respect of active members of a final salary pension fund.

We restrict our attention to the first three risks, as the ambit of this research is the market valuation of pensions in payment.

2. Valuation of background risks

2.0.1. Cardinale et al. (2006) survey the economic theory of background risks. The interaction between background risks and tradable risks, such as covariation and whether background risks are multiplicative or additive, can affect the degree of tolerance for these risks.

2.0.2. Pension liabilities subject to background risks are an incomplete market in financial economics terms. Earlier it was noted that there are elements of the investment risk which are also not amenable to perfect replication. There is no unique risk-neutral measure for incomplete market liabilities. However, it is possible to assign a market value to the liabilities using one of the incomplete-markets valuation approaches discussed in the literature survey.

2.0.3. Regardless of the valuation approach, the background risk processes should be modelled as accurately as their significance dictates in order to generate a sound and robust measure of market value. The interaction between background risks and market risks should also be modelled with care since the valuation may be sensitive to the form of this interaction.

3. Longevity risk

The longevity risk of pensions in payment is the risk that mortality experience is lighter than anticipated in the mortality projection. The trend of mortality improvements in recent times highlights the significance of this risk. In the UK, the recent CMI (2004) working paper showed that the CMI central mortality projections for life-office annuity business in the past 35 years only accounted for only about 50% of the actual mortality improvement.

3.1. Longevity hedging with insurance and reinsurance

3.1.1. Milevsky & Promislow (2001) model longevity risk assuming an insurer is able to hedge the risk by writing life insurance policies. However, life protection products provide a limited longevity hedge since the causes of death for insureds differ from those of annuitants and pensioners due to differences in age at death (Richards & Jones 2004). In SA this difference may be exacerbated by the impact of AIDS. In the US 10% of the life insurance market is on lives aged 70 or older (Ortiz, Stone & Zissu 2006), but the capacity for longevity hedging in respect of these lives falls short of the longevity risk of private and state pension liabilities for this age group.

3.1.2. Olivieri (2002) discusses how various proportional and excess-of-loss life reinsurance arrangements can be used to reinsure longevity risk and the related investment risks of annuities. A dynamic reinsurance hedging programme can then, in theory, be used to complete the pension and life annuity market. In practice reinsurers may be reluctant to assume undiversifiable mortality risks and a limited volume of annuity reinsurance capacity is available. In the UK less than 5% of longevity market is reinsured compared with more than 50% of the life assurance market and some reinsurers view longevity risk as ‘too toxic’ to underwrite (Wadsworth 2005).

3.2. Mortality derivatives

3.2.1. Derivatives on mortality or survival rates or indices may be structured to assist the risk management of insurers. The mortality derivatives market has been slow to develop as a result of the complexity of the risk and lack of natural two-way interest. Various contracts have been proposed in the literature:

- ‘Survivor swaps’ exchange a mortality swap leg linked to a survival index for a fixed leg linked to the mortality projection that equates the expected present value of the legs at each future reset date. Survivor swaps are analogous to zero-coupon inflation swaps, where the survival index takes the place of an inflation index. Indeed, combining

the survival index with an inflation-linked notional will be preferable for real annuity insurers and many pension funds. Dowd et al. (2006) report that over-the-counter (OTC) survivor swaps have been transacted between reinsurers. Lin & Cox (2007) discuss how a combination of swaps can be used to hedge diversifiable and (to an extent) systemic risk between life insurance and annuity portfolios.

- ‘Mortality options’ have a payoff based on a mortality rate or index at maturity. Guaranteed annuity options (GAO) granting policyholders the right to enter into life annuities at predetermined rates were common in the UK in the 1980s. These options subsequently caused large losses for insurers who failed to hedge the longevity and interest rate risk.
- ‘Annuity futures’ could be envisaged, with the underlying based on an insurer poll of annuity rates analogous to the dealer poll used to set a market swap rate for interest rate derivatives. Blake et al. (2006) state that the Association of French Pension Funds is considering such a future. They describe the practical difficulties posed by the contract, including the heterogeneity of life annuity policies, lack of incentives for insurers to quote rates and the slow reaction of annuity rates to changes in longevity.

3.2.2. Mortality derivatives may be embedded within physical investments to produce mortality-linked bonds with risk and return characteristics attractive to both issuers and investors. Insurance securitisations, outlined in chapter 3, provide an ever-expanding source of market price information for mortality risk. Securitisations to date have concentrated on elements of life insurance assets and liabilities that convey limited information on the mortality risk of pensioners.

3.2.3. The Swiss Re ‘Vita’ five year mortality-linked bond issues were fully subscribed¹. The reinsurer in effect purchased options on a predetermined index of population mortality for five countries to hedge part of its mortality risk. The bonds have been described as catastrophe bonds since the strike is set at a level that would normally only be breached by deaths caused by a distinct event, such as a disease epidemic, terrorist attack or natural catastrophe. Therefore the bond would only provide a limited longevity hedge for pension annuity risk in these countries.

¹ “Swiss Re successfully closes its second life catastrophe bond”, Swiss Re, 2005.

3.2.4. Some securitisations, such as the Norwich Union securitisation of equity-release mortgages in the UK² have been promoted as longevity risk hedges. However, the link between mortality and other risks is complex and Richards & Jones (2004) contend that these securitisations “might be felt to be unsuitable for say, backing annuity liabilities, as this would be doubling up of longevity risk on both the asset and liability sides”. The Tarrytown securitisation of US senior life settlements (with life expectancy of insureds under 12 years) is a fairly clean transfer of longevity risk (Ortiz et al. 2006) for old lives, but these lives are generally impaired. The bonds would need to be sold short to be used as a longevity hedge and the market is as yet too thin to do this.

3.2.5. The European Investment Bank (EIB) longevity bond arranged by BNP Paribas did not achieve sufficient investor interest for the issue to proceed. The bond was of the form of a 25-year nominal annuity with payments linked to the survival index for a 65-year-old (in 2003) male life based on England and Wales population mortality statistics. The bond can be viewed as the combination of a nominal annuity bond and a survivor swap. The reinsurer underwriting the longevity risk for the EIB, Partner Re, emphasised that it had limited capacity for further deals. Cairns, Blake, Dawson & Dowd (2005) consider why the bond issue was unsuccessful. While the bond would help hedge longevity risk for UK pension funds and life insurers, these institutions may find it more cost-effective to hedge the risk directly with a reinsurer (subject to the capacity constraints addressed in section 3.1).

3.2.6. It is an open question whether there are natural counterparts for longevity risk. Furthermore, the heterogeneity of longevity risk frustrates attempts to construct derivatives to transfer the risk effectively. Blake et al. (2006) discuss why basis risk between the longevity experience of a portfolio (such as DB pensioners) and a mortality derivative may be unavoidable. The alternative bases for defining a survivor index are:

- (1) national population experience;
- (2) industry experience (e.g. SA CSI, UK CMI studies); or
- (3) portfolio experience.

3.2.7. A population mortality index does not expose either counterpart to asymmetric information, but there may be a high degree of basis risk if the reference population experience is a poor proxy for the portfolio. Industry experience for relevant insurance contracts (e.g. life-office annuities) may be more representative, but this takes longer to analyse and there may be the

² “Life Company Funding: A Hybrid Future?”, Life and Pensions, December 2005.

perception of moral hazard if significant industry participants can trade on their advance knowledge of the results (Blake et al. 2006). The payoff of the mortality derivative may be linked to the experience of the specific portfolio in order to eliminate basis risk, but this is at the cost of the standardisation needed to concentrate potential liquidity in the contract. The market will be limited to OTC deals and antiselection may limit longevity risk takers to counterparts with underwriting expertise (such as reinsurers).

3.2.8. Publication delays to the mortality index increase basis risk and moral hazard, but delays are inevitable given the time required to collect and graduate mortality experience. The relative effect of the mortality indexation lag is reduced for the long-term deals needed to hedge pension liabilities. The Swiss Re and EIB mortality-linked securities base payments on population mortality experience, for which there is a publication delay of a year or more after the end of the exposure period. A bank dealing mortality derivatives has created a longevity index based on US population experience to facilitate prospective mortality derivative trades³.

3.2.9. In summary, while mortality derivatives and reinsurance may allow a pension fund to hedge part of its mortality risk through diversification, exact hedging of longevity risk is not possible without an insurer buy-out of the liability. This weakens the case for using pricing techniques designed for complete markets to value pension liabilities without consideration of the market price of longevity risk.

3.3. Classes of longevity risk models

3.3.1. The impetus for the stochastic mortality models developed in recent years is from regulation (e.g. the Financial Services Agency Integrated Prudential Sourcebook in the UK) and actuarial professional guidance (e.g. Actuarial Society of SA guidance for internal mortality models of SA life insurers). To model the mortality process it is useful to consider the following fundamental sources of mortality risk:

- stochastic variation in the number of deaths each year;
- mortality rate projection risk; and
- mortality model and parameter risk.

3.3.2. The stochastic variation in deaths is diversifiable and the risk can be managed through the law of large numbers. To simplify analysis it is usual to assume that deaths are independent events. The number of deaths over a period τ for a pool of n homogeneous lives experiencing (initial) mortality

³ “CS Longevity Index Commentary, 1 March 2006”, Credit Suisse.

rate q is then a Binomial($n, q\tau$) random variable. If n is large this can be approximated as a Poisson process with mean $\lambda = nq\tau$.

3.3.3. The assumption of independence is not likely to be true since some deaths will be attributable to a single event (such as a disease epidemic), but it is nonetheless reasonable in practice. The mortality experience over a short period for pensioners is dependent in a different way to that of active members who may work together and be exposed to a common catastrophic risk (such as the New York September 2001 terrorist attack).

3.3.4. The assumption of homogeneous lives is reasonable only if the pool is broken down into subgroups by risk factors such as age, sex, amount of pension, etc. Many pension funds have a small number of retired lives and relatively high stochastic variation in the number of deaths. Heterogeneity of risk factors will further exacerbate this stochastic variation. CMI (2006) observes that, for smaller pools (fewer than 5000 lives), stochastic variation due to heterogeneity in annuity size can outweigh mortality projection risk. Reversionary benefits and differences in pension indexation (e.g. higher indexation for retired managers) will exacerbate heterogeneity.

3.3.5. It is usual to model experience over successive periods of $\tau = 1$ year. If more precision regarding the timing of deaths is required, e.g. when pension benefits are paid monthly, deaths may be assumed to be uniform over the year. There will be seasonal fluctuation in mortality experience (e.g. more deaths in winter when pensioners are more susceptible to flu, etc) but this will be minor compared with the stochastic variation in the numbers of deaths.

3.3.6. The second and third risk sources are systemic and cannot be diversified away. The mortality rate projection is the best estimate of future mortality experience for the pool of lives considered and can be viewed as a parameter of the stochastic mortality model. We define parameter risk as the uncertainty relating to this and other parameters in the model. Model risk is the risk that the model has been incorrectly specified.

3.3.7. Cairns, Blake & Dowd (2006) provide a comprehensive survey of the published stochastic mortality models. They class models as follows:

- (1) models that project the force of mortality for each age (or cohort or both) and future time based on past experience;
- (2) models of random variation in discrete time around a deterministic mortality projection; and
- (3) mortality models based explicitly on the risk-neutral pricing paradigm for interest rates and credit.

3.4. Published stochastic mortality models

3.4.1. *The Lee–Carter model.*

The Lee & Carter (1992) model assumes a force of mortality for lives of age x at future time t of

$$\log \mu(x, t) = a(x) + b(x)k(t) + \epsilon(x, t)$$

where $\epsilon(x, t)$ is an error term. The parameters $a(x)$, the general mortality level and $b(x)$ and $k(t)$, the age-specific variation, are constrained to give a unique solution. This model is used by the Bureau of Census in the US, where it explains almost all the US population mortality variation with time (Haberman & Verrall 2006). The model is less effective for UK data as it cannot capture the cohort effect, where the population born in the decades around 1935 has experienced consistently higher mortality improvements than other generations (Willets 2006). Stochastic mortality projections are based on modelling $k(t)$ as a time series.

3.4.2. *P-spline models.*

CMI (2006) discusses the development of two public stochastic mortality models using the Lee–Carter method and the Penalised splines (‘P-splines’) method of Currie et al. (2004). The idea behind P-splines is similar to the spline curves used to graduate mortality curves. Basis splines are fitted in two dimensions (e.g. by age and data period or by age and cohort) over historic data and the projection period using regression or maximum likelihood and a penalty function. The penalty function weighs goodness of fit (to the historical data) against smoothness (over the data and projection range). The model does not define a stochastic process and so cannot generate sample paths of future mortality experience, unlike the Lee–Carter method. Confidence intervals for projected mortality rates for P-splines will include parameter uncertainty. The parameter uncertainty for the Lee–Carter method can only be estimated indirectly, for example by using the parametric bootstrap method of resampling from an assumed distribution for the parameters and refitting the model. In this sense the methods are complementary.

3.4.3. *Parametric mortality curve models.*

Cairns, Blake & Dowd (2005) model a parametric mortality curve in discrete time using a stochastic mortality curve of Perks form:

$$\tilde{q}(t, x) = \frac{\exp(A_1(t+1) + A_2(t+1)(x - t))}{1 + \exp(A_1(t+1) + A_2(t+1)(x - t))}. \quad (45)$$

Here $\tilde{q}(t, x)$ is the realised initial mortality rate and $A_1(t)$ and $A_2(t)$ are correlated random walks with drifts, which can be interpreted as the general and age-specific expected rates of mortality improvement. The initial curve parameter values $A_1(0)$ and $A_2(0)$ are based on an initial mortality projection, so this model is of the second class identified above. Other parametric forms for the stochastic mortality curve have been proposed: Olivieri (2001) uses a Heligman–Pollard law, Milevsky & Promislow (2001) and Ballotta & Haberman (2004) use a mean-reverting Brownian-Gompertz law.

3.4.4. *The Olivier–Smith model.*

Olivier & Smith (2007) model $\ell(s, t, x) = \mathbb{E}[\ell(t, x) | \mathcal{F}_s]$, the expected number of survivors at time t for a cohort of initial age x , given an initial survivor projection $\ell(0, x)$ and the filtration $\mathcal{F}_s$ of deaths to time $s \leq t$. This set-up enables the model to incorporate both stochastic variation in the number of deaths and stochastic variation in the mortality rate around the mortality projection. The number of cohort deaths each year conditioned on the number of cohort survivors at the start of the year is assumed to follow the binomial distribution:

$$\ell(s, t, x) \sim \text{Bin}[\ell(s, t-1, x), p(t-1, t, x)].$$

Here $p(s, t, x)$, the probability of survival to time t of a life in the cohort with initial age x conditional on $\mathcal{F}_s$, has a Gamma deterioration factor $G(s)$ with mean 1 and variance v :

$$p(s, t, x) = \left(\frac{\ell(s-1, t)}{\ell(s-1, t-1)} \right)^{\beta(s, t)G(s)} \quad \forall t \geq s \quad \text{where } G(s) \sim \text{Gamma}[1/v, 1/v].$$

The constant $\beta(s, t)$ is a bias correction factor to ensure that the expected number of survivors is equal to the initial survivor projection $\ell(0, x)$. The number of survivors divided by the projection is therefore a martingale and the bias correction factor is analogous to the risk-neutral drift adjustment in an interest rate model.

3.4.5. Olivier & Jeffery (2004) justify a Gamma deterioration factor on the grounds that it is positive and so ensures that $p(s, t, x) \leq 1$ and the bias correction factor is tractable. For UK CMI data, the standard deviation of $G(s)$ is $\sqrt{v} \approx 5\%$ and a truncated normal deterioration factor would give similar results as this gamma distribution is close to symmetric. Dowd et al. (2006) use a transformed Beta variable for the deterioration factor for US population mortality. They use a standard deviation for the deterioration factor of 2.2%.

3.4.6. The Olivier–Smith model uses the same deterioration in a given calendar year for all cohorts. Cairns (2006) generalises the model using a Gaussian copula over different Gamma deterioration factors for each age cohort. This gives a better fit to UK population mortality since 1960, since it allows age-specific mortality improvement.

3.4.7. *Interest rate and credit models for mortality.*

Cairns, Blake & Dowd (2006) show how stochastic mortality models can be built from a risk-neutral interest rate model framework. Short-rate models describe the behaviour of the spot force of mortality, usually in continuous time. Dahl (2004) considers conditions for which the force of mortality μ_{x+t} gives rise to an affine structure for survival probabilities, i.e.

$$p(s, t, x) = \exp[A(s, t, x) - B(s, t, x)\mu_{x+t}] \quad \text{under } \mathbb{Q}$$

3.4.8. Miltersen & Persson (2005) and Cairns et al. (2006) discuss the conditions for a forward mortality rate framework analogous to HJM to be free of arbitrage. The system is complicated by the introduction of age as an additional dimension, but the conditions are similar to HJM if the Brownian drivers apply to all age cohorts. Cairns (2006) notes that mortality-linked securities dependent on survival probabilities, such as survivor swaps and longevity bonds, are more readily priced within a short-rate framework. By contrast, guaranteed annuity options depend on forward mortality rates and may be more tractable in a forward mortality rate model.

3.4.9. The positive interest framework of Flesaker & Hughston (1996) is a useful basis for mortality models since it ensures that the force of mortality remains positive at all times. Cairns et al. (2006) discuss how such a model might be formulated and calibrated.

3.4.10. Milevsky & Promislow (2001) use the hazard-plus-interest credit modelling theory of Duffie & Singleton (1997) to model the spot force of mortality and the short interest rate, assuming these processes are independent.

3.4.11. Cairns et al. (2006) also develop an analog of the (annual) LIBOR market model that they term the SCOR (survivor credit offer rate). The SCOR depends on the age of the cohort and is equivalent to the annual mortality drag for a retired life in the cohort (i.e. the investment return net of the risk-free rate the life must earn to justify not annuitising).

3.5. Stochastic mortality model considerations

3.5.1. A stochastic mortality model must produce a strictly positive mortality rate at all future times. Furthermore, Cairns et al. (2006) argue that a

model should produce a spot mortality rate at each future time that is monotonically increasing at advanced ages in order to be ‘biologically reasonable’. This is a considerable challenge for models where multiple factors are employed to allow different rates of mortality improvement with age. For example, the copula over different age-cohort deterioration factors used in the Cairns extension to the Oliver–Smith model must be formulated to ensure that this condition is met. Cairns (2006) also notes that very few biologically reasonable short-rate mortality models have a simple analytical form for the survivor bond price as a function of the current short mortality rate.

3.5.2. A more contentious point, also advocated by Cairns et al. (2006), is that the stochastic variation about projection of mortality improvements should not be strongly mean-reverting. They argue that unanticipated mortality improvements should not reduce the potential for further future improvements. This may be the case if better medical technology is responsible for the improvements, although to the extent there is a fixed maximal human life span there may be mean-reversion at very advanced ages.

3.5.3. A related concept is the vitality effect, so called by Olivier & Jeffery (2004) since high mortality in one year implies higher than anticipated mortality in the next year because the surviving lives have been weakened by the initial cause of the high mortality. The opposite herd effect is when high mortality in one year implies lower mortality in the next year when it is the weaker lives that were initially lost. The Olivier–Smith model assumes that the number of cohort deaths in each year is independent of previous years, since these deaths are independent binomial-Gamma mixture variables. It is not obvious how herd or vitality effects could be included through serial correlation in the deterioration factors while retaining a tractable bias-correction factor.

3.5.4. Mortality studies must follow the experience of a cohort of lives to judge whether there are vitality or herd effects, against the alternative hypothesis of independence. This may be difficult in industry studies unless the data from all contributors is sufficiently detailed and standardised.

3.5.5. Cohort studies also eliminate the risk of confounding the mortality experience with changes over time in the nature of the population studied. For example, a study of SA life-office annuitants may not control for all the changes in risk factors from an increasing number of individuals funding their own retirement annuities. Here past experience may not be a good guide to future and stochastic mortality models with more control over expected mortality

improvements may be preferred over those which project past experience (such as P-splines).

3.5.6. Willets (2006) discusses projections based on causes of death (COD). This approach is useful for understanding short-term trends in mortality but the nature of longevity improvement is complex and dynamic and it is unlikely to be a COD stochastic mortality model will be helpful for long-term pension liabilities. COD models may be of use for determining the extent of mortality delta offset between annuitants and insureds over short periods.

3.5.7. The assumption of independent deaths within each year simplifies a stochastic mortality model, giving it flexibility to address the long-term variation in longevity that is significant for pension liabilities. However, it may be considered necessary to include low-frequency, high-severity random catastrophic mortality risk in the model. Klein (2003) points out that a single catastrophe could undo the effect of decades of improvements. The 1918 influenza pandemic would decrease the CS Composite Longevity Index (for the entire US population) by 4.4 years, equivalent to 5.4 years of mortality improvement at the current rate⁴.

3.6. Relationship between longevity and interest rates

3.6.1. The relationship between longevity and the (real) liability discount rate may have a significant effect on the risk of the liability cashflows and hence the market value. Most actuarial models explicitly or implicitly assume no correlation (Richards & Jones 2004). This facilitates the valuation and mitigates the multiplicative effect of longevity as background risk, but perfect matching is not possible and the market cannot be treated as complete.

3.6.2. There has been little research into the relationship between longevity and interest rates. A report commissioned by the Association of British Insurers⁵ suggests there is no clear relationship since there are a number of opposing factors:

- Wealth effects. To the extent that better medical care outweighs diseases of affluence (such as heart disease from obesity), there will be a positive net wealth effect on longevity. To the extent that property and equity wealth gains are realised and invested in bonds to match the increase in pension liabilities, longevity and long-term interest rates may be negatively correlated. This effect will only occur gradually with time. Miltersen & Persson (2005) point to the relationship

⁴ “CS Longevity Index Commentary, 1 March 2006”, Credit Suisse.

⁵ “Report for the ABI on key correlation assumptions in ICA for life offices”, Deloitte and Touche LLP, 2005.

between general economic development and longevity. In time, the development of the SA economy may see long-term interest rates converge to the lower levels of developed markets and longevity converge to the higher levels of more developed markets. This scenario would imply a negative correlation between longevity and yields.

- **Demographic shifts.** An increase in longevity combined with low fertility rates will increase the proportion of retirees drawing state pensions. This will increase the state borrowing requirement unless action is taken to counteract this (e.g. by increasing the age at which the state pension commences). The decrease in the proportion of the population in employment may preclude the use of increased taxation to fund these liabilities. Since most state pension commitments are not prefunded, this will result in a net increase in government debt and higher long-term real and nominal interest rates. This effect may be offset or outweighed by increased private pension investment in bonds to match the increased liabilities of retirees. SA public-service pension liabilities are prefunded through the PIC so there will not necessarily be a net increase in government debt. If there is net ageing within the relatively large private pension sector, interest rates may well decrease in response to an increase in longevity. In SA the effect of AIDS may further complicate the relationship between longevity and interest rates.
- **Catastrophes.** A catastrophic mortality event will decrease longevity and may result in initial risk aversion and a flight to quality assets and possibly a prolonged period of lower growth. Both consequences imply a positive correlation between longevity and long-term interest rates, although emerging markets such as SA have frequently experienced higher long-term interest rates in periods of risk aversion.

3.6.3. On balance it is unclear how what relationship longevity will have to long-term yields, making it difficult to calibrate a model for pricing pension liabilities.

3.7. Valuation of pension liabilities with longevity risk

3.7.1. Earlier it was noted that the valuation of pension liabilities subject to longevity risk is an incomplete markets valuation problem. Scholarly research to date regarding the pricing of mortality-contingent instruments in an incomplete market has relied on diversification of background risks (for example Thomson (2005)) or applied a risk-adjusted pricing approach. The latter

approach values the instrument under the risk-adjusted pricing measure $\mathbb{Q}$ so that the drift adjustment to change measure from $\mathbb{P}$ to $\mathbb{Q}$ is the estimated market price of risk. This market price of risk, λ , may be a vector and time-dependent. It compensates the instrument's holder for risk that cannot be perfectly hedged in the incomplete market.

3.7.2. By contrast, the change of measure from $\mathbb{P}$ to $\mathbb{Q}$ in a complete market with a single source of uncertainty corresponds to a drift adjustment of $\frac{\mu-r}{\sigma}$ where μ is the drift under $\mathbb{P}$ of the instrument and r the risk-free short-rate. The resulting drift for all tradables (standardised by the cash account or other tradables) under $\mathbb{Q}$ is zero in this risk-neutral world.

3.7.3. Lin & Cox (2005) estimate the market price of longevity risk from the cost of US life annuities. They use the risk-adjusted pricing transform of Wang (2002) to infer the market price of longevity risk from market annuity rates given assumptions about the insurer's mortality table and expense loading. The estimated market price of risk of 0.18 for 65-year-old males and 0.23 for 65-year-old females assumes the market price of risk is constant with time and implicitly includes all sources of mortality risk and the correlation with investment risk. It would be possible to find the market price of risk as a function of time by bootstrapping from the prices of annuities commencing at each age after 65, after stripping out assumed expense loadings.

3.7.4. The Wang transform distorts the real-world density of a payoff to generate a risk-adjusted pricing density consistent with market prices. The magnitude of this distortion is the market price of risk. Let λ denote this market price of risk, $F^{\mathbb{P}}$ denote the real-world cumulative distribution function (cdf) and $F^{\mathbb{Q}}$ denote the risk-adjusted cdf. The Wang transform from $F^{\mathbb{P}}$ to $F^{\mathbb{Q}}$ is then:

$$F^{\mathbb{Q}}(x) = \Phi \left[\Phi^{-1}(F^{\mathbb{P}}(x)) + \lambda \right], \quad (46)$$

where Φ is the standard normal cdf.

3.7.5. The Wang transform was originally developed to price different excess-of-loss reinsurance layers consistently. Cox, Lin & Wang (2005) report that the transform is consistent with the market pricing of the Swiss Re life catastrophe bonds when using a market price of mortality catastrophe risk of 0.45.

3.7.6. Wang (2002) shows the transform can be adapted to allow for uncertainty in parameter estimation from sampling error by using the Student-t distribution for the distortion cumulative distribution function in place of Φ in (46). The degrees-of-freedom parameter for the Student-t distribution is then based on the sample size. This is relevant for catastrophe risk estimated

from a small sample of past events where the risk is assumed to be stationary over the short-term of the catastrophe bond. This adjusted transform is not appropriate to capture parameter uncertainty due to mortality changing over time in the case of a longevity bond. It is possible to generalise to a multi-dimensional market price of risk and couple any distortion cumulative distribution function with any parametric or empirical cumulative distribution function for the payoff F to create a very general risk-pricing mechanism for incomplete market risks.

3.7.7. Dowd et al. (2006) analyse the sensitivity of the mark-to-market value of a survivor swap to the market price of risk derived by Lin & Cox. They conclude that the value of the swap is far more sensitive to the mortality projection than to the market price of risk within the range $\lambda \in [0, 0.25]$. A similar conclusion holds for nominal pension liabilities, which are equivalent to the survivorship-linked leg of the survivor swap. For inflation-linked liabilities there will be an even greater sensitivity to the mortality projection and while the market price of risk will be different to nominal liabilities, it is also likely to be less significant than the risk inherent in the mortality projection.

3.7.8. Cairns, Blake & Dowd (2005) use the proposed pricing for the EIB survivor bond to estimate the market prices of risk for the two random factors driving their stochastic mortality curve given by equation (45). They assume that longevity risk is independent of investment risk, the mortality projection used in the proposed pricing corresponds to the real-world payout measure and that the resulting spread of 20 basis points over existing EIB debt is fully accounted for by the longevity market price of risk. The effect of Bayesian parameter uncertainty for the drifts of $A_1(t)$ and $A_2(t)$ is considered and found to be minor in comparison with the expected drifts for the price of a nominal 25-year annuity. Again, a real annuity would be more sensitive to this parameter uncertainty as inflation accounts for half the uncertainty in the survivor index 25 years ahead.

3.7.9. Cairns, Blake & Dowd (2005) note the apparent anomaly between the positive market price of longevity risk implied by the EIB survivor bond and the positive market price of mortality risk implied by the Swiss Re catastrophe bond. They reason that the fundamental difference between the nature of the mortality risk of insureds and annuitants and the transaction costs of arbitrage between the prices of these risks can account for this anomaly. It may therefore be advisable to include explicitly catastrophe risk and its market price of risk in the risk-adjusted pricing model used for both annuities and life insurance. A cause-of-death model would also distinguish sufficiently between

the mortality risk of insureds and annuitants to allow for the market prices of risk to have different signs. This would be complex and probably spurious given the uncertainty about the general level of mortality improvement over the term of the pension liabilities.

4. Credit risk

4.0.1. Until now we have used the asset irrelevance logic of Exley, Mehta & Smith (1997) to value the pension liabilities as derivatives on inflation and survivor indices without considering the assets of the fund. However, this logic does not extend to credit risk and other second-order effects.

4.0.2. DB pensions in payment are life annuities written by the fund. Since the fund sponsor is responsible for ensuring the fund has sufficient assets to meet these liabilities, these liabilities can be also viewed as life annuities written by the sponsor and collateralised with the fund assets. The value of these liabilities is therefore subject to sponsor covenant risk which, following Gordon et al. (2005), may be defined as the willingness and ability to pay contributions sufficiently in advance to ensure that the pensions can be paid as they fall due.

4.0.3. The market value of the pension fund's assets at any time is then:

- the market value of the assets held by the scheme; plus
- the market value of the sponsor covenant.

In the absence of a complete market for pension liabilities, there will always be a positive probability of a deficit at some future date, so the market value of the liabilities must be considered in relation to these two fund assets. The degree to which the sponsor covenant is material to the liabilities will depend on the relationship between the fund assets and covenant. In the (hypothetical) case of 100% self-investment of fund assets there is full reliance on the covenant. The covenant is less significant when assets match liabilities and the unfunded liabilities are uncorrelated or negatively correlated with the value of the covenant.

4.0.4. The sponsor covenant will be less significant to the value of liabilities in respect of pensions in payment if pensions in payment have a prior claim on fund assets on windup. Also, pensions in payment are less exposed to the uncertain future quality of the sponsor covenant compared with active members accruing additional DB benefits or deferred members.

4.0.5. Leaving aside the subjective 'willingness to pay' element of the covenant, the credit derivatives market can be used to hedge and hence price the 'ability to pay' element. Single-name credit default swaps (CDS) are available on a range of names, including government and municipal, and may be available on

the sponsor or another legal entity a close correlation to the sponsor's credit risk.

4.0.6. The notional for the CDS protection is proportional to the future deficit process until the pension liabilities are extinguished. The deficit process is not easily modelled since the funding policy will change in line with the size of the future deficit. The notional should therefore be based on the expected funding level at each future date. This expected funding level and other sponsor covenant risk parameters such as the asset–liability mismatch and risk from correlation of assets with the credit risk of the sponsor will be the result of negotiation between the sponsor and the pensioners (and their trustees). However, it should be recognised that there will probably be a positive relationship between the credit spread and the absolute size of the deficit and so a higher future credit spread will be warranted. This is analogous to the observation that a positive correlation between longevity and real interest rate risk will increase the market value of the liabilities unless offset by a negative market price of mortality risk. The default recovery rate allowed for should reflect the rights of pensioners as members of the fund and the rights of the fund as a creditor of the sponsor (as specified by pension law).

4.0.7. The nominal curve used to discount the projected pension liabilities should therefore be based on the risk-free forward rate for each future year plus the credit spread on the expected deficit (net of recovery on default) for the year to allow for the credit risk of the liability. The market value of pension liability calculated in this way is consistent with the value of the liability if it were to trade in the market or the fair value for a bulk transfer between pension funds. The sponsor would be indifferent between paying this market value to defease the pension liability or raising debt to finance the cost of hedging the liabilities. The pensioners (and their trustees) would agree to this transfer of the liability to a third party posing an identical credit risk.

5. Regulatory and tax effects

Cardinale et al. (2006) consider changes to pension regulation or taxation to be background risks. The effect of these risks on the market value of the pension liability is inherently unquantifiable. Examples are legal changes to the indexation of pensions in payment and deferment, interaction with the state pension, changes in dividend tax credits (as in the UK in 1997) or taxes on fund assets (such as the SA Tax on Retirement Funds), imposition of a pension benefit guarantee fund and levy, changes to pension solvency requirements, etc.

6. Conclusion

6.0.1. This research has shown how the market value of pension liabilities can be determined using models for key investment and mortality risk factors and the principle of risk-adjusted pricing. The path-dependent nature of pension indexation and the complexity of feasible interest and inflation rate and mortality models necessitates Monte Carlo simulation. The liabilities typically extend beyond 30 years, but annual time steps will often suffice to capture the essential features of pension indexation and decrements. The regular cashflow profile of typical pension liabilities means the valuation is not unduly sensitive to incomplete aspects of the market such as the indexation lag, seasonality and curve extension to ultralong maturities.

6.0.2. The limited hedging capacity of the inflation and reinsurance markets has prompted questions about their relevance for valuation. The valuation can be defended if it represents a mid price at which a marginal buyer and seller would transact. Like any asset or liability, the hedge cost of a large pension liability will incur a dealing spread to mid commensurate with the market impact of the transaction. The fundamental aim of the valuation is to anticipate the market price for the liability if an active market were to develop. This is a similar principle to extending an interest rate curve to value liabilities beyond the maturity of the longest bonds or swaps. The principle of no arbitrage and the use of a reasonable and regular market price of risk (and hence market efficiency) help to ensure the valuation is representative of the market if and when a market materialises.

6.0.3. It may therefore be preferable to model mortality and investment risk using separate models with dependent assumptions. Regardless of the form of the model, the key challenge is to set the market price of mortality risk consistently with the sparse prices across disparate markets such as the annuity (re)insurance, mortality derivative and life securitisation markets. Each of these markets has its own specific mortality risk factors and underwriting. In the SA and the UK, the growing pension buy out market also conveys information about the market price of the other background risks. The market prices of risk extracted from these various sources will be sensitive to the assumptions about profit and capital margins.

6.0.4. There will always be incomplete aspects to the market for a complex financial instrument, such as pension liability. The distinction between insurance and capital markets is becoming blurred. The market for pension risks is rapidly becoming deeper, more competitive and more transparent. The risk-adjusted valuation approach in turn becomes more relevant and robust. The

synthesis of actuarial and financial economics valuation principles will continue to extend the limits of what markets consider tradable risks.

APPENDIX A

Specimen real annuity

Nature of the liability

To evaluate how material the various valuation parameters are for inflation-linked pensions, we consider a pension in payment to a 60-year male single life. The pension is payable monthly and indexed annually at 100% of the annual inflation increase with no floor (type-1 LPI) or with an annual 0% floor (type-4 LPI) or 75% of the annual inflation increase with a 0% floor (type-5 LPI).

Mortality

PA(90) rated down 6 years is used as the mortality basis. This is the best-estimate basis of Rusconi (2006). Figure 35 shows the profile of real liabilities discounted with this mortality basis.

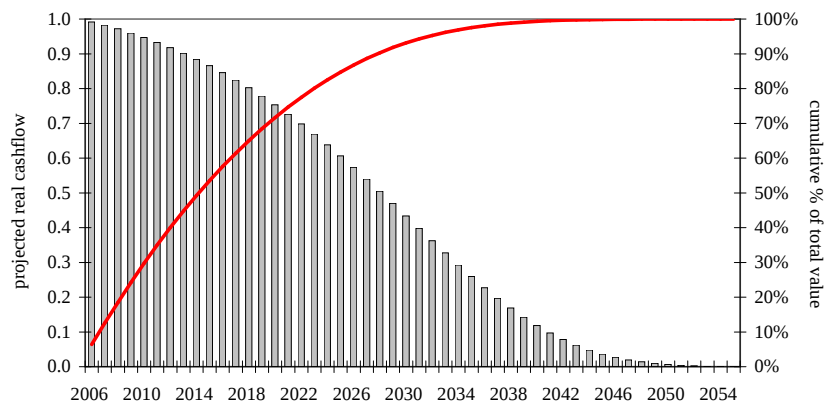


Figure 35: Real cashflows for specimen 60-year male real annuity at 26 June 2006, discounted with best-estimate mortality.

Interest and inflation rates

Liabilities are projected with breakeven inflation and discounted at the swap curve less a 45 basis point valuation margin as discussed in section 4 of chapter 3. The valuation uses the inflation swap curve derived from SA inflation-linked bonds as discussed in section 10 of chapter 4 and shown in Figure 36. The valuation date was chosen because market data were available to the author on this date, rather than because of any particular interest and inflation-curve characteristics.

BEASSA perfect-fit curves				linear forward breakeven inflation curve fitted in Chapter 4			
term	date	zero-coupon swap curve	zero-coupon bond curve	term	date	4-month lag forward CPI	zero-coupon breakeven
		quarterly rate	semiannual rate				annual rate
0.00	26 Jun 06	7.227%	7.292%	0	26 Jun 06	131.083	
0.00	27 Jun 06	7.233%	7.299%	1	26 Jun 07	139.840	6.680%
0.01	29 Jun 06	7.246%	7.311%	2	26 Jun 08	148.864	6.566%
0.02	03 Jul 06	7.269%	7.328%	3	26 Jun 09	158.234	6.476%
0.08	26 Jul 06	7.302%	7.256%	4	26 Jun 10	167.946	6.391%
0.17	28 Aug 06	7.349%	7.213%	5	26 Jun 11	177.995	6.309%
0.25	26 Sep 06	7.570%	7.313%	6	26 Jun 12	188.396	6.232%
0.50	27 Dec 06	7.911%	7.666%	7	26 Jun 13	199.105	6.153%
0.75	26 Mar 07	8.261%	8.028%	8	26 Jun 14	210.349	6.090%
1.00	26 Jun 07	8.527%	8.294%	9	26 Jun 15	222.246	6.042%
1.25	26 Sep 07	8.714%	8.499%	10	26 Jun 16	234.863	6.005%
1.50	27 Dec 07	8.855%	8.708%	11	26 Jun 17	248.189	5.975%
1.75	26 Mar 08	8.961%	8.809%	12	26 Jun 18	262.285	5.950%
2	26 Jun 08	9.049%	8.847%	13	26 Jun 19	277.203	5.930%
3	26 Jun 09	9.237%	8.834%	14	26 Jun 20	293.029	5.914%
4	28 Jun 10	9.305%	8.680%	15	26 Jun 21	309.750	5.900%
5	27 Jun 11	9.333%	8.621%	16	26 Jun 22	327.442	5.889%
6	26 Jun 12	9.327%	8.706%	17	26 Jun 23	346.170	5.879%
7	26 Jun 13	9.312%	8.790%	18	26 Jun 24	365.850	5.868%
8	26 Jun 14	9.267%	8.869%	19	26 Jun 25	385.576	5.843%
9	26 Jun 15	9.210%	8.904%	20	26 Jun 26	405.170	5.805%
10	27 Jun 16	9.195%	8.907%	21	26 Jun 27	424.519	5.755%
11	26 Jun 17	9.165%	8.902%	22	26 Jun 28	443.537	5.697%
12	26 Jun 18	9.120%	8.855%	23	26 Jun 29	462.018	5.630%
13	26 Jun 19	9.070%	8.803%	24	26 Jun 30	479.856	5.556%
14	26 Jun 20	9.013%	8.735%	25	26 Jun 31	496.929	5.475%
15	28 Jun 21	8.948%	8.657%	26	26 Jun 32	513.146	5.389%
16	27 Jun 22	8.876%	8.575%	27	26 Jun 33	528.313	5.298%
17	26 Jun 23	8.798%	8.493%	28	26 Jun 34	542.614	5.204%
18	26 Jun 24	8.717%	8.412%	29	26 Jun 35	557.167	5.116%
19	26 Jun 25	8.635%	8.336%	30	26 Jun 36	572.146	5.035%
20	26 Jun 26	8.553%	8.265%	BESA closing rates, 26 Jun 2006			
21	28 Jun 27	8.472%	8.199%	Nominal bonds			
22	26 Jun 28	8.394%	8.140%	10% 2008	8.730%		
23	26 Jun 29	8.320%	8.085%	13% 2010	8.740%		
24	26 Jun 30	8.252%	8.034%	13.5% 2015	8.840%		
25	26 Jun 31	8.189%	7.988%	10.5% 2026	8.545%		
26	28 Jun 32	8.131%	7.944%	Inflation-linked bonds			
27	27 Jun 33	8.079%	7.905%	3.8% 2008	2.720%		
28	26 Jun 34	8.031%	7.868%	6.25% 2013	2.700%		
29	26 Jun 35	7.987%	7.834%	5.50% 2023	2.670%		
30	26 Jun 36	7.946%	7.802%	3.45% 2033	2.645%		

Figure 36: Market rates for SA bond and swap market, 26 June 2006

Bibliography

- Albrecht, P. (1991), ‘Financial approach to actuarial risks?’, *AFIR* **2**, 227–247.
- Amin, K. & Jarrow, R. (1991), ‘Pricing foreign currency options under stochastic interest rates’, *Journal of International Money and Finance* *10*, 3 **3**, 301–329.
- Anderson, N., Breedon, F., Deacon, M., Derry, A. & Murphy, G. (1996), *Estimating and Interpreting the Yield Curve*, Wiley Series in Financial Economics and Quantitative Analysis.
- Anderson, N. & Sleath, J. (2001), New estimates of the UK real and nominal yield curves. Working paper, www.bankofengland.co.uk/workingpapers/index.htm.
- Ang, A., Bekaert, G. & Wei, M. (2005), Do macro variables, asset markets or surveys forecast inflation better? Working paper, www.nber.org/papers/w11538.
- Aron, J. & Muellbauer, J. (2006), Review of Monetary Policy in South Africa since 1994. CSAE Working Paper Series 2006-07, <http://www.csae.ox.ac.uk/>.
- Babbel, D. F., Gold, J. & Merrill, C. B. (2002), ‘Fair value of liabilities the financial economics perspective’, *North American Actuarial Journal* **6**(1), 12–27.
- Ballotta, L. & Haberman, S. (2004), ‘Guaranteed annuity conversion options and their valuation’. In Fornero, E. and Luciano, E. (eds) *Developing an Annuity Market in Europe*. Edward Elgar Publishing, 167–198.
- Beletski, T. (2006), Inflation-linked products and optimal investment with macro derivatives. Unpublished thesis, <http://kluedo.ub.uni-kl.de/volltexte/2006/2045/index.html>.
- Belgrade, N. & Benhamou, E. (2004a), Impact of seasonality in inflation derivatives pricing. CDC Ixis Quantitative Research working paper, <http://papers.ssrn.com/abstractid=583642>.
- Belgrade, N. & Benhamou, E. (2004b), Reconciling Year on Year and Zero Coupon Inflation Swap: A Market Model Approach. Unpublished research paper, <http://ssrn.com/abstract=583641>.
- Belgrade, N., Benhamou, E. & Khelif, Y. (2004), Valuation of inflation swap volatility under a market model and pricing of real yield options. CDC Ixis Quantitative Research working paper, <http://papers.ssrn.com/abstractid=605608>.
- Belgrade, N., Benhamou, E. & Koehler, E. (2005), ‘Modelling inflation in finance’. In Benaben, B. (ed) *Inflation-linked Products: A Guide for Investors and Asset and Liability Managers*, Risk Books.
- Black, F. (1980), The tax advantages of pension fund investments in bonds. Working paper, NBER Working Paper No. W0533, National Bureau of Economic Research, <http://ideas.repec.org/p/nbr/nberwo/0533.html>.
- Black, F. & Scholes, M. (1973), ‘The pricing of options and corporate liabilities’, *Journal of Political Economy* **81**, 637–659.

- Blake, D., Cairns, A. J. G. & Dowd, K. (2006), 'Living with mortality - longevity bonds and other mortality-linked securities', *British Actuarial Journal* **12**, 153–197.
- Broadie, M., Glasserman, P. & Kou, S. G. (1999), 'Connecting discrete and continuous path dependent options', *Finance and Stochastics* **3**, 55–82.
- Brown, R. & Schaefer, S. (2000), Why Long Term Forward Interest Rates (Almost) Always Slope Downwards. Faculty research paper, London Business School, www.facultyresearch.london.edu/docs/299new.pdf.
- Buhlmann, H. (1987), 'Actuaries of the Third Kind', *ASTIN Bulletin (editorial)* **17**(2), 137–138.
- Cairns, A. J. G. (1998), 'Descriptive bond-yield and forward-rate models for the British government securities market', *British Actuarial Journal* **4**(II), 265–321.
- Cairns, A. J. G. (2001), 'From Financial Economics to Fair Valuation', *AFIR* **11**(1), 135–166.
- Cairns, A. J. G. (2004a), 'A family of term-structure models for long-term risk management and derivative pricing', *Mathematical Finance* **14**(3), 415–444.
- Cairns, A. J. G. (2004b), *Interest Rate Models An Introduction*, Princeton University Press.
- Cairns, A. J. G. (2006), A Generalisation of the Smith-Olivier Model for Stochastic Mortality. Seminar presentation, Heriot-Watt University, www.ma.hw.ac.uk/~andrewc/papers/.
- Cairns, A. J. G., Blake, D., Dawson, P. & Dowd, K. (2005), 'Pricing Risk on Longevity Bonds', *Life and Pensions, October 2005* pp. 41–44.
- Cairns, A. J. G., Blake, D. & Dowd, K. (2005), 'A Two-Factor Model for Stochastic Mortality with Parameter Uncertainty: Theory and Calibration', *Journal of Risk and Insurance* **73**, 687–718.
- Cairns, A. J. G., Blake, D. & Dowd, K. (2006), 'Pricing Death: Frameworks for the Valuation and Securitization of Mortality Risk', *ASTIN Bulletin* **36**, 79–120.
- Campbell, J. Y., Lo, A. W. & MacKinlay, A. C. (1997), *The Econometrics of Financial Markets*, Princeton University Press.
- Cardinale, M., Katz, G., Kumar, J. & Orszag, J. M. (2006), 'Background risk and pensions', *British Actuarial Journal* **12**(1), 79–134.
- Chan, W., Ng, M. & Tong, H. (2006), 'On a simple graphical approach to modelling economic fluctuations with an application to United Kingdom price fluctuation 1265 to 2005', *Annals of Actuarial Science* **1**(1), 103–128.
- Chapman, R., Gordon, T. & Speed, C. (2001), 'Pensions, funding and risk', *British Actuarial Journal* **7**(4), 605–662.
- Clark, G. (1992), 'Asset and liability modelling - the way ahead?', *Presented to the Staple Inn Actuarial Society*.
- CMI (2004), Continuous Mortality Investigation sub-committee, Working Paper 3, Projecting future mortality - a discussion paper. The Faculty of Actuaries and Institute of Actuaries.
- CMI (2006), Continuous Mortality Investigation sub-committee, Working Paper 15, Projecting future mortality - towards a proposal for stochastic mortality. The Faculty of Actuaries and Institute of Actuaries.
- Cochrane, J. H. (2001), *Asset Pricing*, Princeton University Press.
- Cortes, F. (2003), 'Understanding and modelling swap spreads', *Bank of England Quarterly Bulletin, Winter 2003* pp. 407–416.
- Cowley, A. & Cummins, J. D. (2005), 'Securitisation of Life Insurance Assets and Liabilities', *Journal of risk and insurance* **72**(2), 193–226.

- Cox, S. H., Lin, Y. & Wang, S. (2005), Mortality bond pricing. Presentation, First International Conference On Longevity Risk And Capital Market Solutions.
- Currie, I. D., Durban, M. & Eilers, P. H. C. (2004), 'Smoothing and forecasting mortality rates', *Statistical Modelling* **3**(4), 279–298.
- Dahl, M. H. (2004), 'Stochastic Mortality in Life Insurance - Market Reserves and Mortality-linked Insurance Contracts', *Insurance - Mathematics and Economics* **35**(1), 113–136.
- De Gooijer, J. & Vidiella-i-Angeura, A. (2003), 'Non-linear stochastic inflation modelling using seasetars', *Insurance: Mathematics and Economics* **32**(1), 3–18.
- Deacon, M., Derry, A. & Mirfendereski, D. (2004), *Inflation-indexed securities, 2nd edition*, Wiley Finance.
- Dowd, K., Blake, D., Cairns, A. J. G. & Dawson, P. (2006), 'Survivor Swaps', *Journal of Risk and Insurance* **73**(1), 1–17.
- Duffie, D. (2001), *Dynamic Asset Pricing Theory*, Third edition, Princeton University Press.
- Duffie, D. & Richardson, H. R. (1991), 'Mean-variance hedging in continuous time', *Annals of Applied Probability* **1**, 1–15.
- Duffie, D. & Singleton, K. (1997), 'An econometric model of the term structure of interest-rate swap yields', *The Journal of Finance* **52**(4), 1287–1321.
- Dybvig, P., Ingersoll, J. & Ross, S. (1996), 'Long forward rates and zero coupon rates can never fall', *Journal of Business* **69**, 1–25.
- Elder, R., Kapetanios, G., Taylor, T. & Yates, T. (2005), 'Assessing the MPCs fan charts', *Bank of England Quarterly Bulletin, August 2005*.
- Embrechts, P. (1996), Actuarial versus financial pricing of insurance. Working Paper 96-17, Wharton School Center for Financial Institutions, University of Pennsylvania, www.math.ethz.ch/~embrechts.
- Evans, M. D. D. (1998), 'Real rates, Expected Inflation and Inflation Risk Premia', *Journal of Finance* **53**(1), 187–218.
- Evans, M. D. D. (2003), 'Real Risk, Inflation Risk and the Term Structure', *The Economic Journal* **113**(487), 345–389.
- Exley, C. J. (2006), 'Calibrating inflation', *Life and Pensions Magazine, March 2006*.
- Exley, C. J., Mehta, S. J. B. & Smith, A. D. (1997), 'The financial theory of defined benefit pension schemes', *British Actuarial Journal* **3**(4), 835–966.
- Fama, E. (1970), 'Efficient capital markets - a review of theory and empirical work', *Journal of Finance* **25**(2), 383–417.
- Feldman, K. et al. (1998), 'Report of the Fixed Income Working Group', *British Actuarial Journal* **4**(2), 350–383.
- Fisher, I. (1930), *The Theory of Interest*, Augustus M. Kelly, New York.
- Fisher, L. & Weil, R. (1971), 'Coping with the risk of interest rate risk fluctuations - returns to bond holders from a naive and optimal strategy', *Journal of Business* **44**(4), 408–431.
- Fisher, M., Nychka, D. & Zervos, D. (1995), 'Fitting the term structure of interest rates with smoothing splines', *Finance and Economics Discussion Series, 95-1 97-10*.
- Flavell, R. (2002), *Swaps and other derivatives*, Wiley Finance.
- Flesaker, B. & Hughston, L. (1996), 'Positive interest', *Risk* **9**(1), 46 – 49.
- Fuller, W. (1976), *Introduction to statistical time series*, Wiley, New York.
- Garman, M. B. & Kohlhagen, S. W. (1983), 'Foreign currency option values', *Journal of International Money and Finance* **2**, 231 – 237.

- Goldman, M. B., Sosin, H. & Gatto, M. (1979), 'Path dependent options - buy at the low, sell at the high', *Journal of Finance* **34**(5), 1111–1127.
- Golub, B. W. & Tilman, L. M. (2000), *Risk management - approaches for fixed income markets*, Wiley Finance.
- Gordon, T. J. (1999), The price of actuarial values. Presented to the Staple Inn Actuarial Society.
- Gordon, T. J., Evans, A., Freeman, G., Forrester, N., Hall, R., McKay, N. & Shellswell, P. (2005), Allowing for the Sponsor Covenant in Actuarial Advice. The Actuarial Profession - Sponsor Covenant Working Party Final Report.
- Griffiths, J., Imam-Sadeque, F., Ong, A. & Smith, A. D. (1997), Actuaries take note, the yield curve is not flat. Presented to the Joint Institute and Faculty of Actuaries Investment Conference, Edinburgh, 1997.
- Grossman, S. & Stiglitz, J. (1980), 'On the impossibility of informationally efficient markets', *American Economic Review* **70**.
- Gurkaynak, R. S., Sack, B. P. & Wright, J. H. (2006), The U.S. Treasury Yield Curve - 1961 to the Present . The Federal Reserve Board Finance and Economics Discussion Series: 2006-28, <https://www.federalreserve.gov/PUBS/feds/2006/200628/>.
- Haberman, S. & Verrall, R. (2006), Modelling and Measuring Uncertainty in Mortality Forecasts and Lee Carter. Presented to the Institute and Faculty of Actuaries Mortality and Longevity Seminar, May 2007.
- Harrison, J. M. & Kreps, D. M. (1979), 'Martingales and arbitrage in mutiperiod securities markets', *Journal of Economic Theory* **20**(3), 381–408.
- Harrison, J. & Pliska, S. (1981), 'Martingales and stochastic integrals in the theory of continuous trading', *Stochastic Processes and their Applications* **11**, 215–260.
- Head, S. J., Adkins, D., Cairns, A. J. G., Corvesor, A., Cule, D., Exley, J., Johnson, I., Spain, J. & Wise, A. (2000), 'Pension Fund Valuations and Market Values', *British Actuarial Journal* **6**(1), 42–143.
- Heath, D., Jarrow, R. & Morton, A. (1992), 'Bond pricing and the term structure of interest rates a new methodology for contingent claims valuation', *Econometrica* **60**(1), 77 – 105.
- Henrard, M. (2005), Inflation Bond Option Pricing in Jarrow-Yildirim Model. Working paper, [129.3.20.41/econ-wp/fin/papers/0510/0510027.pdf](https://www.federalreserve.gov/PUBS/feds/2006/200628/).
- Heston, S. L. (1993), 'A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options', *The Review of Financial Studies* **6**(2), 327–343.
- Huang, H.-C. & Cairns, A. J. G. (2004), 'Valuation and hedging of LPI liabilities', *British Actuarial Journal* **10**, 627–663.
- Huber, P. P. (1997), 'A Review of Wilkie's Stochastic Asset Model', *British Actuarial Journal* **3**(1), 181–210.
- Hull, J. C. (2003), *Options, futures and other derivatives*, 5th edition, Prentice Hall.
- Hull, J. & White, A. (1993), 'One factor interest rate models and the valuation of interest rate derivative securities', *Journal of Financial and Quantitative Analysis* **28**(2), 235 – 254.
- Ilmanen, A. & Byrne, R. (2003), 'The impact of convexity differences on the interest-rate swap curve', *Journal of Bond Trading and Management* **1**(4), 305–314.

- IMF (2005), 'Does inflation targeting work in emerging markets?', *World Economic Outlook* **2**, Ch 4, 161–186.
- Institute and Faculty of Actuaries (1998), 'Working party report on market based discount rates for pension cost accounting'.
- James, J. & Webber, N. (2000), *Interest Rate Modelling*, Wiley Financial Engineering.
- Jarrow, R. & Yildirim, Y. (2003), 'Pricing Treasury Inflation Protected Securities and Related Derivatives using an HJM Model', *Journal of Financial and Quantitative Analysis* **38**(2), 337–359.
- Jarvis, S., Southall, F. & Varnell, E. (2001), Modern valuation techniques. Paper presented to Staple Inn Actuarial Society, 2001.
- Jensen, B. A. & Nielsen, J. A. (1995), 'Pricing by no arbitrage'. In D R. Cox, D. Hinkley and O E. Barndorff-Nielsen (eds) *Likelihood, Time series with Econometric and other Applications*, Chapman and Hall.
- Kamal, M. & Derman, E. (1999), 'Correcting Black Scholes', *Risk* **12**(1), 82–85.
- Kemp, M. H. D. (1997), 'Actuaries and derivatives', *British Actuarial Journal* **3**(1), 51–180.
- Kendal, S. & Franklin, M. D. (1993), 'Some pension fund topics', *Transactions of the Actuarial Society of South Africa* **9**(3), 34–85.
- Klein, R. (2003), A matter of life and death. The Review, Special Report - Mortality Risk, March 2003.
- Lee, P. J. & Wilkie, A. D. (2000), 'A comparison of stochastic asset models', *AFIR* **10**, 407–445.
- Lee, R. D. & Carter, L. R. (1992), 'Modelling and forecasting U.S. mortality', *Journal of the American Statistical Association* **87**(419), 659–671.
- Leland, H. R. (1985), 'Option pricing and replication with transactions costs', *Journal of Finance* **40**(5), 1283–1301.
- Lin, Y. & Cox, S. H. (2005), 'Securitization of mortality risks in life annuities', *Journal of Risk and Insurance*, vol. **72**(2), 227–252.
- Lin, Y. & Cox, S. H. (2007), 'Natural Hedging of Life and Annuity Mortality Risks', *North American Actuarial Journal* **11**(3), 1–15.
- Lioui, A. & Poncet, P. (2005), 'General equilibrium pricing of CPI derivatives', *Journal of Banking & Finance* **29**(5), 1265–1294.
- Lipsey, R., Steiner, P. & Purvis, D. (1987), *Economics, 8th edition*, Harper and Row.
- Lucas, R. E. & Sargent, T. J. (1978), "New" explanations of the persistence of inflation and unemployment'. In *After the Phillips Curve: Persistence of High Inflation and High Unemployment*, Federal Reserve Bank of Boston Conference Series No. 19, pp. 49–72.
- Maitland, A. J. (1996), 'A review of Thomson's Stochastic Investment Model', *Transactions of the Actuarial Society of South Africa* **11**(2), 437–450.
- Maitland, A. J. (2001), 'An Empirical Approach to Immunisation in South Africa', *South African Actuarial Journal* **1**, 119–138.
- Maitland, A. J. (2002), 'Interpolating the South African Yield Curve using Principal Components Analysis - a descriptive approach', *South African Actuarial Journal* **2**, 129 – 146.
- Malvaez, L. (2005), Valuation of Inflation-Indexed Derivatives with three factor model. Masters Thesis, University of Oxford, eprints.maths.ox.ac.uk.

- McCulloch, H. J. (2000), Long Forward and Zero-Coupon Rates Indeed Can Never Fall, But Are Indeterminate - a Comment on Dybvig, Ingersoll and Ross. Ohio State University Economics Dept. Working Paper No. 00-12.
- McCulloch, H. J. & Kochin, L. A. (1998), The Inflation Premium Implicit in the US Real and Nominal Term Structures of Interest Rates. Ohio State University Economics Dept. Working Paper No. 98-12, revised September 2000.
- McNeil, A., Frey, R. & Embrechts, P. (2006), *Quantitative Risk Management*, Princeton Series in Finance.
- Mercurio, F. (2005a), 'Pricing inflation-indexed derivatives', *Quantitative Finance* **3**(5), 289–302.
- Mercurio, F. (2005b), Pricing Inflation-Indexed Derivatives with Stochastic Volatility. Presentation to The Second Fixed Income Conference, Prague.
- Mercurio, F. & Moreni, N. (2006), 'Inflation with a smile', *Risk* **19**(3), 70–75.
- Merton, R. C. (1973), 'Theory of rational option pricing', *Bell Journal of Economics and Management Science* **4**(1), 141–183.
- Milevsky, M. A. & Promislow, S. D. (2001), 'Mortality derivatives and the option to annuitise', *Insurance Mathematics and Economics* **29**, 299–318.
- Miltersen, K. R. & Persson, S.-A. (2005), Is mortality dead? Stochastic forward force of mortality rates determined by no arbitrage. Working Paper, University of Bergen.
- Modigliani, F. & Miller, M. H. (1958), 'The cost of capital, corporate finance and the theory of investment', *American Economic Review* **48**, 261–297.
- Moller, T. (2002), 'On valuation and risk management at the interface of insurance and finance', *British Actuarial Journal* **8**(4), 787–827.
- Musiela, M. & Rutkowski, M. (1997), *Martingale Methods in Financial Modelling*, Springer Verlag.
- Mussche, P. (2002), 'Drivers of dollar swap spreads', *Risk* **15**(3), 86–88.
- Nelson, C. R. & Siegel, A. F. (1987), 'Parsimonious modelling of yield curves', *Journal of Business* **60**(4), 473–489.
- Olivier, P. & Jeffery, T. (2004), Stochastic Mortality Models. Presentation to The Society of Actuaries of Ireland, October 2004.
- Olivier, P. & Smith, A. (2007), Olivier-Smith Model. Presented to the Institute and Faculty of Actuaries Mortality and Longevity Seminar, May 2007.
- Olivieri, A. (2001), 'Uncertainty in mortality projections - an actuarial perspective', *Insurance Mathematics and Economics* **29**.
- Olivieri, A. (2002), Risk and reinsurance in life annuity portfolios. University of Trieste, Department of Applied Mathematics, publication n. 3/2002.
- Ortiz, C. E., Stone, C. A. & Zissu, A. (2006), Securitized Senior Life Settlements' Macauley Duration and Longevity Risk. Presentation at the Life Insurance Settlement Association's 12th Annual Fall Conference, New York, NY, November 29-December 1, 2006.
- Palin, J. & Speed, C. A. (2003), Hedging Pension Plan Funding Ratio. Presented to the Society of Actuaries Symposium 'The Great Controversy - Current Pension Actuarial Practice in Light of Financial Economics', June 2003.
- Phillips, A. (1960), 'The relation between unemployment and the rate of change of money wage rates in the united kingdom, 1861-1957', *Economica* **27**, 1–31.

- Rebonato, R. (1998), *Interest Rate Option Pricing Models, 2nd ed.*, Wiley Financial Engineering.
- Rebonato, R. (2002), *Modern Pricing of Interest Rate Derivatives*, Princeton University Press.
- Redington, F. (1952), 'Review of the principles of life-office valuations', *Journal of the Institute of Actuaries* **78**, 286–340.
- Reitano, R. R. (1997), 'Two paradigms for the market value of liabilities', *North American Actuarial Journal* **1**(4), 104–122.
- Richards, S. & Jones, G. (2004), Financial aspects of longevity risk. Presented to the Staple Inn Actuarial Society, 2004.
- Rogers, L. C. G. (1997), 'The Potential Approach to the Term Structure of Interest Rates and FX Rates', *Mathematical Finance* **7**, 157–176.
- Rusconi, R. (2006), Non profit annuities - thoughts on South Africa. Presented to Actuarial Society of South Africa sessional meeting, 21 February 2006.
- Sack, B. P. (2000), Using Treasury STRIPS to Measure the Yield Curve. Federal Reserve Bank research paper, www.federalreserve.gov/pubs/feds/2000/.
- Smith, A. D. (1996), 'How actuaries can use financial economics', *British Actuarial Journal* **2**(5), 1057–1193.
- Smith, A. D. (1999), A review of models for pricing caps and floors on inflation over the long term. Presented to the Joint Institute and Faculty of Actuaries Investment Conference, Loch Lomond 1999.
- Smith, A. D. (2001a), Investigating some measures of market efficiency. Institute and Faculty of Actuaries Finance and Investment Conference, Guernsey 2001.
- Smith, A. D. (2001b), Pricing beyond the curve. Presentation to Centre for Financial Research, Judge Institute of Management Studies, Cambridge University, 2001.
- Smith, A. D. (2003), The Smith Model. B&W Deloitte, www.thesmithmodel.com.
- Smith, A. D. & Wilson, T. (2000), Fitting yield curves with long term constraints. Working paper, Bacon and Woodrow.
- Stander, Y. (2005), *Yield Curve Modelling*, Palgrave Macmillan.
- Svensson, L. (1995), 'Estimating forward interest rates with the extended Nelson and Siegel method', *Sveriges Riksbank Quarterly Review* **3**, 13–26.
- Thomson, R. J. (1994), 'A Stochastic Investment Model for actuarial use in South Africa', *Transactions of the Actuarial Society of South Africa* **10**, 72–133.
- Thomson, R. J. (1996), 'Stochastic investment models: the case of South Africa', *British Actuarial Journal* **2**(3), 765–801.
- Thomson, R. J. (2002), 'Editorial', *South African Actuarial Journal* **2**, 147–150.
- Thomson, R. J. (2005), 'The pricing of liabilities in an incomplete market using dynamic mean-variance hedging.', *Insurance: Mathematics and Economics* **36**, 441–455.
- Thomson, R. J. & Gott, D. V. (2007), 'Stochastic models for actuarial use: the equilibrium modelling of local markets', *Proceedings of the 16th AFIR Colloquium (forthcoming)*.
- Tilley, J. A. & Mueller, M. (1991), 'Managing interest rate risk for long liabilities', *AFIR* **1**, 169–184.
- Treynor, J. (1977), 'The principles of corporate pension finance', *Journal of Finance* **32**(2), 627–638.

- UK DMO (2000), The DMO's Yield Curve Model, United Kingdom Debt Management Office.
UK DMO research papers, www.dmo.gov.uk/gilts/public/research/.
- Van Bezoooyen, J., Mehta, S. & Smith, A. (1999), Supply and Demand Imbalances in the UK Government Bond Market. Presented at the Joint Institute and Faculty of Actuaries Investment Conference, 1999.
- Van Bezoooyen, J. T. S., Exley, C. J. & Smith, A. D. (1997), A market based approach to valuing LPI liabilities. Presented to the Joint Institute and Faculty of Actuaries Investment Conference, Edinburgh 1997.
- Wadsworth, M. (2005), The pension annuity market - further research into supply and constraints. Report for the Association of British Insurers, 2005.
- Waggoner, D. (1997), Spline methods for extracting interest rate curve from coupon bond prices. Federal Reserve Bank of Atlanta Working Paper Series, 97-10.
- Waisberg, N., Morris, G., Ovenden, G. & Mothapo, R. (2004), The real missing link. Presented to the Actuarial Society of South Africa Convention, 2004.
- Wang, S. S. (2002), 'A universal framework for pricing financial and insurance risks', *ASTIN Bulletin* **32**, 213–234.
- Whitten, S. & Thomas, R. (1999), 'A non-linear stochastic asset model for actuarial use', *British Actuarial Journal* **5**(5), 919–953.
- Wilkie, A. (1984), 'The cost of minimum money guarantees on index-linked annuities', *Transactions of the International Congress of Actuaries* **22**, 137–150.
- Wilkie, A. (1995), 'More on a stochastic asset model for actuarial use', *British Actuarial Journal* **1**(5), 777–964.
- Wilkie, A. D. (1986), 'A stochastic asset model for actuarial use', *Transactions of the Faculty of Actuaries* **39**, 341–373.
- Willets, R. (2006), Recent developments in mortality. Presented to the Institute and Faculty of Actuaries Mortality and Longevity Seminar, May 2007.
- Wilmott, P. (1994), 'Discrete charms', *Risk* **7**(3), 48–51.
- Wilmott, P. (1999), *Derivatives The theory and practice of financial engineering*, John Wiley & Sons, Ltd.
- Wise, A. (2004), Pensions and economics - the way ahead. Presented to the Staple Inn Actuarial Society, 2004.
- Yakoubov, Y., Teeger, M. & Duval, D. (1999), 'A stochastic investment model for asset and liability management', *AFIR* **9**, 237–266.